



HAL
open science

Le calcul économique dans le processus de choix collectif des investissements de transport

Joël Maurice, Yves Crozet, Claude Abraham, Marc Gaudry, Marc Ivaldi,
Catherine Vibes, Christian Gollier, Laetitia Andrieu, André De Palma,
Nathalie Picard, et al.

► To cite this version:

Joël Maurice, Yves Crozet, Claude Abraham, Marc Gaudry, Marc Ivaldi, et al.. Le calcul économique dans le processus de choix collectif des investissements de transport. *Economica*, 454 p., 2007, coll. Méthodes et approches, Gérard Brun, 978-2717854251. halshs-00187333

HAL Id: halshs-00187333

<https://shs.hal.science/halshs-00187333>

Submitted on 11 May 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Au mois de février 2003, un rapport sur les grands projets d'infrastructure de transport était remis au Ministère de l'Équipement et des Transports. Rédigé, à la demande du ministre, par le Conseil général des Ponts et Chaussées et l'Inspection des Finances, il passait au crible du calcul économique l'ensemble des projets de nouvelles infrastructures de transport terrestre (routes, fer, voies navigables). Ses conclusions ont été d'emblée contestées par tous ceux, notamment les élus locaux, qui voyaient dans ses résultats la remise en cause de la programmation d'infrastructures les concernant.

Le calcul économique est ainsi en partie déconsidéré alors même que s'accroissent les contraintes financières qui pèsent sur les budgets publics. Le recours à des outils d'aide à la décision, propres à rationaliser les choix publics, devrait donc se renforcer. Pour cette raison, suivant une demande du Conseil Général des Ponts et Chaussées (CGPC), les Groupes opérationnels (GO) n° 1 et n° 11 du PREDIT ont lancé un travail de recherche sur les pistes de travail qui pourraient, non pas refonder, mais enrichir le calcul économique. Comment le rendre plus apte à jouer son rôle dans le processus de choix collectif dans le champ des infrastructures de transport ?

En réponse à cette demande, un « Réseau de recherche sur les dimensions critiques du calcul économique » a été constitué. Les contributions des participants de ce groupe, qui a rassemblé près de vingt chercheurs de tous horizons, ont été regroupées dans cet ouvrage. Le tout est précédé d'un préambule rédigé par Claude Abraham, qui dispose du recul nécessaire pour rappeler que les questions posées par la présente recherche ne sont pas nouvelles !

*
**

Joël Maurice, ancien élève de l'École Polytechnique, est Professeur à l'École Nationale des Ponts et Chaussées et directeur du CERAS (Centre de Recherche et d'Enseignement en Analyse Socio-Économique), équipe membre de Paris Sciences Économiques (PSE).

Yves Crozet, est professeur à l'Université de Lyon et directeur du Laboratoire d'Économie des Transports (UMR CNRS n° 5593). Il est également président du GO 1 du Prédit « Mobilités, territoires et développement durable ».

Sous la direction de
Joël MAURICE
Yves CROZET

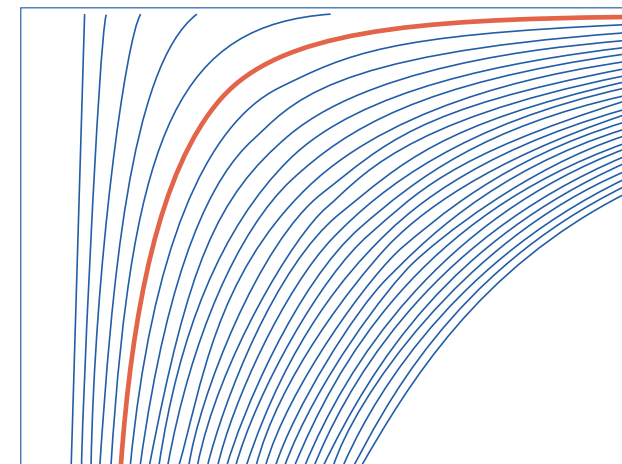
LE CALCUL ÉCONOMIQUE

Collection « Méthodes et Approches »
dirigée par Gérard Brun

LE CALCUL ÉCONOMIQUE

dans le processus de choix collectif
des investissements de transport

Sous la direction de
Joël MAURICE et Yves CROZET



Ouvrage publié avec le soutien du PREDIT,
programme de recherche sur les transports terrestres



9 782717 854251

ISBN 978-2-7178-5425-1

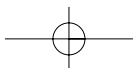
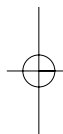
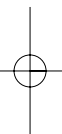
75 €



ECONOMICA



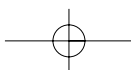
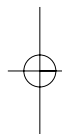
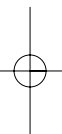
LE CALCUL ÉCONOMIQUE





Collection « Méthodes et Approches »
dirigée par Gérard BRUN

Cette collection axée sur la recherche en transport et en urbanisme a pour ambition de publier des ouvrages contribuant à un renouveau conceptuel dans le domaine des sciences humaines, par le recours à des méthodes nouvelles et à des approches transversales.



Collection « Méthodes et Approches »
dirigée par Gérard Brun

LE CALCUL ÉCONOMIQUE

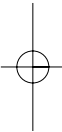
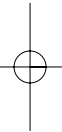
dans le processus de choix collectif
des investissements de transport

Sous la direction de
Joël MAURICE et Yves CROZET

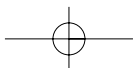
Ouvrage publié avec le soutien du PREDIT,
programme de recherche sur les transports terrestres



 **ECONOMICA**
49, rue Héricart, 75015 Paris



© Ed. ECONOMICA, 2007
Tous droits de reproduction, de traduction, d'adaptation et d'exécution
réservés pour tous les pays.



SOMMAIRE

Préambule.....	VII
Introduction générale.....	1
Première partie : La construction du calcul économique	3
Chapitre 1 : Structure de la modélisation du trafic et théorie économique	6
Chapitre 2 : Analyse Coût-Bénéfice dans un contexte de concurrence intermodale et intramodale	98
Chapitre 3 : Actualisation : prise en compte du temps dans un environnement risqué	112
Chapitre 4 : Intégration du risque et de l'incertitude dans la construction du calcul économique	128
Chapitre 5 : La mesure de l'unité sociale des investissements : l'enjeu du processus de production des valeurs tutélaires	165
Chapitre 6 : L'influence relative des différentes valeurs tutélaires : une étude par la sensibilité des indicateurs socio-économiques	191
Deuxième partie : Des projets à la programmation des investissements.....	221
Chapitre 7 : Évaluation, financement et programmation des investissements	227
Chapitre 8 : Une méthode d'optimisation des programmes d'investissements de transport	249
Chapitre 9 : Choix des projets sous contrainte budgétaire annuelle : essai de récapitulation.....	276
Troisième partie : L'aide à la décision publique	315
Chapitre 10 : De la modélisation des comportements au calcul économique : l'équité des politiques de transport.....	317
Chapitre 11 : L'analyse des projets d'infrastructure de transport dans un cadre d'équilibre général	340

Chapitre 12 : Équité, efficacité et acceptabilité dans la localisation des équipements collectifs.....	361
Chapitre 13 : Équité territoriale, acceptabilité et grandes infrastructures de transport	402
Chapitre 14 : Les figures de l'acceptabilité	432

PRÉAMBULE

Claude ABRAHAM¹

À la demande des responsables du réseau de recherche sur les dimensions critiques du calcul économique, et à la lecture de quelques-uns des travaux produits par ce groupe, c'est bien volontiers que je livre quelques remarques issues de mon expérience en la matière.

1. LES PREMIERS CALCULS DE RENTABILITÉ

C'est à la Direction des Routes, à la fin des années cinquante, que naissent, en France, les premiers calculs dits de rentabilité. Sous l'influence des économistes d'EDF (Pierre Massé, puis Marcel Boiteux), les ingénieurs de la Direction des Routes² justifient, développent et appliquent les premiers calculs de « Bénéfice actualisé ».

Pour un projet routier donné : élargissement, déviation, construction d'une route nouvelle, on évalue, année par année, les « avantages » ou « bénéfices » que procure le projet nouveau : économies de coûts de circulation, gains de temps, amélioration de la sécurité. Ces avantages sont chiffrés à leur valeur présente en actualisant les sommes futures, et l'on compare cet avantage actualisé au coût du projet. Si le « Bénéfice actualisé » est positif, le projet est réputé rentable et doit être réalisé. Comme, en général, le bénéfice actualisé à l'année zéro dépend de la date de mise en service du projet, on appelle année optimale de mise en service, celle où le bénéfice actualisé passe par un maximum, caractérisé, en première approximation, par le fait que l'avantage annuel devient supérieur au produit du coût de l'investissement par le taux d'actualisation. Si plusieurs variantes sont envisageables, on doit retenir celle qui, supposée mise en service à son année optimale, dégage le bénéfice actualisé, (on dirait aujourd'hui la « Valeur actuelle nette, VAN ») le plus grand.

Le programme optimal, celui qui conduit au bénéfice actualisé total maximal, est donc celui qui comporte toutes les opérations rentables,

1. Ingénieur général honoraire des Ponts et Chaussées.

2. Dont je faisais partie, ce qui justifie peut être qu'on m'ait confié la rédaction de ce préambule.

dans leurs variantes les plus rentables, qui ont atteint ou éventuellement dépassé leur année optimale de mise en service.

Le taux d'actualisation retenu, d'abord suggéré par EDF, sera rapidement fixé par le Commissariat général du Plan. Les valeurs du temps, ou des consommations des véhicules, sont déduites, de mesures pour les secondes et, tant bien que mal, des comportements pour les premières. Quant à la prise en compte de la sécurité, les ingénieurs de la direction des routes proposent une méthode d'évaluation, tout en précisant qu'il incomberait au pouvoir politique de dire le droit³.

En dépit de leur apparente simplicité, ces raisonnements marquent une double évolution :

- Une évolution technique : le monde des transports est imprégné de la notion de capacité ; une route, si elle est saturée, doit être élargie ou remplacée. Le nouvel aménagement doit être capable de faire face aux besoins du trafic pendant trente ans⁴. La capacité elle-même est définie comme le trafic de la 30^e heure de pointe qui peut être écoulee sans perturbation notable. La bible des ingénieurs de l'époque est davantage le « Highway Capacity Manual » américain qu'un cours d'économie politique. Ces problèmes de capacité sont encore aujourd'hui, en matière de décision, largement dominants pour les infrastructures non linéaires (ports, aéroport). Et la référence à la trentième heure demeure présente dans de nombreux textes officiels, même pour les infrastructures linéaires⁵.

- Une évolution économique : même pour ceux qui acceptent le principe du calcul économique, les notions de temps de récupération, (durée au bout de laquelle l'investissement initial est « amorti ») ou de taux de rentabilité interne (taux d'actualisation annulant le bénéfice actualisé) sont plus familières que celle de la maximisation du bénéfice actualisé. On retrouve d'ailleurs ces notions dans des textes officiels récents, dont la célèbre Loi d'orientation sur les transports intérieurs (Loti, 1982).

Largement utilisés à la Direction des Routes, les calculs de rentabilité, sous l'influence, en particulier, du Service des Affaires Économiques et Internationales (SAEI) du ministère de l'Équipement, celle du Commissariat Général du Plan, et enfin celle de la Direction de la Prévision du ministère de l'Économie et des Finances, gagnent pro-

3. Jacques Thédié, Claude Abraham, « Le prix d'une vie humaine dans les décisions économiques », *Revue Française de Recherche Opérationnelle*.

4. Pourquoi 30 ans ?

5. Par exemple, la circulaire du 12 décembre 2000 : Instruction sur les conditions techniques d'aménagement des autoroutes de liaison, ou le « Guide des études de trafic interurbain », du SETRA, de mai 1992.

gressivement, dans le domaine des infrastructures linéaires, les autres composantes du secteur des transports.

2. DES QUESTIONS ANCIENNES QUI RESTENT D'ACTUALITÉ

Cependant, l'utilisation systématique des calculs aux fins de programmation se heurte, dès l'origine, à toute une série d'obstacles concernant :

1. Leur légitimité : le résultat du calcul doit-il s'imposer aux décideurs ?
 2. Leur crédibilité :
 - prend-on correctement en compte tous les avantages et désavantages d'un projet ?
 - chiffre-t-on correctement les avantages ou désavantages non marchands ?
 - le taux d'actualisation retenu est-il satisfaisant ?
 - faut-il prendre en compte une « désutilité », un « coût d'opportunité » des fonds publics ?
 3. Leur acceptabilité : comment traiter les problèmes d'équité ?
 4. Les aléas et l'incertitude, qui affectent toutes les composantes du calcul.
 5. Les écarts entre les besoins budgétaires que justifieraient les calculs et les allocations de crédits.
- Reprenons ces cinq points d'achoppement.

1. La légitimité

Sur ce premier point, les premiers tenants du calcul économique s'empressent de remettre les choses à leur place : « *les calculs économiques ne doivent pas avoir pour effet de se substituer aux décisions à prendre : ils constituent une donnée, parmi bien d'autres, du choix des investissements, choix qui ne saurait s'extraire de considérations plus générales de caractère politique et social notamment.* »

Il est permis de faire des investissements peu rentables, mais à condition de savoir ce qu'il en coûte⁶. »

Dès cette époque, Pierre Massé met en garde : « *Il faut se garder de croire que l'emploi de l'appareil mathématique conduit à la vérité absolue. Il garantit simplement – sauf erreur de calcul – que les conclusions*

6. André Laure et Claude Abraham, « L'Étude économique des investissements routiers », *La Route*, Supplément à la revue Travaux 1958.

obtenues sont la conséquence exacte des hypothèses adoptées. Mais il n'ajoute pas une parcelle de vérité aux hypothèses elles-mêmes. Et comme celles-ci sont presque toujours, pour les nécessités du calcul, une schématisation abrégée de la réalité, ce sont elles qui exigent de notre part l'effort le plus difficile et le plus fécond.

On a raison de croire au calcul quand on y voit un instrument qui abrège le chemin, un moyen d'expression plus puissant que le langage ordinaire. On a tort de croire aux chiffres quand on leur attribue une vertu propre, un fantôme d'absolu. Il faut perdre le complexe de la troisième décimale quand (c'est souvent le cas) les hypothèses de départ sont à peine formulées à 10 % près. »

Près de cinquante ans plus tard, dans son rapport d'avril 2000⁷, Marcel Boiteux écrit : *« l'intérêt de faire du bilan socio-économique, non le critère, mais le noyau de l'estimation de la valeur d'un projet, c'est de permettre une analyse des raisons pour lesquelles on est conduit à s'écarter de la solution à laquelle ce seul bilan aurait conduit, et de pouvoir ainsi chiffrer le surcoût de la décision⁸. »*

On notera la réticence évidente des économistes à la prise en compte, par d'autres, de ce qui échappe au calcul, réticence qui n'est pas nouvelle. Déjà, en 1884, à propos des subventions aux liaisons ferroviaires déficitaires, Clément Colson écrivait :

« L'existence de bénéfices indirects n'est pas niable, mais elle a donné lieu à des exagérations très dangereuses. Comme ils sont, de par leur nature, impossibles à évaluer, même approximativement, ce sont eux que l'on invoque toutes les fois que les bénéfices directs d'une affaire apparaissent comme nettement inférieurs aux charges. Ils servent à justifier avec une égale facilité tous les sacrifices demandés au budget, que ces sacrifices soient lourds ou légers, puisque leur justification est tirée de gains impossibles à mesurer. »

Cinquante ans plus tard, on en est toujours au même point. Ainsi par exemple, la circulaire du 3 octobre 95, dite circulaire IDRAC, précise :

« L'évaluation économique des projets ne se substitue pas à la décision politique.

Elle a l'avantage de permettre des comparaisons et par là d'ouvrir le débat sur une base rationnelle. Elle nécessite une grande transparence et une séparation claire de ce qui relève de l'appréciation politique et de ce que l'état de l'art permet d'évaluer. »

7. Actualisation du rapport : *Transports : pour un meilleur choix des Investissements*, Commissariat Général du Plan, avril 2000.

8. C'est nous qui soulignons.

2. La crédibilité

Les questions sont ici nombreuses, abordons-les une par une.

2.1. La nature et l'évaluation des avantages pris en compte

La nature des avantages (et nuisances) pris en compte a évolué au cours du temps. En 1983, par exemple, le Commissariat général du Plan avait confié à un groupe de travail présidé par Edmond Malinvaud une réflexion sur ce que devait être le contenu d'un calcul économique, prenant éventuellement en compte « *les effets distributifs des projets, leurs effets sur l'emploi, la balance des paiements, et d'autres variables économiques pertinentes* ».

Plus récemment, la dispersion, selon les secteurs où les services, des évaluations et des modes de calcul ont conduit le Commissariat Général du Plan à confier à Marcel Boiteux la présidence de deux groupes de travail successifs⁹, dont les rapports ont servi de base à la rédaction d'instructions ministérielles. Luc Baumstark explique, dans un des chapitres la philosophie, et la pratique, qui ont sous-tendu l'évaluation des avantages et inconvénients non marchands pris en compte dans les calculs de surplus.

2.2. Le taux d'actualisation

En ce qui concerne le taux d'actualisation, le Commissariat général du Plan, au cours des 50 dernières années, a systématiquement fixé la doctrine, et vient encore de le faire récemment, à la suite des travaux d'un groupe de réflexion présidé par Daniel Lebègue¹⁰. Ce groupe a recommandé l'adoption d'un taux d'actualisation de 4 %, décroissant progressivement dans le temps au-delà de 30 ans, à condition que les risques soient pris en compte par ailleurs.

2.3. Le « coût d'opportunité des fonds publics »

En raison des distorsions économiques introduites par la perception des impôts, il a été suggéré, à plusieurs reprises, de majorer les dépenses sur fonds publics par la prise en compte d'un coût d'opportunité des fonds publics. Le rapport Lebègue suggère de retenir, provisoirement, un taux de majoration de 30 %.

Ce n'est pas la première fois que cette notion de majoration du coût des fonds publics est suggérée. Il est frappant de noter que, à ma

9. Commissariat Général du Plan, « Transports ; pour un meilleur choix des investissements », Novembre 1994, actualisé en avril 2000 ?

10. « Révision du taux d'actualisation des investissements publics », Rapport du groupe d'experts présidé par Danielle Lebègue, Commissariat Général du Plan, 21 janvier 2005

connaissance du moins, elle n'a jamais été transcrite dans les instructions officielles, et qu'elle n'apparaît, dans l'instruction cadre du 25 mars 2004 ¹¹, que dans son annexe III, avec une rédaction qui laisse l'impression qu'on s'intéresse surtout aux projets privés subventionnés.

Toutefois, il pourra être tenu compte du coût d'opportunité des fonds publics dans l'évaluation d'un projet sans l'intégrer directement dans sa valeur actualisée nette mais en s'assurant que la valeur actualisée nette ainsi calculée (sans prise en compte du coût d'opportunité des fonds publics) par euro public dépensé est supérieure ou égale au coût d'opportunité des fonds publics.

3. L'acceptabilité

Une des hypothèses fortes, et implicites, du calcul classique du surplus est celle qui veut que l'on puisse faire la somme algébrique des surplus de tous les intéressés, quels qu'ils soient, que ces surplus soient positifs ou négatifs, ce qui revient à admettre :

- que les pertes des uns « puissent être compensées » par les gains des autres. Mais le fait qu'elles puissent être compensées ne signifie pas pour autant qu'elles le soient ;
- que les gains sont additionnables ;
- et que le surplus d'un individu « privilégié », et celui d'une personne « défavorisée », sont « commensurables ».

Incontestablement, les calculs économiques intègrent mal les problèmes d'équité et, partant, d'acceptabilité des projets par ceux qui n'en tirent nul avantage, au contraire. Deux questions particulières peuvent être distinguées en matière d'équité.

- Des questions d'équité spatiale : à coûts égaux, les projets sont d'autant plus rentables que le trafic y est plus important. Aussi, pour les relations interurbaines, les calculs de rentabilité privilégient la concentration des investissements sur les itinéraires très fréquentés, et permettent rarement de justifier des équipements dans les zones peu denses, et mal desservies.

- Des questions d'équité sociale : la valeur du temps particulièrement élevée des usagers à haut revenu justifie-t-elle les investissements qui en résultent ? Et comment tient-on compte des usagers à faibles revenus, incapables de bénéficier des infrastructures mises à leur disposition ? Le surplus collectif, somme algébrique de surplus indivi-

11. Ministère de l'Équipement, « Instruction cadre relative aux méthodes d'évaluation économique des grands projets d'infrastructure de transport », 25 mars 2004. Mise à jour 27 mai 2005.

duels, dont certains sont d'ailleurs négatifs, et non compensés, est-il réellement représentatif d'un véritable intérêt collectif ?

La non prise en compte de ces problèmes expliquerait l'écart entre la rentabilité des projets, calculée selon les méthodes classiques, et leur acceptabilité¹². Un bon projet n'est-il pas d'abord un projet acceptable, et donc réalisable dans un délai raisonnable ?

Cela expliquerait et éclairerait un fait noté par Jean-Pierre Giblin : « *les différents groupes de pression font peu de cas des bilans socio-économiques lorsqu'il s'agit de choisir ou de hiérarchiser des projets*¹³. » Le même auteur ajoute : « *une bonne partie des déconvenues qui rendent le calcul économique peu crédible serait évitée si l'on était capable d'introduire des considérations d'équité à côté de celles d'efficacité, qui caractérisent les approches économiques utilitaristes.* »

Jean-Pierre Giblin a sans doute raison, et il faudra sûrement progresser dans cette direction. Vincent Piron et moi-même avons proposé la prise en compte d'un nouveau concept, celui d'« amertume », écart entre ce qu'un usager aurait accepté de payer, et le prix qu'on lui réclame¹⁴. Mais ces compléments d'analyse, sur lesquels Jean-Pierre Giblin fait également de très intéressantes suggestions, n'enlèvent rien aux calculs classiques qui demeurent, quoi qu'on en ait, le *noyau incontournable* de l'évaluation, pour reprendre l'expression de Marcel Boiteux. À trop déprécier la valeur de ce noyau, on risquerait fort de n'avoir plus grand-chose à se mettre sous la dent !

Peut-on dire que les calculs, cependant, ont évolué, se sont améliorés et généralisés, et que leur utilisation s'est effectivement répandue ?

4. La prise en compte du risque

Si l'on sait assez bien comment les investisseurs privés, leurs financiers, dans le cas des financements de projets, incorporent l'évaluation des risques dans leurs calculs, on sait moins ce que devrait être le comportement des investisseurs publics. L'instruction cadre de mai 2005, dans son annexe III, apporte des éléments de réponse, en distinguant les risques spécifiques, propres à chaque projet, et les incertitudes liées à l'environnement externe du projet, qui seraient incorporées

12. On en a vu une illustration à l'occasion des débats relatifs au projet de liaison ferrée « CDG Express ».

13. Jean-Pierre Giblin, « De l'eau a coulé sous les ponts depuis Jules Dupuit », Note interne au CGPC, mars 2004.

14. Voir, en particulier : Vincent Piron et Claude Abraham, « Amertume et acceptabilité des péages », *Cahiers Scientifiques des Transports*, Automne 2001.

dans le taux d'actualisation. Les articles de Ch. Gollier et A. de Palma tentent d'apporter des éclairages complémentaires sur cette question.

5. *Le problème des insuffisances budgétaires*

Ce problème des insuffisances budgétaires, si elles sont systématiques, pose des questions de fond : si les pouvoirs publics ne peuvent, ou ne veulent, financer tous les programmes dont les calculs indiquent qu'ils amélioreraient le bien commun, c'est que les calculs sont faux, et que, en particulier, le bénéfice actualisé est surévalué. Il faut alors remettre en cause, de mon point de vue, soit l'évaluation des avantages et des nuisances, soit leurs perspectives de croissance, soit les taux d'actualisation, soit l'évaluation du risque, soit le coût d'opportunité des fonds publics.

On peut d'ailleurs discuter du caractère accidentel ou systématique de l'écart entre les besoins budgétaires révélés, et justifiés par les calculs, et les allocations effectives de crédit. Je suis de ceux qui pensent que cet écart est systématique, même si l'on prend en compte un coût d'opportunité des fonds publics, et une évaluation raisonnable du risque, sauf à ne retenir que les projets dont le bénéfice actualisé a plus de 70 ou 80 % de chances d'être atteint ou dépassé.

Quoi qu'il en soit, le problème de la recherche du programme d'investissement optimal, et de la décentralisation possible des calculs, en cas d'insuffisance de crédits, reste un problème majeur. L'instruction cadre du 25 mars 2004 et sa mise à jour du 27 mai 2005 indiquent que les projets seront retenus en les classant par valeur décroissante de la valeur actualisée nette par euro public investi, tout en étant muette sur le problème de l'année optimale de mise en service, et du choix entre variantes.

De nombreux économistes s'interrogent sur le caractère optimal de cette formule. J'en avais moi-même, avec André Laure¹⁵, proposé une autre, à la suite d'une démonstration dont la rigueur n'est peut-être pas évidente, et qui conduisait à retenir, chaque année, les projets par valeur décroissante du rapport avantage de l'année de mise en service divisé par coût de l'investissement. Ce qui revenait à prendre en compte un coût fictif des fonds publics, tel que l'année de mise en service permise par les disponibilités budgétaires, et l'année optimale de mise en service, coïncident.

15. Claude Abraham et André Laure, « Étude des programmes d'investissements routiers », *Annales des Ponts et Chaussées*, novembre-décembre 1959.

Concernant ce problème, on trouvera ci-après, dans les articles rassemblés dans la seconde partie de l'ouvrage, de nouvelles analyses approfondies, et sûrement plus rigoureuses, encore que leurs conclusions ne convergent pas. Preuve s'il en fallait que les recherches sur le calcul économique ont encore de l'avenir !

INTRODUCTION GÉNÉRALE

Au mois de février 2003, un rapport sur les grands projets d'infrastructure de transport était remis au ministère de l'Équipement et des Transports. Rédigé, à la demande du ministre, par le Conseil général des Ponts et Chaussées et l'Inspection des Finances, il passait au crible du calcul économique l'ensemble des projets de nouvelles infrastructures de transport terrestre (routes, fer, voies navigables). Ses conclusions ont été d'emblée contestées par tous ceux, notamment les élus locaux, qui voyaient dans ses résultats la remise en cause de la programmation d'infrastructures les concernant. Se faisant très largement l'écho de la protestation des élus, le CIADT¹ du 18 décembre 2003 annonçait le lancement de 50 grands projets d'infrastructure pour les 20 années à venir : un programme où les questions de financement et de rentabilité socio-économique étaient peu détaillées. Cet exemple, parmi beaucoup d'autres, illustre la perte de crédibilité des processus élaborés dans les années 60 pour garantir une certaine rationalisation des choix budgétaires.

Le calcul économique est donc aujourd'hui en partie déconsidéré alors même que s'accroissent les contraintes financières qui pèsent sur les budgets publics. Le recours à des outils d'aide à la décision, propres à rationaliser les choix publics, devrait donc se renforcer. Mais les choses ne sont pas si simples. La décision publique ne peut être le décalque du classement des projets par les experts. Il est au contraire logique que la capacité du calcul économique à chiffrer les coûts et les surcoûts, et donc à rendre commensurables et comparables différents projets, soit source de tensions. Le principal obstacle que rencontre le calcul économique provient donc de sa capacité à bousculer le jeu des intérêts. Pour cette raison, sa légitimité est régulièrement mise en doute, et il serait naïf de croire que les groupes de pression, quels qu'ils soient, soient un jour prêts à se rendre aux arguments du calcul économique. Mais cela ne signifie pas qu'il doive être écarté, au contraire. Le décideur public doit se donner les moyens de savoir ce que seront les implications de ses choix.

Pour cette raison, suivant une demande du conseil général des Ponts et Chaussées, les Groupes opérationnels (GO) n° 1 et n° 11 du PREDIT ont lancé un travail de recherche sur les pistes de travail qui

1. . Comité interministériel d'aménagement et de développement du territoire.

pourraient, non pas refonder, mais enrichir le calcul économique. Comment le rendre plus apte à répondre aux questions qui se posent aujourd'hui en matière de développement des infrastructures de transport ?

En réponse à cette demande, un « Réseau de recherche sur les dimensions critiques du calcul économique » a été constitué à la fin de l'année 2004. À la suite du préambule, rédigé par Claude Abraham, les contributions des participants de ce groupe ont été regroupées dans cet ouvrage organisé en trois parties :

- la construction du calcul économique (hypothèses de trafic, structures de marchés, valeurs tutélaires, taux d'actualisation, risque et incertitude) ;
- la programmation des investissements (comment passer d'un ensemble de projets à un programme ordonné ?) ;
- l'aide à la décision publique (mesure des impacts, efficacité, équité et acceptabilité, questions d'aménagement du territoire).

PREMIÈRE PARTIE

LA CONSTRUCTION DU CALCUL ÉCONOMIQUE

Dans le domaine des transports, le choix des investissements a un impact très important sur le développement économique et social, mais aussi sur les finances publiques. Les autoroutes, les routes, les voies ferrées classiques et à grande vitesse, les aéroports, marquent durablement l'évolution et le développement du territoire. Les choix effectués aujourd'hui engagent donc l'avenir et sont porteurs d'irréversibilités. L'évaluation, *a priori* et, ce qui est moins connu, *a posteriori*, constitue donc une obligation intellectuelle et même morale, face à l'usage de l'argent public et aux enjeux financiers, économiques et sociaux. L'évaluation économique des choix d'investissements relève donc d'une exigence démocratique. Non seulement les évaluations sont rendues publiques, mais elles doivent éclairer les choix collectifs. Aussi, lorsqu'il a été progressivement acclimaté en France, après la Seconde Guerre mondiale, et lors des tentatives de rationalisation des choix budgétaires (RCB), le calcul économique a-t-il été présenté comme une chance pour la démocratie, un moyen de rendre plus transparents et plus cohérents les choix publics. « Pouvoir sans volonté au service d'une volonté sans pouvoir »¹, l'État se trouvait doté, par le calcul économique, d'un moyen d'étalonner les priorités collectives. Dans sa conception la plus optimiste, le calcul économique allait clarifier les processus politiques.

Ainsi, d'un point de vue réglementaire, la loi d'orientation des transports intérieurs (LOTI) du 30 décembre 1982 (article 14) et son décret d'application du 17 juillet 1984 impliquent la réalisation d'évaluations socio-économiques *a priori* et *a posteriori*.

« [...] Les grands projets d'infrastructures et les grands choix technologiques sont évalués sur la base de critères homogènes intégrant les impacts des effets externes des transports relatifs notamment à l'environ-

1. Selon la formule de S.C. Kolm lorsqu'il décrit le rôle de l'État dans la tradition économique libérale.

nement, à la sécurité et à la santé et permettant de procéder à des comparaisons à l'intérieur d'un même mode de transport et entre différents modes ou combinaisons de mode. Ces évaluations sont rendues publiques avant l'adoption définitive des projets concernés. Lorsque ces opérations sont réalisées avec le concours de financements publics, un bilan des résultats économiques et sociaux est établi au plus tard cinq ans après leur mise en service. Ce bilan est rendu public [...] », (extrait de l'article 14 de la Loti).

Plus récemment, la loi sur l'air et l'utilisation rationnelle de l'énergie (Laure) du 30 décembre 1996 et la loi sur l'orientation pour l'orientation, l'aménagement et le développement du territoire du 25 juin 1999 prévoient l'introduction dans les évaluations des effets sur l'environnement, la sécurité et la santé, toutes choses que le calcul économique sait faire par la monétarisation des effets externes. Cette étape clé du calcul économique a largement mobilisé les administrations françaises, qui ont établi des valeurs tutélaires des coûts externes comme le bruit et la pollution, mais aussi la vie humaine et le temps. Des réflexions ont ainsi été engagées au début des années 90 sous l'égide du Commissariat Général du Plan. Les publications des rapports Boiteux : *Transports : pour un meilleur choix des investissements* en 1994 et *Transports : choix des investissements et coût des nuisances* en 2001, constituent des références en la matière.

- Une des premières composantes concrètes de ces attentes se trouve dans les trafics. La présence d'un trafic élevé sur une infrastructure est en effet le premier signe tangible de son utilité sociale. Aussi, tout exercice de calcul *a priori* de la rentabilité d'un investissement doit-il se fonder sur des prévisions de trafic. Il s'agit d'un exercice très délicat qui, lorsqu'il conduit à des résultats erronés, a des impacts importants sur la qualité des choix collectifs. Pour cette raison, le chapitre 1 (Marc Gaudry) développe longuement toutes les subtilités de ce travail, il nous aide aussi à comprendre les principales sources d'erreur et comment on peut tenter d'y remédier.

- Le chapitre 2 s'intéresse aussi, plus indirectement, aux prévisions de trafic. Marc Ivaldi et Catherine Vibes y montrent que les prévisions doivent aussi s'efforcer de prendre en compte d'éventuels changements structurels des conditions de la concurrence. Ainsi, le développement des compagnies aériennes dites Low Cost, pourrait conduire à une structure de marché inédite lorsque l'on s'intéresse à la pertinence d'une ligne, voire d'un réseau ferroviaire à grande vitesse.

- Dans le calcul économique, le choix du taux d'actualisation est une donnée sensible. C'est pourquoi le chapitre 3, confié à Christian Gollier, revient sur les principales variables qui déterminent le taux

d'actualisation. Il insiste sur les raisons qui ont récemment poussé la puissance publique française à modifier le taux d'actualisation tutélaire pour l'évaluation des projets publics. Il insiste sur le fait que ce changement invite surtout à une meilleure prise en compte des risques.

- Prolongeant cette interrogation essentielle sur la question des risques, Laetitia Andrieu, André de Palma et Nathalie Picard montrent dans le chapitre 4 comment il est possible d'évaluer les risques et d'intégrer les valeurs ainsi obtenues dans le calcul économique. Il s'agit là d'un champ bien balisé par les acteurs privés, notamment financiers, mais paradoxalement mal connu dans la sphère publique. Cette situation n'est pas durable quand on sait que la puissance publique est de plus en plus adverse au risque.

- Il est ensuite nécessaire d'insister sur le fait que les taux de rentabilité interne (TRI) calculés ne sont pas seulement des valeurs financières. Le calcul économique produit également des TRI socio-économiques dont le principe est la valorisation, dans les projets, des avantages non marchands. C'est le fruit d'une démarche de monétarisation des principaux effets identifiés, à savoir les gains ou pertes de temps pour les usagers, la diminution ou l'accroissement des pollutions classiques liées aux transports, les améliorations ou dégradations de la sécurité, du bruit (surtout en milieu urbain) et des émissions de CO₂. Le chapitre 5 est consacré par Luc Baumstark à un rappel des enjeux liés à la définition de valeurs tutélares telles que celles qui ont été définies par les deux rapports Boiteux.

- L'estimation monétaire de ces différents avantages relève d'un caractère nécessairement incertain, du fait de la nature non palpable de ces effets. Pourtant, dans un bilan socio-économique, leur part est souvent prépondérante dans les résultats finals. Les valeurs tutélares prises en compte dans la valorisation de ces avantages ont un caractère décisif. À ce titre, des tests de sensibilité sont nécessaires. En poussant le raisonnement sur d'éventuelles modulations des valeurs tutélares, on peut même tenter un calcul économique « à l'envers ». Ce à quoi s'essaye Guillaume Chevasson dans le chapitre 6.

CHAPITRE 1

STRUCTURE DE LA MODÉLISATION DU TRAFIC ET THÉORIE ÉCONOMIQUE¹

Marc Gaudry²

1. MODÉLISATION DU TRAFIC, THÉORIE ÉCONOMIQUE ET ÉVALUATION DES PROJETS

Modifications de réseaux et calcul financier et économique. Il s'agit ici d'isoler la structure typique de la modélisation du trafic et d'en relier les pratiques fondamentales et articulations-clés aux concepts de base de la théorie économique. La perspective de cette modélisation est celle de l'utilisation des résultats des modèles obtenus « en situation de référence » pour évaluer des projets d'investissement dans les infrastructures ou celle de la détermination des niveaux de service des modes de transport. Plus généralement, la perspective est celle du calcul financier et économique associé aux modifications d'infrastructure ou de service envisagées « en situation de projet ».

Structure et articulations clés seulement. Il ne relève donc pas de notre propos de faire directement l'histoire de cette modélisation, d'en répertorier les nombreuses variantes urbaines et interurbaines, ou même internationales, traitant des voyageurs et des marchandises, ou d'en développer certaines pistes de réflexion nouvelles dont on pourrait pressentir ici les amorces, voire les arrhes. Il nous faudra faire appel à des concepts simples maîtrisés au cours des études universitaires de premier cycle ou définis intuitivement ; il faudra beaucoup simplifier pour mettre l'accent sur les charnières-clés de la méthodologie ou des outils et leurs implications pour l'évaluation des projets.

1. L'auteur remercie Ariane Dupont-Kieffer de l'INRETS dont les nombreuses suggestions ont grandement contribué à améliorer la première version de ce texte.

2. INRETS-DEST et Université de Montréal-AJD.

2. L'ÉCONOMIE, LES FLUX ET LE RÉSEAU EN PÉRIODE DE RÉFÉRENCE : UN CADRE D'ÉQUILIBRE PARTIEL

Une approche agrégée compatible avec les développements désagrégés récents. Nous supposons disponibles toutes les informations de nature fine sur les flux (déplacements individuels de personnes ou envois de marchandises), leurs causes et les conditions de transport sur les réseaux d'infrastructures ou de services. Cependant, notre description sera ici d'abord de nature agrégée pour respecter l'évolution historique des méthodes, agrégées à l'origine, depuis un demi-siècle et la stratégie de modélisation qui en est née et qui a donné naissance à l'approche en quatre étapes.

Activités, flux, réseaux, zones. Supposons donc donnés, durant une période de référence souvent journalière ou annuelle (i) les niveaux et la distribution spatiale des populations et activités économiques localisées qui expliqueront les flux de transport ; (ii) les flux de transport par mode et combinaison de modes entre toutes les localisations $(1, \dots, i, \dots, j, \dots, Z)$ appelées zones ; (iii) la topologie et les caractéristiques des liens³ s constitutifs de chaque réseau (capacités, prix et conditions de transport et de sécurité) reliant toutes ces localisations i ou j .

Activités données. Les modèles de trafic s'inscrivent dans une perspective d'équilibre partiel où l'on cherche à répondre à des questions précises de modification et dimensionnement fin des réseaux du genre : à **niveaux et localisations** d'activités donnés, « combien de véhicules particuliers et de poids-lourds passeront sur ce nouveau lien s^* s'il est construit et de combien les autres liens verront-ils leur usage modifié ? ».

Demandes dérivées et conditionnelles. Pour répondre à la question, on estimera des fonctions de demande définies de i à j conditionnellement à ces diverses activités $\mathbf{A}_i \equiv (A_{i1}, \dots, A_{ia}, \dots, A_{iA})$ présentes dans toutes les zones $\mathbf{A} \equiv (\mathbf{A}_1, \dots, \mathbf{A}_i, \dots, \mathbf{A}_j, \dots, \mathbf{A}_Z)$. Notons dès maintenant que la demande sera définie de i à j mais qu'on s'intéressera nécessairement à sa réalisation sur des liens s du réseau parce que c'est sur les arcs orientés du réseau que se déterminent les conditions de transport de i à j . C'est sur les arcs qu'interagissent les demandes reliant plusieurs paires origine-destination et que s'établissent les niveaux de service effectifs qui génèrent les quantités demandées par paire origine-destination.

3. Le niveau de précision de notre présentation n'exige pas de distinguer entre liens non orientés et arcs orientés.

Sans rétroaction sur les activités. Dans le cadre fondamental et strict de la modélisation du trafic, on ne modifie pas la localisation des activités suite à des modifications du réseau, mais des accommodements partiels sont parfois aménagés, voire implantés dans les logiciels de planification du trafic. C'est dire que les perspectives de l'équilibre général sont au mieux lointaines, même si des efforts récents de modélisation (*e.g.* Bröcker, 2000) d'équilibre général calculable d'un réalisme croissant fournissent des représentations de l'ensemble de l'économie, niveaux et distribution spatiale des activités compris, et des flux de transport : on trouvera l'exemple sans doute le plus complet dans Mercenier *et al* (2001). Nous y reviendrons lorsque nous discuterons des effets de rétroaction des modifications du réseau sur la localisation et le niveau des activités.

2.1. La représentation des flux par un système à quatre matrices

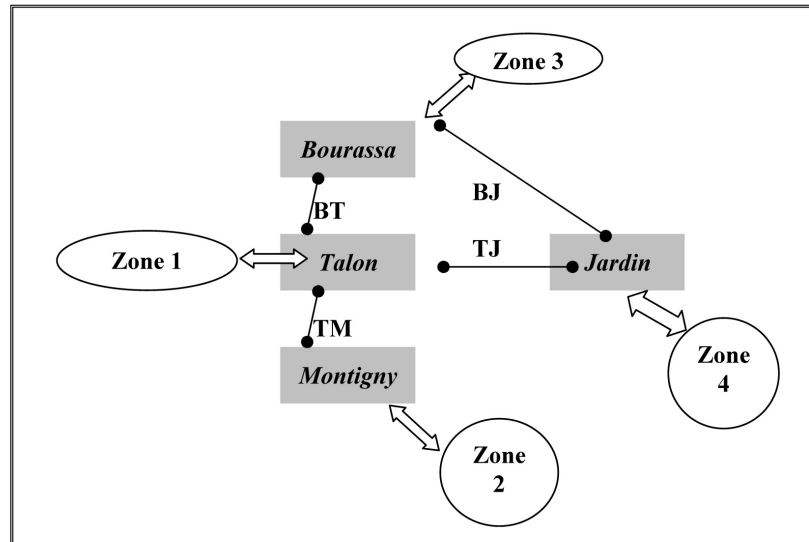
On peut représenter l'état des flux modélisés en période de référence à l'aide de matrices inspirées des tableaux interindustriels dont les cases pourront être associées aux quatre étapes des modèles classiques de trafic. Formulons un exemple inspiré du transport en commun urbain.

A. Comptage des flux. Supposons que les flux entrants ou sortants d'une zone (1,... 4) soient inscrits lors de leur *entrée/sortie* à une station (Bourassa, Talon, Jardin, Montigny) et lors de leur *passage sur les liens* non orientés (BT, TJ, TM, BJ) d'un réseau unimodal reliant ces stations, mais ne soient pas comptés à l'intérieur des stations lors des correspondances. Par exemple, 4 flux (C1, C2, D1, D2) allant de la Zone 4 vers les autres zones sont ainsi inscrits au Tableau 2 lors de leur usage du réseau physique représenté au Tableau 1 (voir page suivante).

B. Inscription des flux dans la matrice OD-ENL. Dans la matrice présentée au Tableau 2, le flux C1 qui va à la zone 3 est inscrit lors de son entrée dans la station Jardin, lors de son passage sur les liens TJ et BT et lors de sa sortie à la station Bourassa. Le flux C2 qui a comme destination la zone 2 est inscrit successivement lors de son entrée à la station Jardin, lors de son usage des liens TJ et TM et lors de sa sortie à la station Montigny. Les flux D1 et D2 en direction des zones 3 et 1 respectivement, sont en fait des flux directs (sans correspondance) passant respectivement par les liens BJ et TJ en direction des zones 3 ou 1.

Cette matrice comptable contient en fait 4 matrices : la matrice *E* des comptages à l'entrée du réseau, la matrice *L* qui désigne les flux de

Tableau 1
Représentation d'un réseau de 4 stations reliant 4 zones par 4 liens



sortie vers l'extérieur du réseau, la matrice N des flux internes au réseau et la matrice $O-D$ (supposée connue) des flux origine-destination. On peut les représenter ainsi :

N	L
E	O-D

Au Tableau 2, on reconnaît facilement \mathbf{FE} et \mathbf{FL} , les vecteurs des flux entrant de l'extérieur et sortant vers l'extérieur du réseau, ainsi que \mathbf{NE} et \mathbf{NL} , les vecteurs des flux entrant dans le réseau ou en sortant. Les vecteurs totaux $\mathbf{TE} = [\mathbf{FE} + \mathbf{NE}]$ et $\mathbf{TL} = [\mathbf{FL} + \mathbf{NL}]$ désignent alors naturellement pour le premier les flux totaux entrant dans une station en provenance soit de l'extérieur ou d'une autre station et, pour le second, les flux sortant soit vers une autre station ou vers l'extérieur. (Voir page ci-contre.)

C. Propriétés comptables élémentaires du système : un exemple.
Considérons au Tableau 3 l'exemple d'un système simplifié sans lien direct entre Jardin et Bourassa, système où les flux directs sont soulignés et les flux en correspondances sont indiqués en gras. La matrice $O-D$ choisie n'est pas symétrique.

Tableau 2

Inscription des flux ayant comme origine la zone 4

Matrice I		Destination : réseau					Destination : extérieur				Total	
		O/D	B	T	M	J	NL	1	2	3	4	FL
Origine : réseau	B								D1+C1			
	T	C1		C2			D2					
	M							C2				
	J	D1	C1+C2 +D2									
Origine : extérieur	NE											
	1											
	2											
	3											
	4					D1+D2 C1+C2		D2	C2	D1+C1		
FE												
Total	TE											

Si on note e' le transposé d'un vecteur unitaire et $diag(.)$ la diagonale d'une matrice, les propriétés de cette matrice II sont, pour le cas où il n'y a pas d'accumulation de voyageurs sur le réseau :

- i) définition des flux totaux entrants : $TE = FE + NE$; (1)
- ii) définition des flux totaux sortants : $TL = NL + FL$; (2)
- iii) comme $diag(N) = 0$, le flux total : $e'FE' = e'FL = 2\ 175$; (3)
- iv) et l'identité comptable du système : $TE' = TL$, (4)

c'est-à-dire que tout ce qui passe par un point Bourassa, Talon, Montigny ou Jardin du réseau en provenance de l'extérieur ou d'un autre point du réseau est égal à ce qui en sort vers un autre point ou vers l'extérieur. Voir tableau page suivante.

Ici $FE' \neq FL$ puisque la structure des flux entrant de l'extérieur n'est pas la même que celle des flux sortant vers l'extérieur, mais l'Annexe 1 utilise un tel cas particulier III fondé sur une matrice $O-D$ symétrique – pour lequel l'identité comptable (4) est toujours vérifiée.

D. Les variables qui nous intéressent dans le processus classique à 4 étapes. Ce cadre comptable permet d'abord de bien isoler ce dont il sera question à chaque étape du processus classique de conduite des études de trafic : d'abord le **Total** de déplacements $T = e'FE' = e'FL = \sum_{i \in I} \sum_{j \in J} T_{ij} = T_{IJ}$; ensuite l'objet des analyses de la première étape, celle de la **Génération**, puisque les marges⁴ droites de la matrice E désignent les productions ou **émissions** $T_i = \sum_j T_{ij} = O_i$ (en brun) et les marges

4. Au sens strict, puisque la matrice $O-D$ est un tableau de fréquences, les « marges » sont les vecteurs des fréquences *marginales* des flux origine-destination.

Tableau 3
Flux pour un système à 4 zones, 4 stations et 3 liens

Matrice II		Destination : réseau					Destination : extérieur				Total		
		O/D	B	T	M	J	NL	1	2	3	4	FL	TL
Origine : réseau	B		150 500 75			725			25 200 25			250	975
	T	25 200 25		400 500 50	75 300 100	1675	250 150 100					500	2175
	M		300 200 100			600		400 500 50				950	1550
	J		25 400 250			675					75 300 100	475	1150
	NE	250	2000	950	475	T_i T_j	500	950	250	475		2175	
Origine : extérieur	1		100 25 50			175		50	25	100			
	2			100 200 300		600	100		200	300			
	3	150 500 75				725	150	500		75			
	4				25 400 250	675	250	400	25				
	FE	725	175	600	675	2175							
Total	TE	975	2175	1550	1150								

inférieures de la matrice L désignent les *attractions* $T_j = \sum_i T_{ij} = D_j$ (en vert). De plus, la matrice $O-D$ contient les T_{ij} observés (donc $T_{ij} \geq 0$) analysés à l'étape de la *Distribution* et la matrice N comprend les flux par lien qui seront les objets de l'*Affectation*. En modifiant les conventions d'écriture de la matrice N et/ou la matrice $O-D$ de façon à distinguer entre les modes (par exemple avec des polices ou des couleurs), ce qui aurait beaucoup alourdi le cadre comptable, il aurait aussi été possible d'y identifier explicitement les valeurs modales N_m ou T_{ijm} ($m = 1, \dots, M$) expliquées à l'étape du *Choix modal*.

Ce cadre permet aussi de mettre en lumière, en le dédoublant, un problème de prévision familier en analyse input-output. Pour le constater, il y a lieu de relier notre cadre comptable, élaboré à l'origine dans Gaudry (1973), à la théorie économique pertinente, celle de la comptabilité nationale et de l'analyse input-output.

E. Relation à la théorie économique. Quesnay (1759) avait une simple matrice 4x4 dont les lignes et colonnes ne s'équilibraient ou ne « balançaient » qu'après beaucoup d'efforts, mais il ne distinguait pas entre flux intermédiaires et finaux comme l'a fait plus tard Leontief (1941) en formulant les partitions N et L .

En nous inspirant de Blankmeyer (1971), nous réinterprétons simplement L des demandes finales comme une matrice de « flux vers l'extérieur » du réseau, ce qui appelle naturellement à reconnaître une contrepartie E de « flux vers l'intérieur » et leur combinaison évidente dans le dernier quadrant de la matrice. Cette spatialisation des matrices de flux de transport diffère profondément d'une simple démultiplication spatiale des cases de N largement pratiquée en économie régionale (Moses, 1955) et en économie des transports (Cascetta et Di Gangi, 1996), démultiplication qui ne change pas la structure de Leontief mais change la taille de la matrice des flux N dont les dimensions deviennent alors égales au produit du nombre de « secteurs » par le nombre de zones.

On sait que, dans cette structure classique de type Leontief, il est habituel de définir des coefficients technologiques et de démontrer, en reliant les flux finaux aux flux intermédiaires par l'intermédiaire des demandes totales, que la variation du vecteur des demandes finales n'implique pas une variation proportionnelle des flux intermédiaires. Comme les flux intermédiaires N désignent maintenant, dans notre formulation, les flux sur les liens du réseau, cette absence connue de proportionnalité est pertinente pour les prévisions de transport.

F. Demande finale et flux sur le réseau. On la retrouve bien dans la nouvelle formulation, mais dédoublée en quelque sorte, selon qu'on s'intéresse au vecteur des demandes finales d'entrée ou de sortie, parce qu'on ajoute au cadre input-output $[N ; L]$ les matrices des flux à l'entrée E qui autoriseront naturellement une seconde définition des matrices de coefficients⁵ technologiques A .

On montre dans l'Annexe 1 que, pour faire une prévision des flux intermédiaires N^* à partir d'une hausse prévue des éléments des demandes finales sortantes, \mathbf{FL}^* , on calcule successivement :

$$\mathbf{TL}^* = (\mathbf{I}-\mathbf{A})^{-1} \mathbf{FL}^* \quad (5)$$

et

$$\mathbf{N}^* = \mathbf{A} \bullet \mathbf{B}^* \quad (6)$$

en utilisant comme d'habitude la définition de A fondée sur le vecteur des flux sortants et la matrice identité I . Mais naturellement, des formules analogues sont aussi définissables *mutatis mutandis* par rap-

5. Cascetta & Di Gangi (1996) définissent une matrice diagonale dite « de commerce interzonal » T et la multiplient ensuite par une matrice A diagonalisée et des vecteurs d'output redéfinis de manière conforme. Cela a pour effet de redéfinir (5) comme suit : $\mathbf{FL} = (\mathbf{I}-\mathbf{T} \bullet \mathbf{A})^{-1} \mathbf{T} \bullet \mathbf{TL}$, où les vecteurs \mathbf{FL} et \mathbf{TL} ont aussi été redéfinis de manière conforme.

port aux flux entrants : elles seront différentes des premières si la matrice $O-D$ n'est pas symétrique.

Par contre, il n'est pas clair que modifier le vecteur des demandes finales sortantes TL (ou symétriquement celui des demandes entrantes) ait un sens si l'on ne modifie pas en même temps le vecteur des demandes entrantes TE (ou symétriquement sortantes), sauf à montrer qu'il n'y pas en général de proportionnalité entre demandes finales (productions ou attractions) et flux sur les arcs du réseau *à moins qu'on ne modifie la matrice $O-D$ elle-même*, tout en conservant inchangée la structure des choix des itinéraires inscrite en N .

Dans les applications, la prévision portera souvent indépendamment sur les émissions et les attractions de manière distincte et spécifique. Elle n'impliquera donc pas en général de proportionnalité et exigera de dérouler à nouveau entièrement la séquence des quatre étapes pour obtenir les flux sur le réseau N .

Mais avant de considérer explicitement cette séquence, il nous faut préciser la structure d'offre et de demande qui sous-tendra l'analyse dans le cas général, celui de la présence de congestion. Les notions simples de l'équilibre du marché Demande-Offre ne suffiront pas et il faudra introduire aussi un équilibre sur le réseau qui, combiné au premier, définira un (ou plusieurs) équilibre (s) Demande-Performance-Offre.

2.2. La structure Demande-Performance-Offre des modèles de trafic

2.2.1. Procédures structurelles et équilibres multiples

Des procédures. En effet, pour rendre compte des flux et, simultanément, des conditions de transport (comme le temps de transport) compatibles avec ces derniers et pour faire des prévisions d'impact de modifications du réseau, la modélisation du trafic utilise une combinaison de *procédures* complexes aux caractéristiques particulières.

Trois niveaux de procédures structurelles. Pour le comprendre, il est utile de distinguer trois niveaux d'un système : *Demande*, *Performance* et *Offre*. Et nous désignons par « procédure » l'ensemble des opérations nécessaires au calcul (parfois conjoint) d'une ou plusieurs variables – une fonction⁶ étant la forme la plus simple d'une procédure – impliquées dans un équilibre de ce système à trois niveaux.

6. Les trois niveaux sont « structurels » exactement comme si l'on avait trois équations structurelles d'un système d'équations simultanées.

Nous présentons à la Figure 1 un schéma synthétique de système à trois niveaux, schéma plus élaboré que celui qu'on trouve dans Quinet (1998) et moins élaboré que le schéma initial (Florian et Gaudry, 1980, 1983) qui comprend des niveaux supplémentaires et documente en détail les divers équilibres qui nous intéressent. Il est aussi possible de décliner cette structure en l'explicitant selon des perspectives stratégiques, tactiques ou opérationnelles (Florian *et al.*, 1988) définies en décidant, pour résoudre un problème particulier, de ce qui est endogène et de ce qui ne l'est pas.

La demande complexe et la distinction entre origine-destination et lien. On pourrait dire que la première procédure à définir, *Dem* (\cdot), est celle du calcul de la demande D . Elle comprend une part d'estimation de paramètres au sens économétrique classique mais aussi l'usage de mécanismes comme celui de l'affectation qui associe aux liens du réseau les quantités demandées de i à j .

Ce vecteur des demandes par mode dépend du niveau prédéterminé des *activités* A (e.g. la population dans le cas des études sur les voyageurs, le PIB dans le cas des travaux sur les marchandises), des *caractéristiques* des consommateurs ou des marchandises Y , du niveau de *service réalisé* C (e.g. le temps et la fréquence de service, le niveau des incidents et accidents). Ce service effectif est généralement différent du service planifié T et du prix effectivement payé P (fonctions des vitesses et des itinéraires réellement choisis) compte tenu des taux des taxes et péages par lien F .

L'offre d'infrastructure et de services : un vecteur planifié. D'un point de vue traditionnel de marché, la seconde procédure à considérer, *Sup* (\cdot), désigne la façon de déterminer ou de modifier l'offre, tant d'infrastructures que de services, sur un réseau souvent multimodal. Comme cette offre est en général en bonne partie centralisée, au moins pour les infrastructures, sinon pour les services, il est préférable de parler de procédure plutôt que de donner l'impression, comme en concurrence parfaite, que l'offre est définie par une simple courbe de coût avec une technologie donnée.

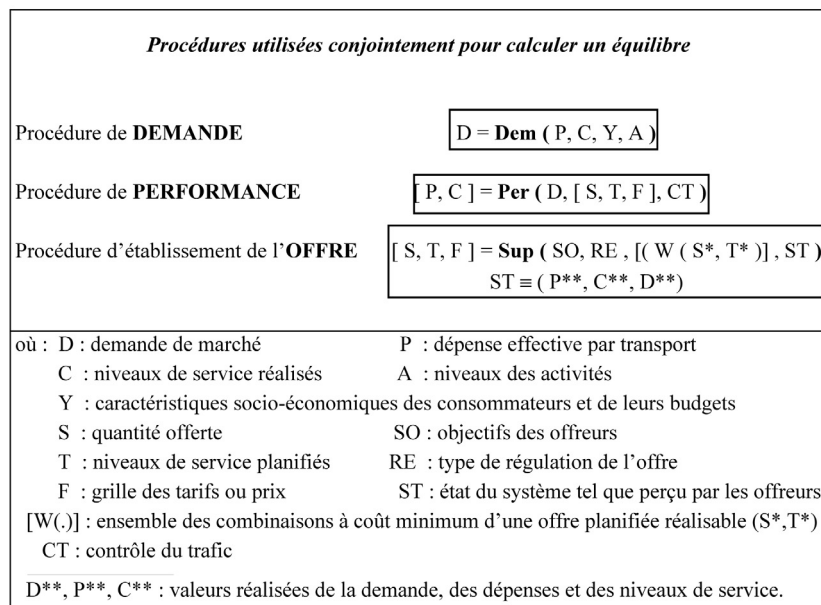
L'offre est alors, sur chaque lien donné s , un vecteur de *quantités* S (e.g. sièges-km, capacité routière), de *tarifs pratiqués* F (e.g. taxes sur le carburant ou péages) et de *niveaux de service planifiés* T (e.g. fréquence de service et temps de transport des autobus défini par une vitesse commerciale prévue ; flux libre de congestion, d'incident ou d'accident sur le lien en question).

Les facteurs qui déterminent l'offre d'infrastructure et de services sont nombreux et comprennent, parmi d'autres, ceux qui sont identi-

fiés à la Figure 1 : des objectifs SO et la réactivité même des prestataires assujettis à une régulation RE . Nous y supposons que les responsables de l'offre réagissent à l'état du système ST et sont capables de calculs complexes $[W(.)]$ pour minimiser le coût de production de diverses configurations de l'offre, même si la « fonction de production » n'est pas explicitement formulée.

Figure 1

Les trois niveaux structurels des modèles de trafic



Source la plus récente ; Gaudry, 1999.

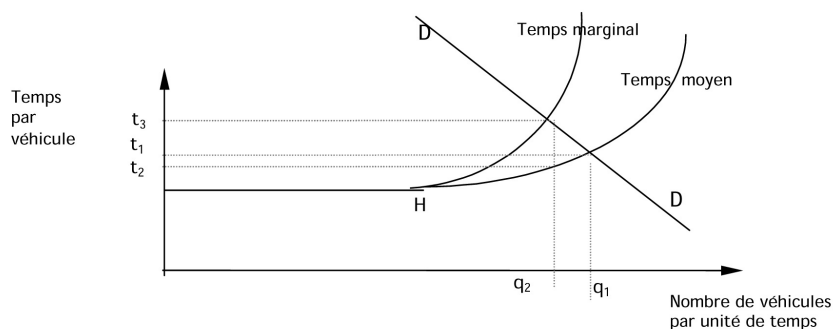
La performance réalisée. La troisième procédure, *Per* (\cdot), est « insérée » entre les deux premières qu'elle réconcilie : elle détermine la performance réelle du réseau et dépend autant de la demande effective D que de l'offre planifiée $[S, T, F]$, ainsi que de facteurs qui lui sont propres comme les règles de gestion en temps réel du réseau et le contrôle du trafic CT .

La forme la plus simple de cette procédure est celle qui détermine le niveau de congestion d'une voie de circulation en fonction de l'usage, pour un lien aux caractéristiques physiques (*e.g.* nombre de voies) et de gestion (*e.g.* régulation des feux aux extrémités, type de stationnement autorisé, signalisation et marquage) données, comme la courbe de coût moyen dessinée à la Figure 2 : au-delà d'un point H , le temps moyen croît avec le flux de véhicules, ce que l'on appelle la congestion. Les grands logiciels de planification des réseaux permet-

tent de choisir pour chaque lien la bonne courbe parmi des dizaines offertes au menu.

Équilibres. Une formulation qui utilise trois niveaux structurels, dont chacun détermine un vecteur de sorties endogènes apparaissant comme variables explicatives à d'autres niveaux, peut donner lieu à divers types d'équilibres plus complexes que l'équilibre Demande-Offre traditionnel. Dans une analyse du sens économique des courbes de débit-temps (Gaudry, 1976) qui faisait remarquer⁷ qu'elles n'étaient pas, malgré leur apparence, des courbes d'offre, nous avons distingué notamment entre les équilibres de réseau et les équilibres de marché, parmi tous les équilibres possibles de ce système à 3 niveaux.

Figure 2
Courbe débit-temps de parcours



Par exemple, un équilibre de réseau (Demande-Performance) est réalisé au point (t_1, q_1) d'intersection des courbes de demande routière et de congestion : les conditions de transport $[P, C]$ qui déterminent la quantité demandée sont bien endogènes. Le calcul de cet équilibre est normalement effectué en formulant un problème de maximisation de l'aire définie par la différence entre la courbe de demande et la courbe de débit-temps, ce qui maximise bien le surplus du consommateur⁸.

Mais cette courbe débit-temps n'est que l'enveloppe inférieure des solutions admissibles pour une firme d'utilisateurs (d'autobus, par exem-

7. Opinion qui a été immédiatement partagée par Martin Beckmann et Harold Kuhn lors de la présentation de la communication en novembre 1974. Voir un résumé partiel et quelques explicitations à l'Annexe 2.

8. Comme en situation d'équilibre général avec fonctions de coûts moyens croissants (convexes vers le bas), ce calcul du point fixe d'intersection maximise aussi simultanément la surface entre la courbe de congestion et le temps à l'équilibre, mais il ne s'agit pas ici simultanément d'un surplus du producteur car la courbe de débit-temps n'est pas une courbe d'offre, malgré son apparence, et l'usage des mêmes techniques de calcul de points fixes pour trouver l'équilibre.

ple) qui peut vouloir faire circuler ses véhicules plus lentement que ce qui serait réalisable, pour des taxis à la recherche de clients, ou pour d'autres cas plus complexes reliés à l'offre et à la demande de conduite des véhicules particuliers dans les ménages, marché où l'offre et la demande de conduite⁹ sont difficilement identifiables.

2.2.2. *Relation du système des modèles de trafic à la théorie économique*

Que penser d'un système à trois niveaux alors que nous utilisons normalement des systèmes Demande-Offre à deux niveaux ? Et que penser d'une offre vectorielle distinguant entre quantité et niveau de service pour un prix donné alors que l'offre est souvent formulée comme une simple quantité (scalaire) de qualité unique implicite ?

Niveaux et économétrie du déséquilibre. Dans la structure à trois niveaux, P et C désignent en particulier la congestion, les files d'attente ou les prix du marché noir qui équilibrent la quantité offerte à celle qui est demandée. Pourquoi ajouter ce niveau intermédiaire entre offre et demande ? Cet ajout, tel qu'on le pratique explicitement dans les transports, complique les équilibres mais évite les convolutions de l'économétrie du déséquilibre.

Jusqu'au début du 19^e siècle, la discussion économique se concentrait sur la formulation qui visait à expliquer une variable quelconque, par exemple Q, en fonction d'autres variables. Ces relations entre Q et les variables explicatives n'étaient, en l'absence de structure claire Demande-Offre, ni des formes réduites ni des formes structurelles. Elles mélangeaient les considérations propres à la Demande et à l'Offre : nous les trouvons donc confuses, voire incohérentes.

Ce n'est qu'avec Ricardo (1817) que la phrase « offre et demande » apparaît dans un titre de chapitre et que la structure moderne s'impose. Le premier à l'employer dans le contexte de la détermination des prix avait été Steuart-Denham (1787), comme nous le rappelle Groenewegen (1987), mais il faut en fait attendre Cournot (1838) pour avoir une première exposition symétrique de l'offre et de la demande. Du point de vue de l'estimation des fonctions, il fallut attendre Working (1927) pour comprendre leur identification et la façon de les recouvrer à partir des données.

Dans le domaine des transports, il a été fort éclairant au 20^e siècle d'introduire un niveau intermédiaire pour rendre compte des différen-

9. À toutes fins utiles, la demande d'usage des voitures particulières représentée par une courbe comme DD à la Figure 2 peut décrire la solution complexe d'un problème d'offre et de demande de conduite dans les ménages.

ces entre les valeurs réalisées et les valeurs planifiées du service, du taux de remplissage (confort), du prix et du niveau de sécurité. Cette nouvelle distinction s'est avérée pertinente tant au niveau théorique qu'empirique lors de l'estimation des systèmes d'équations simultanées à 3 niveaux décrivant pour la première fois le fonctionnement de réseaux de transport en commun (*e.g.* Gaudry, 1979, 1980) dans ce cadre à trois couches.

Elle est applicable à d'autres domaines, comme l'étude des économies centralement planifiées (Gaudry et Kowalski, 1990) où il est préférable de prendre partie pour l'équilibre modélisé par les queues (Kornai et Weibull, 1977) plutôt que pour « la demande insatiable » sortie d'un équilibre stationnaire vide de comportement économique structurel comme celui de Kornai (1982). Elle s'applique aussi à l'étude des marchés publics des services de santé où les queues rétablissent les équilibres¹⁰.

Elle a en tout cas l'avantage d'éviter les surprises de l'économétrie du déséquilibre dont l'usage fait conclure à Portes *et al* (1987) que l'économie polonaise était « en état d'offre excédentaire la moitié du temps durant les années 1960 et durant la période 1976-1978 » ! Pour éviter l'ironie de Podkaminer (1989) devant « une telle mine de révélations sur les économies centralement planifiées », il aurait été préférable de construire un modèle de Performance explicatif de la longueur des queues pour les voitures, le logement, *etc.* plutôt que d'étudier l'ensemble des combinaisons possibles des quantités observées selon qu'on affecte chacune sur la courbe de demande, d'offre ou sur les deux fonctions simultanément.

Formulation vectorielle des biens : demande, offre et logistique. On sait bien qu'une fonction de demande de transport qui ne comprendrait pas les niveaux de service en plus du prix serait risible, mais on ne réalise pas toujours que ce genre de formulation pratique se généralise dans l'étude microéconomique de tous les marchés pour lesquels il faut décrire des biens complexes par des vecteurs de caractéristiques, y compris temporelles (Lancaster, 1966, 1971).

Il y a longtemps que le marketing quantitatif, notamment des modes de transport¹¹, étudie précisément la demande pour les biens « non scalaires » dont les caractéristiques ont toutes une valeur et un prix : l'usage étonnant du mot « hédonique » pour désigner les consen-

10. Et on hésite devant les résultats de régression de ceux qui concluent que le nombre de médecins augmente les coûts (au lieu de réduire les queues) et qu'il faut donc le réduire.

11. Voir dans Gaudry *et al* (1998) l'histoire des modèles multiplicatifs de parts de marché dans les transports de 1966 à 1969.

tements à payer pour chacun des éléments du vecteur des caractéristiques des biens est bien plus récent et avoisine 1974 (Rosen).

Aujourd'hui, c'est la détermination du niveau de service à fournir avec le bien qui conduit et tire toute la chaîne logistique d'appui à l'offre. Par exemple, Michelin, qui vend les mêmes pneus partout, décide du temps d'attente des garages et concessionnaires dans chaque pays (3 heures au Japon, une semaine au Brésil, 2-4 jours en France et en Europe, *etc.*). Il ajuste ensuite la combinaison de modes d'acheminement et d'entrepôts à cette exigence du « produit » (Cromières, 2004) : ses 100 000 livraisons par jour dans le monde impliquent des temps d'attente très variables entre les pays, mais qui font chaque fois partie de son vecteur d'offre.

L'offre vectorielle d'un opérateur de transport public n'est pas différente, même si son niveau n'est pas déterminé par les mêmes motifs de maximisation du profit. La quantité (*e.g.* un siège-km d'un certain niveau de confort) et la qualité, entendue comme un niveau de service comprenant plusieurs dimensions dont on sait qu'elles sont importantes dans la fonction de demande (*e.g.* temps de marche, de correspondance, d'attente, de transport dans le véhicule), sont des éléments distincts qu'on ne peut résumer par un simple produit de la quantité par une qualité sans dimension, comme on le fait parfois.

3. LA SÉQUENCE CLASSIQUE À QUATRE ÉTAPES

3.1. *Grandes tendances de la modélisation du trafic depuis 50 ans*

Des premières années jusqu'en 1971. La formulation de la séquence classique si réputée et d'usage universel aujourd'hui est associée à un fourmillement d'études préparatoires aux développements des réseaux métropolitains et interurbains aux États-Unis durant les années 1950.

Les travaux séminaux les plus complets et les plus connus (CATS, 1959-1962) tirèrent profit des expériences précédentes plus limitées et des débuts simultanés de la modélisation du trafic à l'aide de méthodes analytiques requérant l'usage des ordinateurs. Ils furent conduits **à partir de 1955** sous l'égide de l'agence de planification régionale du Grand Chicago : paradoxalement, le *Chicago Area Transportation Study* (CATS) désigne en fait simultanément les trois tomes classiques publiés entre 1959 et 1962 par l'organisme de ce nom et l'organisme lui-même dont le mandat et les travaux de modélisation continuent à ce jour.

On trouve dans ces trois tomes la première séquence classique présentant, dans une étude *en coupe transversale* des déplacements de personnes en agglomération urbaine durant *une période choisie*, les quatre étapes complètes structurées et reliées dans l'ordre conventionnel : **Génération, Distribution, Choix modal, Affectation au réseau**. On y trouve aussi une modélisation très simple des effets des modifications des réseaux sur l'utilisation du sol, moins intéressante que la célèbre procédure de Lowry (1964) plus tardive et toujours en usage aujourd'hui.

Nous décrirons brièvement le contenu représentatif des quatre étapes tel qu'il s'imposa durant les quelques années suivantes et tel qu'on peut le trouver dans le premier manuel universitaire sur ces méthodes et pratiques (Meyer et Straszheim, 1971). Cet excellent manuel est encore utilisé, faute d'un successeur d'envergure comparable, en combinaison avec des traités ou articles spécialisés qui mettent l'accent sur des aspects particuliers : la demande et le choix modal (Ortúzar et Willumsen, 2001), la distribution (Erlander et Stewart, 1990) ou l'affectation (Potts et Oliver, 1972 ; Boyce *et al*, 2004).

Raffinements sans perte des repères depuis 1972. Depuis la publication en 1971 par Brookings de ce manuel représentatif de l'état de l'art, des raffinements importants ont progressivement modifié chacune des étapes, ou plusieurs d'entre elles, mais sans en changer l'esprit : on le constatera plus bas. Il faut sans doute nommer dès maintenant en particulier, en guise d'introduction : (i) le développement des modèles de **choix discret**, important après 1972 (Domencich et McFadden, 1975) ; (ii) l'usage des **formes fonctionnelles souples** dès 1976 (Gaudry et Wills, 1978) ; (iii) et un ensemble de **techniques d'affectation** plus complexes que celle du simple calcul du chemin le plus court, comme les méthodes d'équilibre (Netter, 1972 ; Nguyen, 1974), de choix discret¹² (Dial, 1971), ou même de théorie des jeux.

Toutes ces innovations, plus rigoureuses et moins *ad hoc* que l'état de l'art de 1971, s'appuyaient sur des travaux théoriques fondamentaux antérieurs, tant mathématiques que statistiques, ou sur le climat méthodologique qu'ils engendraient. Sur les fondements du **choix discret**, on pense à Finney (1947, 1964) pour l'analyse Probit et à Warner (1962) et Rassam *et al* (1970) pour l'analyse Logit. On pense aussi naturellement, pour la vision abstraite¹³ des modes qui traverse ces

12. Malheureusement, les travaux précurseurs d'Abraham et Coquand (1961) sur le Probit (et une approximation du Logit) n'étaient connus qu'en France.

13. Réduite aujourd'hui à l'imposition de coefficients génériques sur les variables comme le tarif et le temps de transport des modes, elle était novatrice en son temps puisqu'elle permettait de

modèles de choix discret, à Quandt et Baumol (1966). En ce qui concerne l'usage des *formes souples*, Anscombe et Tukey (1954), Tukey (1957) et Box et Cox (1964) sont aussi des précurseurs dont les travaux seront repris et étendus aux transports. S'agissant de *l'affectation* sur le réseau, l'ouvrage de Beckmann *et al* (1956) est fondamental du point de vue méthodologique, sinon algorithmique.

Si les contenus des quatre étapes ont été souvent lentement modifiés, suite à la formation de courants minoritaires proposant des méthodes, algorithmes et pratiques nouvelles, leurs repères sont par contre toujours reconnaissables et continuent d'encadrer la modélisation du trafic aujourd'hui. Les travaux de McFadden sur le choix discret, reconnus par le prix Nobel de sciences économiques 2000, ont été un des fers de lance de cette seconde vague de modélisation du trafic.

Voyageurs et marchandises. Centrée au début sur les problèmes d'agglomérations urbaines et de déplacement des voyageurs, la modélisation a atteint à ce jour des niveaux de détail inimaginables il y a 30 ans en utilisant des techniques de ventilation des marchés agrégés, raffinements qu'on ne pourra vraiment refléter ici. Par ailleurs, si les mêmes techniques sont utilisées aujourd'hui pour la modélisation du trafic des marchandises, c'est avec quelques adaptations rendues nécessaires par la *plus grande hétérogénéité des marchandises*¹⁴ et la *relative rareté des données* comparables à celles dont on dispose pour les voyageurs. Et, bien sûr, on n'isole pas encore directement le motif de déplacement des marchandises comme on le fait couramment pour les personnes : mais la question a-t-elle un sens ?

3.2. La génération et la distribution

3.2.1. Considérées successivement

La Génération. La première étape est celle de l'explication des *productions ou émissions* et des *attractions*, *i.e.* des vecteurs T_i et T_j

prédire la demande de services nouveaux comme par exemple le marché d'avions à décollage et atterrissage court (ADAC) reliant les centres villes par l'intermédiaire d'aéroports aménagés sur des parkings urbains jouxtant les CBD (*Central Business District*), ce qui était impossible en économie classique. En économie classique, il est impossible de dire *quoi que ce soit* sur le marché d'un nouveau bien sans expérience historique de la consommation de ce bien même si on connaît tout ce qu'on veut sur le marché des millions de biens existants. La simplification scalaire de la nature des biens adoptée au 19^e siècle, de pertinence déjà douteuse dans les transports, avait vécu.

14. Par exemple, dans un modèle interurbain, l'usage de 20 ou 30 segments du marché des voyageurs serait considéré comme généreux alors que le même nombre pour expliquer le comportement de marchandises serait considéré comme à peine adéquat au regard de la diversité des besoins et des comportements des chargeurs en termes de transport.

qui sont les marges de la matrice $O-D$ au Tableau 2, par les niveaux d'activité qui ont cours dans ces zones ($i, j = 1, \dots, Z$) :

$$T_i = p (A_{i1}, \dots, A_{ia}, \dots, A_{iA}) + e_i \quad (7)$$

et

$$T_j = a (A_{j1}, \dots, A_{ja}, \dots, A_{jA}) + a_j \quad (8)$$

où, dans le cas des voyageurs, les activités dominantes sont représentées par la population en (7) et les emplois en (8) et, dans cas des marchandises, par des variables comme la valeur ajoutée ou l'emploi. Ces équations contiennent généralement très peu de variables (souvent une ou deux) et sont toujours linéaires. Sauf erreur de notre part, on n'utilise encore pour ces régressions que les valeurs non nulles des vecteurs de Z observations : nous n'avons jamais vu de Tobit utilisé à cette étape pour exploiter les observations nulles qui, elles aussi, contiennent de l'information.

Plus étonnant, elles ne contiennent généralement pas de variables sur la disponibilité du parc de voitures particulières ou d'autres équipements de transport et aucune variable de prix ou de distance ou de temps de transport. La pratique de cette étape n'a guère changé depuis ses origines malgré quelques essais épars et minoritaires d'introduire dans les équations (7) et (8) des variables d'accessibilité ou de centralité qui permettraient ensuite de modéliser à cette étape¹⁵ un effet de retour des modifications de réseau.

Du point de vue de la *théorie économique*, les résultats sont interprétables comme des coefficients d'input ou d'output fixes : on explique la production et le besoin en « déplacements » ou en « tonnes ». Mais on n'explique pas (encore) des passagers-km ou des tonnes-km : on ne sait pas où vont les émissions et comment seront satisfaits les besoins définis par l'attraction, car on répondra à ces questions à la seconde étape.

Il faut avouer que les travaux des économistes pour représenter le rôle des transports dans la fonction de production des entreprises en sont toujours à leurs premiers balbutiements : le transport n'est pas (encore) un facteur de production et il n'a aucune productivité propre, et encore moins de productivité différenciée par mode. Il n'y a pas encore de substitution entre les formes de transport et les stocks ou d'autres intrants comme les communications, ce qui implique que les

15. L'usage de variables d'accessibilité dans les procédures explicatives de l'utilisation du sol est autre chose et concerne un autre niveau de la modélisation.

flux tendus sont sans justification théorique, sauf l'exception visible de Baumol et Vinod (1970).

Lors d'un examen que nous avons fait en 2000 de plus de 800 modèles calculables d'équilibre général, aucun de ceux qui incorporaient le transport (moins de 2 %) ne le traitaient autrement que comme une « taxe » : ce n'était un facteur de production pour aucun économiste.

Il apparaît comme facteur de production et de localisation dans nombre d'études de développement régional mais son rôle y est toujours *ad hoc* et d'interprétation difficile parce que la mesure du capital transport est alors commune à toutes les activités de la zone et ne fait pas partie au sens strict, mesurable et différencié, de la fonction de production des firmes et des ménages ; ces fonctions sont d'ailleurs souvent moins bien formulées que (7) et (8) et leur identification économétrique est souvent au mieux ignorée.

La distribution. Les sources T_i et les besoins en personnes et tonnes T_j établis, on passe à l'étape de la distribution qui expliquera en fait la **distribution des longueurs** des déplacements, par une modélisation sous contrainte des flux de la matrice O-D elle-même. On part de :

$$T_{ij} = g(T_i, T_j, U_{ij}) + \varepsilon_{ij} \quad (9-A)$$

qu'on reformule ensuite pour refléter l'imposition des doubles contraintes décrites plus bas :

$$T_{ij} = g_c([\hat{E}_i, \hat{A}_j], T_i, T_j, U_{ij}) + \varepsilon^*_{ij} \quad (9-B)$$

où U_{ij} est une mesure de l'utilité (ou de la désutilité) associée au parcours de i à j et $[\hat{E}_i, \hat{A}_j]$ désignent les facteurs d'ajustement qui font respecter les doubles contraintes, respectivement sur les émissions et les attractions. Dans les premiers modèles, la distance de i à j est utilisée comme métrique de U_{ij} mais des mesures plus complexes, comme le coût généralisé (une combinaison habituellement linéaire d'argent et de temps) ont rapidement vu le jour pour cette « impédance ».

La forme fonctionnelle universellement utilisée pour (9) est celle d'un modèle gravitaire, donc multiplicatif, dont le seul paramètre estimé est celui de U_{ij} , parce que les paramètres de puissance des « masses » à l'origine et à la destination y sont arbitrairement fixés à 1. En milieu urbain, les élasticités ainsi estimées par rapport à la distance varient généralement de -1 à -3 selon le motif de déplacement des personnes : le motif travail est celui pour lequel la distance freine le moins et le motif personnel celui pour lequel il freine le plus ; en milieu

interurbain, elles sont supérieures en valeur absolue à -1 tant pour les personnes que pour les marchandises.

Une dimension fondamentale de l'estimation de ce paramètre de U_{ij} est le fait qu'elle soit faite sous les doubles contraintes suivantes pour que les valeurs calculées T_{ij}^* fournissent une destination à toutes les émissions et satisfassent exactement à tous les besoins d'attraction :

$$T_i = \sum_j T_{ij}^* = O_i \quad (10)$$

$$T_j = \sum_i T_{ij}^* = D_j \quad (11)$$

En conséquence, puisque les « marges » de la matrice sont reproduites, l'estimation ne produit qu'une distribution des longueurs des flux, ce qui fournit le kilométrage « manquant » aux quantités de personnes ou de tonnes modélisées à l'étape précédente. Ces contraintes sont souvent jugées comme étant raisonnables et un effort d'importance est fait pour les imposer lors de l'estimation, surtout depuis que Wilson (1967, 1970) a dérivé ce modèle à doubles contraintes de la maximisation de l'entropie.

Du point de vue de la *théorie économique*, toutefois, l'assujettissement à ces doubles contraintes impose, peut être involontairement, une structure arbitraire très particulière aux erreurs de régression ε_{ij}^* en (9-B) et nous pensons que l'estimateur est de ce fait biaisé.

Mais il engendre de surcroît et presque involontairement une propriété très particulière du modèle (9-B) en apparence gravitaire : suite à des modifications des coûts de transport sur une liaison ij , les flux seront réaffectés sur d'autres chemins pour satisfaire aux doubles contraintes : les termes \hat{E}_i et \hat{A}_j sont en effet fonctions de l'ensemble des $\{U_{ij}\}$ et contiennent donc autre chose que des utilités U_{ij} associés à la paire ij considérée, même si cela n'est pas apparent en (9-B). Il se produira donc une substitution entre les chemins qui changera la structure spatiale de la matrice $O-D$. Cette propriété automatique disparaît naturellement si on n'impose pas les doubles contraintes, dans quel cas on revient à (9-A), ou si on n'en impose qu'une seule.

En fait, les doubles contraintes transforment profondément un modèle de demande où, en apparence, la taille du flux pour une paire origine-destination ij ne dépend que de ce qui se passe en i , en j , ou en ij . Plus précisément, en (9-A), le coût de transport de i à j n'influence que le flux sur cette liaison et n'a aucun effet sur les autres flux ik ou nj . Les doubles contraintes modifient cette propriété d'utilité séparable entre les flux sans que la structure de la substitution-complémentarité de (9-B) n'ait été explicitée autrement que comme une obligation comptable. Pour ces raisons, beaucoup de chercheurs préfèrent

regrouper les deux premières étapes en une seule et faire jouer aux prix un rôle clair sur le niveau et la longueur des flux justement en partie parce qu'ils n'utilisent pas les doubles contraintes.

3.2.2. *Combinées en une étape*

Si on remplace les « marges » par leurs facteurs explicatifs dans (9), et si on ajoute une variable b_{ij} indicatrice de la présence éventuelle d'une frontière entre les zones considérées, le modèle devient :

$$T_{ij} = g(A_i, A_j, U_{ij}, b_{ij}) + \varepsilon_{ij}, (i, j = 1, \dots, Z) \quad (12)$$

qui est en principe de forme quelconque et vise l'explication conjointe du niveau et de la distribution des flux.

Or il existe depuis longtemps pour ce modèle une pratique de la forme multiplicative, dite aussi « gravitaire », mais sans doubles contraintes, comme chez Newton¹⁶ (1684). Déjà Lill (1891) utilisait un modèle explicitement gravitaire des flux interurbains de voyageurs par train, comme l'ont fait approximativement nombre d'études de l'entre-deux guerres dont l'étude sur les marchandises de Morgenstern (1936) qui établit des flux internationaux en tonnes et en valeur avant de discuter du choix modal. Appliqué aux flux interrégionaux et internationaux, le modèle s'est enrichi très tôt d'une mesure des « effets-frontières » dans l'explication des flux de transport et, un peu plus tard, dans l'explication des flux commerciaux et de communication.

Le succès empirique de ce modèle dans sa forme multiplicative est extraordinaire : des milliers d'applications, tant voyageurs que marchandises (à condition que leur niveau d'agrégation des flux soit suffisant), ainsi qu'aux flux commerciaux en valeur – et à toutes les échelles (urbaine, interurbaine, internationale). On peut dire qu'aucun modèle économique n'a eu plus de succès que le modèle gravitaire.

Du point de vue de la *théorie économique*, il semble compatible avec beaucoup de théories du commerce international. Toutefois, le fait que les flux indicés de i à j ne dépendent que de variables indicées en i , en j ou en ij , en fait une structure compatible avec l'axiome IIA (*Independence from Irrelevant Alternatives*) de Luce qui exige que les rapports de probabilité de choix entre deux « alternatives » ou choix ne dépendent que de leurs caractéristiques en propre et pas de celles des autres possibilités. Du point de vue de l'interprétation économique, il s'ensuit qu'il n'y a aucune concurrence ou complémentarité

16. Chez Newton, la constante gravitationnelle G , appelée aussi constante universelle de gravitation, comme son nom l'indique, ne varie pas par paire de corps considérée. Le modèle (12) comprend une constante de régression.

entre les flux (les marchés origine-destination) puisque le prix applicable à une liaison n'affecte que le flux sur cette liaison : le prix du transport Paris-Londres est sans effet sur le flux Paris-Rome, et inversement.

Des efforts considérables ont été faits ces dernières années pour évaluer la robustesse de ce modèle du point de vue de (i) l'usage de prix précisément autres que les prix associés à la liaison considérée ; (ii) la forme fonctionnelle ; (iii) la présence de corrélation spatiale entre les erreurs associées à divers flux ; (iv) le relâchement des contraintes unitaires sur les variables d'activité comme la population ou le PIB.

En particulier, ses travaux très généraux sur des flux voyageurs en Allemagne et au Canada (Gaudry *et al.*, 1994), sur les flux internationaux de voyageurs européens à l'intérieur de l'Europe des 15 (Gaudry *et al.*, 1998b) et sur bases de données du commerce interrégional européen et nord-américain (Gaudry, 2004) ont montré que : (i) la forme multiplicative des variables d'activité était une excellente approximation, très proche de la forme optimale, mais qu'imposer à ces variables des valeurs d'élasticités unitaires nuisait à l'ajustement, qui pouvait révéler une certaine différence entre les élasticités associées à la production et celles qui sont associées à l'attraction ; (ii) l'élasticité du flux par rapport au terme d'utilité globale¹⁷ U_{ij} était mieux estimée avec des formes mathématiques plus souples qu'avec un simple produit, ce qui était déjà connu (Kau et Sirmans, 1979) ; (iii) que l'insertion¹⁸ de prix autres que les prix associés à la liaison considérée était parfois possible mais que les contributions de ces nouvelles variables étaient plutôt faibles et éloignaient peu des propriétés IIA de la structure de base. On a raison de se demander s'il existe un modèle plus robuste dans l'ensemble de la littérature économique.

D'ailleurs, ce modèle conjoint de « génération-distribution », agrégé dans sa forme multiplicative simple, est toujours au cœur de nombreuses procédures d'estimation de la demande, par exemple Arduin (1989). Sa performance est excellente même quand la distance est utilisée comme mesure de U_{ij} car la distance est corrélée avec les indicateurs les plus sophistiqués de l'utilité de l'ensemble des modes offerts sur les liaisons, comme le *logsum* que nous définirons sous peu. Plusieurs modèles ont conservé une modélisation agrégée de la

17. La formulation et le mot *impédance* (au signe anticipé négatif) ont progressivement été remplacés par une formulation centrée sur l'*utilité* (au signe anticipé positif).

18. Directe ou indirecte par le truchement de la corrélation spatiale entre les résidus de régression.

génération-distribution, même s'ils l'ont combinée avec des modèles de choix discret des étapes ultérieures.

3.3. Le choix modal et l'affectation

3.3.1. Considérés successivement

Le choix modal, d'abord agrégé. La plupart des déplacements sont, au sens strict, intermodaux parce qu'ils se font par une combinaison de modes (e.g. accès à pied et autobus ; taxi et avion) analogue à celle du transport (combiné ou) intermodal des conteneurs et caisses mobiles. Dans ces conditions, si le problème du choix entre modes pour aller d'une origine à une destination est formulé comme un choix entre modes *principaux*, il est entendu que choisir un mode principal n'exclut pas l'usage d'autres modes pour une petite portion du trajet, ou qu'on peut définir pour les cas limites des combinaisons particulières de modes principaux articulés en configurations complémentaires (e.g. *park and ride*, *kiss and ride*).

Formulés dans cette perspective des modes principaux, les premiers modèles de choix modal, formulés pour expliquer le choix entre le transport en commun (TEC) et la voiture particulière (VP), utilisaient au milieu des années 1950 des fonctions assez simples établissant une relation, souvent variable en fonction du niveau de motorisation donné dans la zone d'origine, entre la part de la voiture et le ratio des temps de transport par les deux modes considérés, ce qu'on peut écrire :

$$(\text{part}_{\text{VP}})_{ij} = f \left(\left[\frac{(\text{temps}_{\text{VP}})_{ij}}{(\text{temps}_{\text{TEC}})_{ij}} \right], [\text{taux d'équipement VP}]_i \right) + \mu_{ij} \quad (13)$$

où le rapport des temps de transport VP/TEC trace pour la part une « courbe de diversion » ou de détournement.

La forme linéaire. Typiquement linéaires et estimées par la méthode des Moindres Carrés Ordinaires (MCO), ces fonctions posaient des problèmes de prévision. La simulation des effets d'une modification du ratio des temps de transport ou d'une autre variable (ou leur combinaison) pouvait¹⁹ prédire une part négative ou une part plus grande que 100 %.

Par ailleurs, le genre de formule (13) se prêtait mal aux cas où plusieurs modes sont en concurrence, même si elle pouvait être élargie pour tenir compte des prix des modes. On est donc rapidement passé à des formulations qui, par construction, rendaient impossibles des pré-

19. Ce qui est impossible par construction dans l'échantillon utilisé pour l'estimation, mais normal dès qu'on sort du domaine échantillonnal à des fins de simulation ou de prévision.

visions de parts négatives et ne sommant pas à 100 % tout en accommodant en même temps plusieurs modes, chacun d'eux étant caractérisable par un vecteur de propriétés (*e.g.* tarif, temps de trajet, fréquence).

La forme multiplicative. La première formulation complète et cohérente qui s'imposa, surtout en modélisation interurbaine, fut le modèle de parts multiplicatif (McLynn *et al*, 1968) et ses diverses variantes, en fait identique au modèle « *Multiplicative Interactive Competition* » (MIC) déjà connu de la littérature du marketing quantitatif :

$$p_i = \frac{\prod_k X_{ik}^{\beta_{ik}}}{\sum_{j \in C_n} \left(\prod_k X_{jk}^{\beta_{jk}} \right)} \quad (14)$$

et qui permettait d'expliquer la part p_i du mode i ($i, j = 1, \dots, M$) à l'aide de ses k caractéristiques X_{ik} en comparant son « attractivité » à celle de l'ensemble des possibilités regroupées au dénominateur.

Rapidement, ce modèle devint par ailleurs le cœur de la modélisation des flux commerciaux (Armington, 1969) et il est encore utilisé dans ce contexte, en particulier dans les modèles calculables d'équilibre général des flux commerciaux internationaux, car il autorise pour chaque type de bien des importations et des exportations **sur le même lien**. En dépit de certaines théories sur la spécialisation internationale, comme Heckscher-Ohlin-Samuelson (H-O-S), on observe que la plupart des biens (*e.g.* voitures, machines-outils, nourriture) sont à la fois exportés et importés par beaucoup de pays et se comportent *de facto* comme les flux modaux **bidirectionnels** des modèles de transport et de communication sans qu'on ait vraiment besoin de la concurrence monopolistique entre produits légèrement différenciés pour le comprendre.

La forme du Logit Multinomial Linéaire. Peu après une véritable vague de ces modèles multiplicatifs, un autre modèle de parts apparut, qui évitait aussi les inconvénients de la forme linéaire de (13) : le modèle Logit multinomial *appliqué à des parts*. Ce modèle :

$$p(i) = \frac{\exp(V_i)}{\sum_{j \in C} \exp(V_j)} \quad (15)$$

comprend des fonctions « d'utilité représentative » (pour employer la terminologie postérieure qui s'imposera après 1975 dans le cas d'*applications aux choix discrets*) linéaires :

$$V_i = \beta_{i0} + \sum_n \beta_{in}^i X_n^i + \sum_s \beta_{is} X_s + u_i \quad (16)$$

où les variables explicatives utilisées dans chaque fonction V_i sont partagées entre des variables X_n qui la décrivent et varient en conséquence entre les modes (*e.g.* tarif, temps de marche, temps d'attente et temps dans le véhicule), et des variables X_s socioéconomiques communes qui ne varient pas entre les modes ou « alternatives » (*e.g.* le revenu ou le sexe des voyageurs). L'indice de mode est utilisé en exposant afin de faciliter plus loin une écriture plus générale qui autoriserait que les caractéristiques de tous les modes, comme leurs prix, soient utilisées dans toutes les fonctions d'utilité représentative (dont l'identifiant se trouve en indice).

Utilité et armature du modèle de demande. Du point de vue de la *théorie économique*, les auteurs de ce modèle de parts (Ellis et Rasmussen, 1970) reliaient pour la première fois l'interprétation des fonctions (16) à l'utilité, relation qui allait sous peu être rendue biunivoque par les démonstrations de Domencich et McFadden.

Ces démonstrations permettraient de déduire la forme même du modèle de demande à partir d'hypothèses sur la distribution du terme d'erreur de (16) plutôt que d'accepter la forme d'un modèle de demande sur une base *ad hoc*. Elles établiraient aussi qu'il y a relation biunivoque entre par exemple : (i) la forme Logit (15) et des erreurs qui obéissent à des lois de distribution Weibull (1939, 1951) indépendantes entre les fonctions d'utilité représentatives ; (ii) la forme Probit et des distributions normales ou gaussiennes (Gauss, 1823) des mêmes erreurs ; (iii) ... et ainsi de suite pour la distribution arc tangente et d'autres distributions. Une révolution à plusieurs égards²⁰.

20. Du point de vue de l'histoire de la théorie économique, on peut constater que c'est la deuxième fois que la résolution des problèmes pratiques de modélisation de la demande de transport apporte une contribution fondamentale aux sciences économiques. Toutefois, dans le premier cas, celui de Dupuit (1844) qui est un auteur d'inspiration et de tendance marginaliste, l'utilité du consommateur est vraisemblablement déjà d'esprit plutôt *ordinal* alors qu'ici elle est fermement de type *cardinal*. Cela signifie que l'industrie qui s'est développée depuis 1975 sur la base du modèle Logit utilise tous les jours de manière convaincante des modèles qui nient en pratique une conquête de la théorie de la demande du 19^e siècle, celle de l'optimum du consommateur calculé sans référence au niveau de l'utilité. On abandonne ainsi la perspective ordinale qui ne s'intéresse qu'aux ratios des utilités marginales sans jamais en mesurer le niveau et on la remplace par une mesure très précise du niveau de l'utilité des alternatives qu'on discute à la cinquième décimale près ! Est-ce un progrès ou autre chose ?

Cette révolution : (i) inséra l'aléatoire dans la formulation même des fonctions d'utilité, traditionnellement déterministes ; (ii) établit l'existence d'un lien analytique *biunivoque*, et donc nécessaire, entre la forme de la fonction de demande et la formulation de fonction d'utilité sous-jacente dont elle était dérivée au sens strict, lien sans précédent comparable dans l'analyse des systèmes de demande où dominait alors le modèle de Rotterdam (Barten, 1969) parmi divers systèmes complets de demande (Brown et Deaton, 1971).

Comme l'a fait remarquer lui-même Deaton (1974) dans une étude qui lui valut le Prix Ragnar Frisch de 1979, la pratique de l'estimation des fonctions de demande était toujours prisonnière du dilemme suivant : soit on choisissait une fonction d'utilité, et alors la forme de la fonction de demande était indéterminée, soit on imposait à des fonctions de demande *ad hoc* et arbitraires des contraintes qui les rendaient théoriquement conformes aux propriétés souhaitées de l'une ou l'autre fonction d'utilité, et alors la liste des variables était indéterminée car on ne savait pas quelles variables paramétrer²¹. Mais on ne *dérivait* pas²² la forme exacte de la fonction de demande de ces fonctions d'utilité déterministes qui supposaient une parfaite connaissance des préférences et définies sur des biens « scalaires ». McFadden établissait un tandem utilité-demande en affirmant : « La forme de la fonction de demande est *cela si et seulement si* la fonction d'utilité est *ceci* ; et inversement. »

Émergence des choix discrets. Si toutes les formulations mentionnées dans les équations de (13) à (15) sont bien des modèles de parts parce qu'elles sont appliquées à des données *agrégées*, l'étude de *Charles River Associates* de 1972 (Hall et McFadden, 1972) remaniée²³ et publiée ensuite comme livre (Domencich et McFadden, 1975), non seulement resserrait le lien avec l'utilité décrit plus haut mais estimait

21. Pour citer le texte lui-même : « [There is] a basic dichotomy between utility and demand which runs through all of the empirical literature. Demand models may be derived either by the selection of a utility function or by the arbitrary specification of a system of equations which may then be modified according to the utility theory. If we choose the first course of action, we have the difficulty of the **choice of functional form** ; if we choose the second, we have the difficulty of **choice of variables** to parametrize. In both cases, because of the mathematical complexity of a conversion from one to the other, a decision may have unexpected or untoward consequences » (*op. cit.*, p. 342).

22. Par contre, la théorie de la dualité établissait déjà bien des points de passage entre utilité et demande, mais la relation entre la forme de la fonction d'utilité et celle de la fonction de demande n'y est pas toujours aussi clairement bidirectionnelle que dans le cas du modèle Logit ou Probit.

23. La remarquable annexe qui montre les difficultés de formuler des fonctions de demande modales agrégées quand les biens sont de nature vectorielle n'a malheureusement pas été reproduite dans le célèbre livre. Pour un exposé des problèmes rencontrés durant les années 1960 avec les fonctions agrégées en milieu urbain et un résumé des principales formulations, voir la première partie du rapport fait dans le cadre du projet européen STEMM (Gaudry *et al.*, 1998).

les paramètres de modèles Logit linéaires aux fonctions (16) réinterprétées avec données individuelles ou *discrètes*, données qu'avait utilisées Warner (1962) dix ans avant cette réinterprétation.

Dans la foulée, confortée par l'explicitation du lien avec l'utilité, une vague considérable de travaux firent du Logit Linéaire, appliqué le plus souvent aux choix discrets (individuels), la base d'une reprise du processus classique qui, commençant avec succès par le choix modal, s'étendit bientôt aux quatre étapes de la séquence classique, mais alors avec un succès bien moindre, à mesure que les questions posées se rapprochaient de l'explication de la décision de se déplacer et remontaient une à une les marches des étapes précédentes.

En effet, s'il est facile de décrire les caractéristiques des modes qui n'ont pas été choisis (ce qui est essentiel à la calibration du modèle de choix modal), il est beaucoup plus difficile de décrire (dans un modèle de distribution) les caractéristiques des déplacements pour motif achats ou travail qui *n'ont pas été faits*, et plus difficile encore de faire (pour un modèle de génération) la liste des caractéristiques des activités concurrentes aux déplacements effectués mais qui, elles, n'ont pas été réalisées.

Quant à l'affectation, l'usage du Logit (Dial, 1971), dont *la Loi d'Abraham* (e.g. Abraham, 2001) est en France un cousin proche²⁴, n'y fut jamais dominant comme il l'était et l'est toujours pour le choix modal, car d'autres méthodes étaient considérées comme plus convaincantes à divers égards ou, en situation de congestion, plus favorables du point de vue algorithmique et du coût de calcul.

La floraison baroque de 1977. On peut dériver le Logit Multinomial de plusieurs manières, mais aucune n'a eu autant d'influence que celle de Domencich et McFadden dont le succès est sans doute imputable au rapport bidirectionnel intime établi entre l'armature du modèle et l'utilité aléatoire²⁵.

La vague du Logit a aussi été favorisée par la facilité relative d'estimation de ses paramètres, par opposition par exemple à ceux du

24. Contrairement à ce qu'on a pu dire ou écrire (Leurent, 1999), ce modèle cesse d'être un « Logit logarithmique » dès que la mesure d'impédance qu'il utilise possède plus d'un terme : par exemple, l'usage d'une combinaison linéaire de temps et d'argent en fait un modèle unique qui n'est plus un cas particulier du Logit Box-Cox Standard [(15) et (19)]. Voir le détail dans Gaudry (2006).

25. On peut obtenir le Logit comme sous-produit de la dérivation du *modèle gravitaire* de Wilson (1967) : il y apparaît comme l'arrangement le plus probable si on impose une contrainte de coût total à la maximisation de l'entropie. On peut aussi le trouver en réarrangeant le modèle « de *demande aléatoire rationnelle* » de Theil (1975), comme l'a souligné Truong (1981). Enfin, Leonard (1982) l'a dérivé de la *théorie asymptotique des extrêmes*.

Probit qui comporte des intégrales multiples²⁶ sans pour autant fournir des résultats très différents, et par le développement de trois compléments qui ont émergé simultanément en 1977 : (i) celui des *structures hiérarchiques emboîtées* ; (ii) celui des *formes fonctionnelles souples* ; (iii) et celui des *coefficients de régression aléatoires*. Quelques mots sur chacun, dans l'ordre de leur popularité à ce jour, afin de pressentir leur rôle considérable dans les modèles de trafic et leurs propriétés en situation de prévision pour un projet : on ne peut comprendre ce comportement sans prendre conscience des trois compléments.

Hiérarchisation des choix : du Logit Multinomiale au Logit Hiérarchique. Le plus connu de ces compléments, bien décrit dans les manuels (*e.g.* Ortúzar et Willumsen, 2001), est celui des hiérarchies de modèles emboîtés (*nested*). Elles ont été créées principalement pour pallier²⁷ cette insupportable propriété du modèle Logit qui est l'égalité des élasticités croisées de la demande.

Cela signifie que, si on utilise la formulation multinomiale (15)-(16), l'implantation d'un service d'autocars sur horaire ou d'une piste cyclable entre Paris et Lyon obtiendra toujours son marché en capturant sur cette liaison la même proportion de chacun des modes existants (train, avion, voiture). De même, l'ajout d'un nouvel itinéraire dans un modèle de choix entre trajets par un même mode (et en l'absence de congestion) réduira dans une même proportion les flux des itinéraires existants.

En formulant une hiérarchie à deux niveaux, on contourne en partie le problème : si, par exemple, on suppose qu'au niveau supérieur le choix se fait entre l'avion, la marche à pied et l'ensemble des modes terrestres motorisés, et qu'au second niveau il ne se fait qu'entre ces modes terrestres, qu'arrivera-t-il ? L'élasticité croisée de la part de marché sera la même entre les modes d'un niveau donné, mais elle ne sera plus identique entre chacun des modes du niveau inférieur et un mode appartenant au niveau supérieur : dans notre exemple, elle ne sera plus identique entre tous les modes terrestres motorisés particuliers et l'avion ou la marche à pied. Il faut donc,

26. Dans un des modèles nationaux suédois formulés par Staffan Algers, la région de Stockholm comprend 80 zones où l'on peut faire ses achats. Si l'on arrive à estimer un Logit avec 80 alternatives, c'est bien autre chose d'estimer un Probit, même linéaire. Et que dire dans le cas de villes plus grandes et de fonctions d'utilité non linéaires.

27. La justification technique met à contribution les corrélations supposées entre les erreurs non observées associées à chaque fonction d'utilité de type (16). Malheureusement, les nombreuses et diverses hiérarchies ainsi concevables ne sont pas strictement comparables par des tests statistiques : sans véritable mesure commune et continue entre elles, elles semblent fonctionner ou ne pas fonctionner dans des situations-types.

parmi toutes les hiérarchisations données, en formuler certaines qui donneront des résultats raisonnables.

L'instrument critique et la pierre d'angle de cette hiérarchisation est la découverte en 1977, simultanément et indépendamment, par McFadden²⁸ et par Williams (1977), que l'espérance maximale de l'utilité de l'ensemble des alternatives (en gros ce que l'on peut attendre de mieux du menu des choix disponibles) est égale au logarithme naturel du dénominateur du modèle Logit (le *logsum*), expression appelée aussi officiellement *valeur inclusive I* des alternatives considérées :

$$I \equiv \{ \ln(\sum_m e^{V^m}) \}. \quad (17)$$

On insère alors cette valeur inclusive, une sorte d'indice agrégatif non linéaire, comme variable explicative du choix au premier niveau. Cette opération peut en principe être répétée si l'on a un troisième niveau.

Les hiérarchies sont d'usage commun aujourd'hui et il n'est guère de modèles un peu complexes, définis pour une agglomération ou au niveau national, qui n'en comprennent. Elles sont très utilisées pour comparer l'attractivité de diverses destinations commerciales, chacune se voyant attribuer, à partir d'une origine donnée, son propre « panier » de modes de transport (la valeur inclusive de l'utilité des modes de transport utilisables pour y accéder). Si la très grande majorité des modèles hiérarchiques sont de forme linéaire (16), un nombre grandissant utilisent des formes souples (*e.g.* Deepack et Laferrière, 1994, ou Lapparent, 2006), thème de notre prochain paragraphe.

Formes fonctionnelles souples Box-Cox. Une autre façon d'améliorer le Logit Multinomial, cheville ouvrière du choix modal et étape généralement décisive de toute étude de la demande de transport, est en effet de soumettre à l'épreuve des données l'hypothèse de linéarité de (16) en utilisant à cette fin la transformation non linéaire la plus utilisée en économétrie selon Davidson et MacKinnon (1993), la transformation de Box et Cox (1964). Elle est définie pour toute variable positive Var_v , comme²⁹ :

28. Dans une communication présentée à la *Third International Conference on Behavioural Travel Modelling* (Tanunda, Australie) en avril 1977 et publiée l'année suivante (McFadden, 1978).

29. Nous ignorons donc ici le coefficient μ défini dans l'article de Box et Cox (1964) et nous en tenons à la définition commune. Par ailleurs, lorsque la variable dépendante y d'un modèle n'est pas transformée mais que ses variables explicatives le sont, on dit parfois Box-Tidwell (1962) plutôt que Box-Cox, mais nous négligerons cette pratique. Il est aussi possible, à certaines conditions, de transformer une variable continue (autre qu'une constante) qui contient des observations nulles à condition d'ajouter à la variable transformée une indicatrice compensatoire associée aux observations nulles.

$$Var_v^{(\lambda)} \equiv \begin{cases} \frac{(Var_v)^\lambda - 1}{\lambda} & , \text{ if } \lambda \neq 0, \\ \ln(Var_v) & , \text{ if } \lambda \rightarrow 0. \end{cases} \quad (18)$$

ce qui permet de réécrire (16) dans la forme Logit Box-Cox Standard :

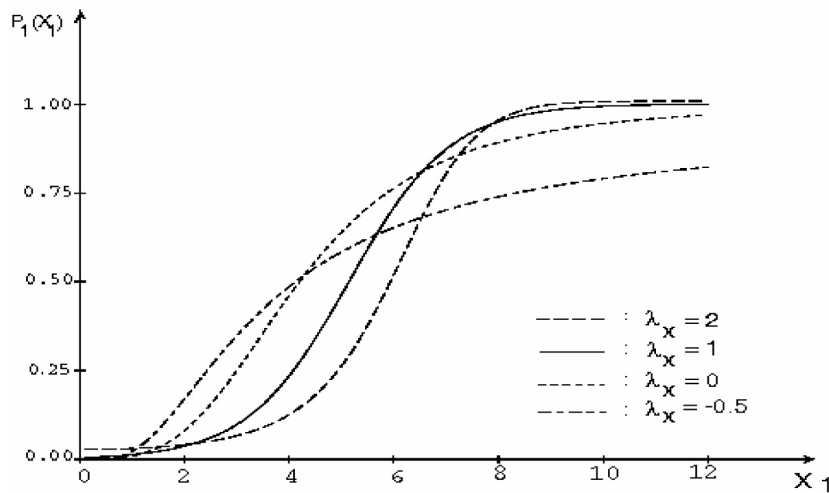
$$V_i = \beta_{i0} + \sum_n \beta_{in}^i X_n^{i(\lambda_{in}^i)} + \sum_s \beta_{is} X_s^{(\lambda_{is})} \quad (19)$$

où nous avons négligé les indices t des observations pour alléger l'écriture. Aux qualités suivantes – un meilleur ajustement aux données, des propriétés théoriques plus générales que celles de la forme linéaire et la possibilité de modifier profondément les résultats de régression en affectant jusqu'aux signes³⁰ des coefficients des variables, s'ajoute le fait que l'usage des transformations rend les valeurs du temps variables selon le point de référence du tarif et du temps et la quantité de temps épargné : la première minute gagnée a une valeur qui dépend de la base de la modification et les minutes suivantes n'ont plus la même valeur que la première. On constate aussi que le Logit Box-Cox (15), (18)-(19) comprend comme cas particulier le modèle multiplicatif précédent (14).

Une autre manière de comprendre le rôle de la non linéarité des variables est de constater qu'elle implique que la courbe de réaction de la part modale par rapport à une modification d'une variable qui nous intéresse, comme la vitesse de transport du mode i , n'est plus symétri-

30. Il faut se rappeler qu'en régression multiple le signe du coefficient d'une variable explicative ne dépend pas seulement de la corrélation simple entre cette variable et la variable dépendante, ou de sa variance propre : elle dépend aussi de la covariance entre toutes les variables explicatives. Comme les variables transformées voient leurs variances individuelles et leurs covariances modifiées par ces transformations, c'est l'existence même de la corrélation statistique qui est déterminée par leur usage. En effet, il est tout à fait possible, comme le savent tous les praticiens de la régression tant classique que logistique, que le signe d'une variable explicative soit positif et significatif si la variable est insérée de manière linéaire mais que ce même signe devienne négatif et significatif si la variable est insérée sous forme logarithmique. Pour des exemples de tels changements extraits de la littérature économique et des transports, tant en régression logistique que classique, voir Gaudry *et al* (1998). Plus généralement, seules les données peuvent décider de la forme optimale et donc conjointement de l'existence de la corrélation statistique (de sa force, de son signe et de son degré de fiabilité) et de celle de la forme de la relation, existences toutes deux douteuses en l'absence de tests conjoints.

Figure 3
Logit Linéaire vs Logit Box-Cox



que par rapport à un point d'inflexion localisé à $p_i = 1/2$, comme l'indique la courbe en gras du cas linéaire sur la Figure 3.

Ce schéma illustre, dans une représentation à deux modes, l'effet sur la probabilité de choix p_i de diverses valeurs d'une transformation Box-Cox appliquée à une variable X_i . Si en effet on croit que le phénomène qui nous intéresse (réaction à la vitesse améliorée d'un mode de transport, *etc.*) implique une représentation asymétrique de l'effet sur la probabilité qui nous intéresse, alors la variable explicative responsable de cet effet n'intervient pas linéairement dans la fonction d'utilité du modèle Logit, et cela indépendamment du nombre de modes considéré. Toutefois, la même variable explicative peut induire une ***courbure de forme différente selon le mode*** considéré et spécifique à chacun d'entre eux.

Du point de vue ***économique***, on peut dire que la linéarité représente un cas de rendements marginaux constants parce que, par exemple, la minute marginale du temps de transport a alors le même effet sur la probabilité de choix du mode concerné indépendamment de la longueur du déplacement, contrairement à ce qui se passe dans le modèle Logit Box-Cox Standard caractérisé par (19). C'est tout différent dans un modèle non linéaire et l'impact sur la valeur du temps peut être décisif tout en redressant des résultats linéaires invraisemblables (Gaudry *et al*, 1989). Du point de vue économique on pourrait penser, tant en théorie de la production qu'en théorie de la consomma-

tion, que la linéarité est, comme dans la nature, la moins vraisemblable des formulations³¹.

Généralisation du Logit Box-Cox Standard. Élaboré en 1976-1977 et publié par Gaudry et Wills en 1978, ce modèle qui généralisait le modèle en usage (Rea *et al*, 1977) fut immédiatement utilisé pour faire les prévisions de demande pour le réseau aérien canadien (Transport Canada, 1979). Malgré les coûts de calcul plus importants que ceux de la version linéaire, le modèle Box-Cox Logit Standard est maintenant d'usage courant dans nombre de pays, tant à l'échelle urbaine qu'interurbaine, et autant pour des problèmes de voyageurs que de marchandises.

Dans sa version généralisée, ce modèle rend aussi possible une incarnation pratique du Logit Universel de McFadden (1975) qui cherche à utiliser tous les prix (ou niveaux de service) de tous les modes dans toutes les fonctions d'utilité, selon :

$$V_i = \beta_{i0} + \sum_n \beta_m^i X_n^{i(\lambda_m)} + \sum_n \sum_{j \neq i} \beta_m^j X_n^{j(\lambda_m)} + \sum_s \beta_{is} X_s^{(\lambda_{is})} \quad (20)$$

Ce modèle Logit Box-Cox Généralisé (15) et (20) ramène en fait le Logit dans le giron des systèmes complets de demande. On l'applique aux flux marchandises à travers les Pyrénées (Gaudry *et al*, 2006). En modifiant profondément la structure des erreurs de (16) par l'ajout de variables « manquantes » en situation d'utilité séparable et donc responsables par leur absence³² des corrélations supposées entre ces erreurs, variables dont la présence rend aussi possible la complémentarité entre les modes, il pourrait devenir un concurrent sérieux du Logit hiérarchique. En effet, par structure, il viole l'axiome IIA et rétablit des élasticités croisées normales entre les modes. On sait par ailleurs que la non linéarité a des effets profonds sur le comportement du Logit en prévision, comme l'a bien démontré l'article de Mandel *et al* (1997). Sur la Figure 3, on constate que modifier X_I de 2 à 4 ou à 6 implique des gains très différents de la part de marché p_I selon la valeur estimée du paramètre de puissance Box-Cox.

31. Rétrospectivement, on peut penser que la nature multiplicative de la loi de Newton n'était pas intuitive dans la mesure où notre intuition perçoit plus facilement les corrélations entre variables linéaires qu'entre logarithmes des mêmes variables. C'est peut-être aussi pourquoi son application aux flux de transport prit si longtemps après 1684.

32. En économétrie, il est rarissime que la corrélation entre les erreurs soit un phénomène aléatoire pur : elle est normalement l'effet de variables pertinentes absentes. En ce sens, le Box-Cox Logit Généralisé rétablit leurs présences.

Coefficients aléatoires et « Mixed Logit ». Formulé et exploré dès 1977³³ par Johnson (1979a, 1979b), le Logit à coefficients aléatoires, aussi appelé « *Mixed*³⁴ *Logit* », est resté dans l'ombre jusqu'à sa redécouverte récente et sa reformulation au goût du jour (McFadden et Train, 2000).

Il pose des problèmes difficiles de formulation des hypothèses sur la distribution et l'estimation des valeurs des paramètres (Johnson, 1977, 1978) et sur la matrice d'information des variables (Cirillo, 2005)³⁵. S'il est déjà difficile d'avoir, pour un individu, une idée juste de la valeur relative pour lui-même des coefficients du tarif, du temps de marche, d'attente ou de transport, comment alors avoir des notions claires et infirmables sur leurs distributions dans la population ? Le problème est plus difficile encore pour les variables socioéconomiques comme le sexe ou le revenu : la distribution du coefficient de la variable indicatrice « homme » est-elle arc tangente ou semblable à celle de la testostérone ? Et cette distribution varie-t-elle par mode de transport ou est-elle nécessairement « générique » ?

Plus important peut-être du point de vue de son avenir semé d'embûches (Hensher et Greene, 2003), une chose au moins est claire : ce modèle revient en partie à prendre en compte la courbure associée à chaque variable. Dans ce sens, Orro *et al* (2005) ont montré que sa récente popularité était peut-être due justement au fait que les véritable relations ne sont pas linéaires et qu'il faut plutôt estimer leurs courbures au lieu de les supposer connues et d'introduire des coefficients aléatoires : ces auteurs nous ramènent eux-mêmes explicitement à la pertinence des transformations de Box et Cox discutées plus haut.

Intersection entre choix modal et affectation. Il est clair qu'en combinant les composants formulés dans les étapes précédentes aux parts de marché ou probabilités de choix entre les modes définies à cette étape-ci, on obtient un modèle de demande par mode et par paire origine-destination. Par exemple, la combinaison de (12) et d'une représentation générale d'un modèle de parts, Logit ou autre, permet d'écrire ainsi ce résultat :

$$T_{ij_m} = T_{ij} \bullet p_{ij_m}. \quad (21)$$

33. Source : discussion avec l'auteur en avril 1977 à la *Third International Conference on Behavioural Travel Modelling* de Tanunda en Australie.

34. Pas au sens de « *mixed up* » comme on pourrait le craindre si l'on a conscience des rendements marginaux décroissants, voire négatifs, en modélisation de la demande qui a toujours été un exercice de construction d'individus représentatifs identifiables par les segments de marché.

35. « *Its information matrix does not have a closed form and hence its efficiency bound is not defined.* »

qui, plus explicitement, se formule³⁶ :

$$T_{ijm} = f(A_i, A_j, U_{ij}) \bullet \{U_{ijm} / \sum_m U_{ijm}\}, m = 1, \dots, M \quad (22)$$

avec un indice de l'utilité globale des modes U_{ij} défini sans prendre le logarithme³⁷ de la somme comme en (17) :

$$U_{ij} = \sum_m U_{ijm}, \quad (23)$$

indice qu'on adaptera à nos fins immédiates, sans perte de généralité, à l'aide d'une spécification plus simple de ses termes que celles qu'on trouve en (19) ou (20), afin de mettre en lumière le groupe des X_{ijmk} dont nous voulons discuter la provenance :

$$U_{ijm} = \exp [\sum_k \beta_k X_{ijmk}], \quad (24)$$

En fait, nous avons toujours supposé plus haut dans l'exposé que l'ensemble $\{X_{ijmk}\}$ des K caractéristiques des M modes était bien disponible pour toutes les paires origine-destination et qu'elles avaient une valeur modale *unique* pour chaque paire. Mais nous n'avons pas expliqué comment ces valeurs de réseau (le prix et le temps, pour l'essentiel) étaient calculées. Nous voulons le faire brièvement dans le cadre d'une présentation simple de la quatrième étape, celle du choix d'itinéraire appliqué à chaque réseau modal et qui permettra de calculer les valeurs des X_{ijmk} .

On voit bien que, si les demandes sont définies de i à j mais que les réseaux sont composés de liens s partagés par les flux de plusieurs paires origine-destination, la tâche de cette étape d'affectation par réseau sera bien de générer des valeurs *uniques* des caractéristiques immédiatement utilisables dans les fonctions de demande et, simultanément, reflétant bien l'usage du réseau pour ces niveaux de demande compatibles avec elles au sens d'un équilibre Demande-Performance.

36. Le problème de savoir si l'on doit estimer les paramètres des deux modèles conjointement ou séquentiellement est autre chose. Laferrière (1988) a montré les gains importants d'une estimation conjointe. On peut conjecturer qu'ils seraient plus grands encore si l'on s'intéressait dans (22) aux variables qui sont présentes dans les deux parties du modèle. Si, en 1975, les modèles de choix modal utilisaient surtout des variables de réseau, elles incorporent de nos jours beaucoup de variables socio-économiques dont on peut penser que leur rôle premier est plutôt dans la partie génération-distribution du modèle. Pour stopper cette dérive « *all in* » des modèles de choix modal, le développement de l'estimation conjointe introduite par Laferrière sera nécessaire malgré les coûts de calcul importants qu'elle implique. On peut penser que certaines variables socioéconomiques comme le sexe ou le revenu auront simultanément un rôle dans les deux parties du modèle, ce que les tests de sous-estimation conjointe clarifieraient.

37. On peut ensuite tester la notion de valeur inclusive en établissant la forme optimale du membre gauche de (22).

On voit aussi aux Figures 1 et 2 que ces valeurs d'équilibre impliquent par définition un calcul compliqué parce que les niveaux des caractéristiques varient avec les niveaux des flux. Dans le cas contraire de niveaux de service ou de performance invariants aux niveaux des flux, la procédure d'équilibre en sera simplifiée. Encore faudra-t-il que, si plusieurs itinéraires sont utilisés pour une liaison donnée, leurs valeurs soient colligées et exprimables par des mesures uniques par paire origine-destination, conformément aux exigences des formulations des fonctions de demande.

L'affectation aux réseaux : des courbes de trafic « détourné » à l'équilibre. Dans un premier temps, on utilisa des « courbes de diversion » de type (13) pour expliquer le choix entre des infrastructures de classes différentes, comme les autoroutes et les voies urbaines, en reliant la part du trafic qui utilisait un genre d'infrastructure au ratio des temps de transport et parfois à la distance. Mais l'usage du temps de transport exigeait de mettre en œuvre une procédure de calcul qui les restituait pour les utiliser dans (13) : elle était à toutes fins utiles manuelle.

Une solution beaucoup moins approximative et bien plus satisfaisante apparut en 1957 avec le début du calcul des chemins les plus courts dans un réseau et l'implantation informatisée d'algorithmes adéquats (Moore, 1957 ; Dantzig, 1960), algorithmes et logiciels d'affectation utilisés pour la première fois à grande échelle dans CATS (1959-1962). Cette solution est toujours appliquée de nos jours quand il n'est pas important d'obtenir une affectation sur plus d'un chemin reliant deux localisations. Par ailleurs, comme le chemin le plus court est unique par définition, sa restitution par l'algorithme rencontre automatiquement le souhait de produire une mesure unique utilisable dans la fonction de demande.

Dans le cas où plusieurs chemins semblent utilisés et qu'il faille en tenir compte, deux types de cas se présentent selon qu'on cherche ou non à modéliser la congestion. En l'absence de congestion, il faut rendre compte de l'usage apparent³⁸ et simultané de plusieurs chemins de longueurs différentes. Un calcul du chemin le plus court est donc par hypothèse insuffisant.

38. Cet impression fondée sur les données peut n'être due par exemple qu'au niveau d'agrégation utilisé lors de la définition de la taille des zones entre lesquelles on a décidé de modéliser les flux.

La première parade³⁹ consiste à utiliser un modèle Logit (Dial, 1971) mais cette solution n'est pas complètement satisfaisante parce que, en plus de produire des flux qui obéiront à l'axiome IIA (l'ajout d'un nouvel itinéraire détournant la même proportion des flux concurrents, même s'il partage l'essentiel du parcours avec un autre), le modèle exige qu'on lui fournisse la liste des itinéraires à considérer. Ensuite, pour produire une mesure unique de leur utilité, on peut utiliser dans la fonction de demande le *logsum* approprié par paire origine-destination.

Mais, par structure, la formule Logit affectera un certain flux à tous les itinéraires que l'on décidera d'avancer. Ce problème peut être difficile à régler : aujourd'hui par exemple, les voyageurs sur le réseau aérien intérieur allemand n'utilisent pas plus de 3 itinéraires par paire origine-destination (les zones étant définies au niveau urbain) mais les mêmes voyageurs utilisent régulièrement 16 itinéraires vers les États-Unis. On peut bien sûr faire la liste des itinéraires utilisés dans chaque cas et construire deux modèles distincts, mais on préférerait que le nombre d'itinéraires utilisé fasse partie de la réponse du problème d'affectation. On désire aussi prendre en compte le niveau variable de la congestion.

C'est ce qui se passe indirectement avec un algorithme d'équilibre⁴⁰ qui affecte les flux au réseau : (i) soit de manière à ce que personne ne puisse améliorer son itinéraire et que cet itinéraire soit pour chacun le plus court possible, ce que l'on appelle un équilibre *des usagers* (*user optimal*), *de comportement* ou égoïste⁴¹ ; (ii) soit de manière à ce que le coût moyen du transport par usager soit le plus faible possible, ce que l'on appelle un équilibre *du système* (*system optimal*), ou équilibre *de contrôle* centralisé des véhicules.

L'algorithme qui calcule l'équilibre *de comportement* (Nguyen, 1974) décide implicitement des chemins utilisés même s'il ne les énumère pas parmi les propriétés de la solution à l'optimum⁴² : au moins le

39. Nous ne voulons pas dire qu'un modèle Logit de choix entre itinéraires n'est pas utilisable en présence de congestion. Nous cherchons simplement à en discuter avec des niveaux de caractéristiques invariants par rapport aux flux.

40. Dans le langage courant, l'expression « calcul d'équilibre » désigne un équilibre des usagers parce que c'est celui qui est utilisé pour sa conformité avec le comportement des automobilistes, mais cette pratique linguistique commune n'interdit pas de parler aussi, au sens technique, d'équilibre dans le second cas.

41. Le mot n'implique pas qu'un usager puisse devenir altruiste sans contrôle centralisé car il n'a aucune idée du coût marginal de son déplacement et de la façon dont il varierait dans un jeu avec les autres usagers, altruistes ou pas, en l'absence de contrôle centralisé.

42. Certains chercheurs qui veulent avoir une idée des chemins utilisés les identifient à partir des dernières itérations de l'algorithme de Frank et Wolfe (1956) généralement utilisé, mais il s'agit là d'une manière *ad hoc* de compléter le contenu de la solution optimale par une information imparfaite qu'elle ne fournit pas d'elle-même.

flux sur chaque lien est-il unique à l'équilibre, ce qui signifie que l'analyste n'a pas à décider des liens utilisés, car ils sont sélectionnés par l'algorithme. Et, comme à l'équilibre tous les chemins utilisés pour une liaison, quels qu'ils soient, sont de longueur égale, le temps de transport restitué est nécessairement unique de i à j en présence de congestion.

On trouvera à la Figure 4 la représentation graphique des deux solutions définies par Wardrop (1952) et calculables pour l'ensemble des paires origine-destination en présence de congestion : l'équilibre de comportement et l'équilibre du système.

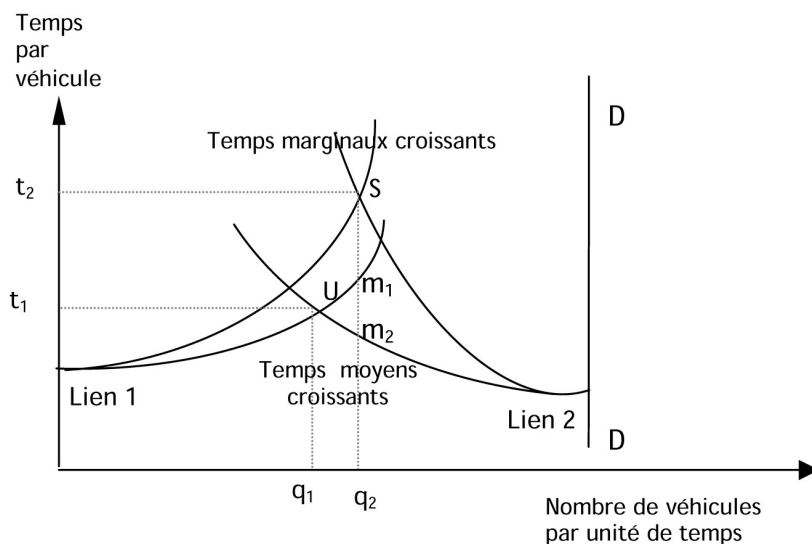
L'usage d'une demande *fixe* DD (qui sert aussi d'ordonnée à la courbe de débit-temps du lien 2) n'implique aucune perte de généralité ; les courbes de débit-temps approximatives permettent malgré tout de distinguer les deux équilibres qui nous intéressent. On distingue facilement le premier au point U : il s'agit bien d'un équilibre au coût moyen (égoïste ou non coopératif) ; dans le second cas, l'équilibre au point S résulte bien d'une affectation au coût marginal des véhicules.

Une explicitation peut aider à comprendre. Dans cette figure, on suppose que la demande DD est invariante au temps de transport et on l'affecte sur deux liens 1 et 2 aux caractéristiques différentes. Comme les deux liens sont représentés « face à face », même s'ils sont parallèles dans la réalité, on peut voir qu'au point U , la totalité du trafic est affectée (q_1 véhicules au lien 1 et le reste au lien 2) et que le temps de transport par véhicule est égal à t_1 sur chaque lien. Cette solution est un équilibre de comportement, stable, parce que le temps de transport est identique sur les 2 liens concurrents et que la totalité de la demande est affectée (voir figure 4).

Par contre le point S , au niveau duquel le temps marginal est t_2 sur chaque lien, est un équilibre du système : q_2 véhicules utilisent le lien 1 et le reste utilise le lien 2 . Toutefois, cet équilibre n'est pas soutenable sans contrôle centralisé parce que le temps des usagers est beaucoup plus élevé sur le lien 1 (au point m_1) qu'il ne l'est sur le lien 2 au point m_2 . Cela signifie que, par rapport au cas précédent, le report de trafic $q_1 q_2$ effectué du lien 1 au lien 2 pour affecter toute la demande minimise bien le coût total du transport mais n'est pas une solution stable si les usagers cherchent toujours à minimiser leur temps propre.

Le fait que l'usager se comporte de façon à minimiser son temps de transport en choisissant son itinéraire en fonction du coût moyen de chaque lien (puisque c'est ce coût qu'il encourt et peut évaluer) plutôt que de façon à minimiser le coût total pour tout le système, conduit

Figure 4

Solutions graphiques : équilibres au coût marginal et au coût moyen par lien

parfois à des bizarreries. La plus connue est « le paradoxe de Braess » qui montre que dans certains cas l'ajout d'un lien nouveau sur un réseau existant augmente le coût total de transport (Braess, 1968, ou en anglais, Braess *et al*, 2005). Cette situation est semblable à celle que l'on retrouve en microéconomie quand on y montre que le meilleur usage des ressources par les firmes exige que le rendement marginal de chaque usage soit le même, pas que le rendement moyen de chaque usage le soit, faute de quoi une réaffectation sans coût des ressources peut améliorer le rendement total.

Naturellement, les économistes se saisissent du graphique et, fascinés par la différence apparente entre le coût moyen et le coût marginal sur chaque lien de la solution de comportement, recommandent *ad nauseam* l'usage de taxes optimales qui rétabliraient *après leur implantation*⁴³ l'usage efficient des ressources. Encore faudrait-il connaître ces écarts délictueux au bon moment, au bon niveau, pour tous les liens d'un réseau réel. Les implantations de péages urbains, comme ceux que l'on trouve en Scandinavie ou à Londres, n'obtiennent pas l'approbation politique sur la base d'un marketing de la tarification au

43. On peut aussi définir le niveau de ces taxes par lien de manière graphique.

coût marginal mais sur d'autres bases plus réalistes⁴⁴ et plus susceptibles de susciter l'approbation des électeurs⁴⁵.

Notre présentation simplifiée des procédures d'affectation qui calculent un équilibre conforme à la représentation du système de la Figure 1 ne précise pas de quelle tranche du temps (*e.g.* pointe du matin, flux annuel moyen) il s'agit et traite les temps de transport comme certains. Il existe bien sûr une littérature raffinée sur l'affectation stochastique et sur l'affectation dynamique⁴⁶ et sur la prise en compte d'autres éléments, comme le coût, dans les calculs d'équilibre. Mais, comme pour les étapes précédentes, ces raffinements et complexités⁴⁷ sont trop nombreux pour nous retenir. Considérons plutôt une autre manière simple de formuler le problème : traiter ensemble les étapes du choix modal et de l'affectation.

3.3.2. *Considérés comme un seul problème*

Un réseau multimodal. Au lieu d'expliquer d'abord un choix entre modes principaux et d'affecter ensuite les flux sur chaque réseau modal « pur », on peut supposer qu'il n'y a qu'un réseau « abstrait du point de vue du mode », dont les liens possèdent certaines caractéristiques (comme le prix et le temps de transport), et y affecter les flux de la matrice O-D tous modes pertinents confondus en suivant une certaine règle d'affectation : Logit, chemin le plus court, équilibre de comportement, *etc.* Le calcul des personnes-km ou des tonnes-km par mode est alors fait *ex post* à partir des flux affectés, en faisant les sommes appropriées sur les liens modaux « purs ». Cette façon de faire, qu'on peut appeler *micro-affectation multimodale*, est surtout associée à des travaux français sur les déplacements voyageurs mais rencontre maintenant ailleurs qu'en France un succès certain, notamment dans le domaine des marchandises.

Dans la première formulation pour les voyageurs de ce qu'on appelle aussi le modèle *prix-temps*, Abraham *et al* (1969) supposent une distribution parétienne des valeurs du temps des individus com-

44. L'Annexe 2 résume une étude sur le sens économique des courbes de débit-temps (Gaudry, 1976) et propose l'hypothèse que les économistes ne sont pas compris des électeurs parce que la courbe débit-temps sur laquelle repose l'argumentaire des premiers n'est pas vraiment une courbe d'offre pour un bien homogène de qualité constante.

45. Comme la fluidité du trafic, un concept mal défini, ou le transfert obligatoire des péages à un fonds dédié à l'amélioration du réseau routier local.

46. Une des questions importantes que nous avons évitées est celle de la reproductibilité des affectations. Certaines procédures produisent un équilibre qui dépend de la séquence de chargement du réseau, solution qui n'est donc pas reproductible si on charge le réseau dans un autre ordre. Ces problèmes se posent naturellement aussi pour l'affectation dynamique par tranche horaire.

47. L'usage de plus d'un critère d'affectation (*e.g.* temps et argent) introduit des complications sérieuses, voire dirimantes, pour l'unicité des solutions d'équilibre (Dafermos, 1983).

mune à toutes les liaisons du réseau. Ils calculent ensuite analytiquement la part modale de l'avion par rapport au train, ce qui ramène leur formulation finale dans le giron des modèles de parts puisqu'aucune microaffectation n'y est effectuée. Dans l'application postérieure de la même idée, on suppose typiquement une distribution des valeurs du temps log-normale et on affecte chaque individu de l'échantillon au réseau multimodal « abstrait » par le chemin le plus « court » après avoir calculé pour chacun un coût généralisé construit à partir d'une valeur particulière de son temps tirée de la distribution. On trouvera une élaboration de l'idée dans les deux articles de Marche (1980). Le modèle MATISSE (Morellet et Marchal, 1995) semble avoir généralisé cette idée en l'appliquant à des centaines de segments fins du marché interrégional pour lequel elle avait été formulée.

Dans le domaine des marchandises, où le prix du transport est si important relativement au temps de transport, le dernier-né des modèles de microaffectation multimodale est NODUS (Jourquin, 1995) dont la dernière version 5.0 autorise même des calculs d'équilibre en présence de congestion (Beuthe *et al*, 1999 ; Jourquin et Beuthe, 2006).

Plusieurs analystes utilisent en fait ce genre d'affectation multimodale *désagrégée* qui détermine le choix du mode et de l'itinéraire en la combinant avec un modèle de Génération-Distribution *agrégé* de forme multiplicative (Arduin, 1989 ; Chopinet, 1997), recherchant en quelque sorte le meilleur des deux mondes⁴⁸.

Sens économique. De par leur structure même, les modèles qui combinent Génération-Distribution dans une étape et Choix modal-Affectation dans une autre, et en multiplient les résultats comme en (21), partagent en deux les effets des modifications aux réseaux de transport envisagées dans un projet : l'effet d'*induction* (ou de génération) dans la première partie et l'effet de *report* (ou de détournement ; *diversion* en anglais) modal dans la seconde.

Une telle décomposition des effets de modifications de réseaux est commune dans les transports même si l'architecture de combinaison des étapes du processus classique varie de modèle en modèle. Nous reviendrons plus bas sur cette décomposition des effets de prix ou de temps *particulière aux transports* (et aux communications) et nous

48. Si l'on utilise pour les valeurs du temps une distribution asymétrique, comme la loi log-normale, la courbe implicite de la part de marché d'un mode sera asymétrique, ce qui fait penser à celle d'un Logit Box-Cox de la Figure 3, comme le reconnaît par exemple Jean-Pierre Arduin depuis longtemps dans ses communications depuis 1989 avec l'auteur, afin d'expliquer la qualité et le succès du « modèle SNCF-SOFRERAIL » tant en France qu'à l'étranger.

tenterons de la comparer aux décompositions classiques de Hicks ou de Slutsky, bien connues de tous en microéconomie.

Nous l'exprimerons aussi à l'aide d'une mesure calculable pour tout modèle de transport, indépendamment de son architecture particulière, celle des élasticités d'induction et de report modal qui permettent de comparer rapidement entre eux les résultats de plusieurs modèles de transport.

3.4. *Inversion de la séquence classique et estimation des matrices O-D*

Matrice O-D manquantes. Avant d'arriver au sens microéconomique de la *décomposition transport* et à l'intérêt de l'exprimer sous forme d'élasticités, il faut dire quelques mots sur des études importantes qui organisent les étapes classiques dans l'ordre inverse de celui que nous avons considéré. Elles visent à reconstituer, à partir de comptages sur les liens, la matrice *O-D* pertinente d'un projet.

Intérêt de la question. Le problème à résoudre est donc celui d'estimer, parmi le très grand nombre de matrices possibles, la matrice O-D qui est la plus compatible avec les comptages observés. Pour mieux comprendre l'importance de la réponse pour nos fins, considérons à la Figure 6 un réseau constitué de 6 noeuds, dont 4 sont aussi des zones d'origine et de destination, et de 6 liens.

Le premier dessin 6.1 indique sur chaque lien les flux et les temps de transport observés en période de référence. Le problème est que ces flux auraient pu être générés par plusieurs matrices O-D, par exemple les matrices A et B de la Figure 5. Pour le comprendre, supposons d'abord que les usagers choisissent le chemin le plus court. Affectons ensuite un par un chacun des flux O-D de la matrice A au réseau en inscrivant tout passage sur un lien utilisé : en faisant ensuite la somme des inscriptions par lien, on retrouve bien les comptages totaux indiqués en 6.1. Mais cela est vrai aussi si l'on refait l'exercice avec la matrice B. Les matrices A et B sont donc compatibles avec les comptages observés en 6.1.

Remarquons maintenant que, si le temps de transport sur le lien $5 \rightarrow 6$ est modifié de 2 à 3 minutes, les flux affectés au réseau selon la même règle d'affectation demeureront inchangés si la matrice A est vraie, ce qu'on montre dans le second dessin 6.2 – identique au premier, mais seront profondément modifiés si la matrice B est vraie, comme l'indique 6.3, le troisième dessin de la Figure 6.

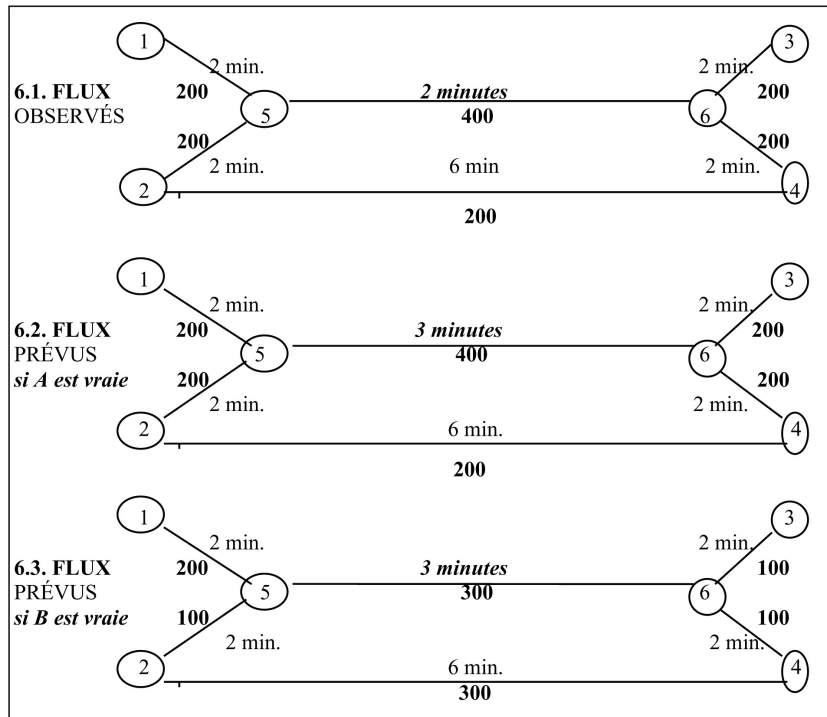
Figure 5

Matrices O-D compatibles avec les flux et les temps observés à la Figure 6.1

		Destination				Destination			
		Zone	3	4			Zone	3	4
Origine	1	0	200		Origine	1	100	100	
	2	200	200			2	100	300	
Matrice A				Matrice B					

Figure 6

Flux observés et prédits suite à des modifications du temps de transport



Il est donc très important pour un calcul de rentabilité des liens anciens ou nouveaux d'un réseau de connaître aussi bien que possible la vraie matrice O-D définie dans la situation de référence du projet. Ce problème d'estimation de la vraie matrice à partir des comptages sur les liens a d'abord été posé indépendamment, et à la même période, dans plusieurs pays dont le Canada, le Danemark et les États-Unis.

Modèles structurels et régressions multivariées. La formulation du problème, adoptée entre 1969 et 1972, tant avec des données urbaines (à Montréal : Arbour *et al*, 1969, 1971a, 1971b ; à Silkeborg : Overgaard, 1972) qu'avec des données interurbaines (en Colombie britannique : Wills, 1971 ; aux États-Unis : Low, 1972), peut être résumée⁴⁹ en deux équations :

$$L_s = \sum_i \sum_j pr(s)_{ij} \bullet T_{ij} \quad (25)$$

et

$$T_{ij}^{(\lambda_y)} = \sum_k \beta_k X_k^{(\lambda_k)} , \quad (26)$$

où $pr(s)_{ij}$ désigne la proportion des déplacements de i à j qui utilise le lien s et L_s désigne le nombre total des véhicules passant par ce lien, quelles que soient leur origine ou leur destination.

Si, après inversion de la transformation Box-Cox de la variable dépendante de (26), on substitue le résultat obtenu dans la première expression (25), comme on le fera plus tard (Wills, 1978), on peut alors écrire une forme assez générale inspirée du modèle multiplicatif (21) :

$$L_s = \sum_i \sum_j pr(s)_{ij} \bullet \left\langle \lambda_y \left[\sum_k \beta_k X_{ij,k}^{(\lambda_k)} \right] + 1 \right\rangle^{1/\lambda_y} + \varepsilon_s , \quad (27)$$

où l'on constate que l'ensemble $\{\lambda ; \beta\}$ des paramètres du modèle de demande sont vraiment estimés à partir des comptages et qu'ils produisent (en sorties intermédiaires) les valeurs de la matrice O-D « calculée » par la grande parenthèse du membre droit. Dans cette formulation, le critère de minimisation des erreurs, ou de maximisation de la vraisemblance, est défini en fonction des différences entre le flux total affecté sur un lien et le comptage qui y est observé. Dans certains cas, on généralise (27) en introduisant au milieu du membre droit un modèle de répartition modale⁵⁰ dont les paramètres sont aussi estimés dans la foulée par la même procédure. Dans ce cas, on a complètement inversé la séquence classique.

La limpidité de la formulation (27) met bien en lumière l'objectif de cette coulée de travaux : simplifier la modélisation du trafic (Bendtsen, 1974). En pratique, les premiers travaux d'estimation des matrices, faits entre 1969 et 1973, utilisaient tous une forme gravitaire (voir aussi

49. Le lecteur intéressé à la naissance de l'ensemble des problématiques diverses, bien vivantes pour la plupart avant 1978, peut consulter Gaudry et Lamarre (1979).

50. Il est loisible au chercheur de définir alors une mesure multimodale des $pr(s)_{ij}$ et des comptages.

Jentsen et Nielsen, 1973) en (26), c'est-à-dire le cas multiplicatif particulier obtenu en fixant $\lambda_y = \lambda_x = \mathbf{0}$ (pour tout X_k) dans (27).

Autres méthodes. À partir de 1977, certains chercheurs ont préféré à la méthode structurelle (27) des méthodes d'estimation des matrices « les plus probables » basées sur des critères d'information (van Zuylen, 1978), d'entropie (Willumsen, 1978) ou d'autres⁵¹ hypothèses statistiques (Spiess, 1987). Ils ont aussi élaboré, y compris en France (Debaille, 1977, 1979), des méthodes qui combinent la mise à jour d'anciennes matrices⁵² et de comptages récents.

Ils ont inventé et diffusé d'autres méthodes publiquement documentées comme celles qui combinent comptages et enquêtes⁵³ sur rue (DOT, UK, 1982), une approche qui a de nombreux adeptes (e.g. Kurihara et Sullivan, 1987). Elle se prête aussi à l'usage de méthodes statistiques sophistiquées comme celles qui calculent (DETR, UK, 1998)⁵⁴ les écarts-types des flux estimés en supposant une distribution hypergéométrique des erreurs sur les liens, mais sans toutefois encore prendre en compte leur corrélation spatiale ou les autres dépendances découlant de la topologie du réseau.

Du point de vue analytique, le problème le plus difficile est peut-être celui de l'endogénéité des $pr(s)_{ij}$ de l'expression (25) en présence de congestion, problème qui a été abordé assez tôt (Nguyen, 1977). Il existe maintenant de nombreux travaux qui résument les diverses méthodes⁵⁵ utilisées. On peut nommer celles de Cascetta et Nguyen (1988) et d'Abrahamsson (1998). Il s'agit d'un champ de recherche encore ouvert et qui donne naissance à des travaux innovants (e.g. Bierlaire, 2002).

Cette problématique prend aujourd'hui une importance croissante avec la privatisation de nombreux opérateurs et la difficulté de coordonner la cueillette d'informations, en particulier quand plusieurs pays sont concernés par un projet. En ce sens, l'estimation des matrices

51. Les grands logiciels offrent souvent de telles méthodes de reconstitution des matrices O-D qu'il faut établir avant même de commencer le travail d'estimation de la demande et d'affectation au réseau – on suppose par des procédures autres que celles qui ont servi à l'estimation de la matrice !

52. Ou d'informations sur la structure de la matrice, provenant par exemple des plaques de véhicules occupant des parkings (Bierlaire et Toint, 1995).

53. *Road side interviews (RSI)*, où l'on interroge les voyageurs sur leurs origines et destinations.

54. James Clavering (Clavering et Kirby, 1994) semble être un des principaux responsables de ces développements.

55. L'étude préparée à la demande de la DG VII pour encadrer les travaux du consortium MYSTIC (Gaudry, 1998) en distingue quatre en plus de leurs combinaisons et des procédures de « dérivation » des matrices N à l'aide de techniques input-output appliquées dans l'esprit de l'Annexe 1. L'existence d'éventuelles dérivations par les méthodes d'équilibre général calculable y est anticipée mais aucune n'est répertoriée.

origine-destination à partir des comptages sur les liens fait maintenant partie de la modélisation du trafic.

4. SENS ÉCONOMIQUE DE CETTE MODÉLISATION PROPRE AUX TRANSPORTS ET AUX COMMUNICATIONS

À l'évidence, la modélisation du trafic utilise un ensemble de techniques propres qui permettent de calculer l'impact de modifications des infrastructures de transport sur le niveau de la demande et modélisent souvent très explicitement aussi des effets de substitution et de complémentarité entre les modes ou les itinéraires qui composent cette demande globale. L'économiste se demande alors si ces techniques reviennent vraiment à celles qui lui sont déjà familières.

Pour satisfaire cette curiosité, nous procéderons en trois étapes. Dans la première, nous rappellerons la nature de la décomposition normalement structurante des modèles de trafic, décomposition que nous appellerons « décomposition transports ». Dans la seconde, nous comparerons sur un seul graphe cette dernière avec les décompositions communes en microéconomie, celles de Hicks et de Slutsky.

Dans la troisième, nous exprimerons la décomposition transports de la demande par mode sous forme d'une somme d'élasticités, forme applicable à tout modèle de trafic, voire à tout modèle de demande. Nous soulignerons que cela est vrai, même si la structure d'un modèle de demande ne fait aucun usage explicite des deux composants-clés des modèles de trafic documentés plus haut : l'*Induction* et le *Report*.

4.1. La décomposition transports : une représentation graphique

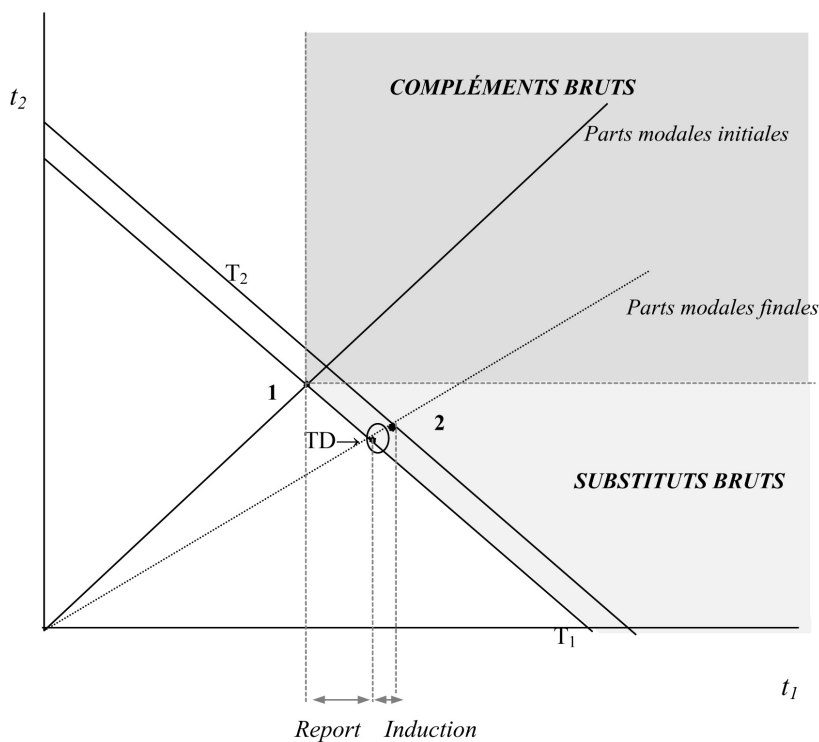
i) Induction et report modaux : un cas à deux modes. La décomposition d'une variation de la quantité demandée pratiquée par les modèles de trafic pour représenter l'effet d'une baisse de prix d'un mode est présentée à la Figure 7. Pour comprendre ce graphique, il faut d'abord remarquer que les deux lignes droites tracées à partir de l'origine, et passant respectivement par les points 1 et 2, représentent le partage modal initial et final ; de plus l'intersection de l'une de ces raies et de toute droite T_1 ou T_2 conserve le partage modal entre les modes 1 et 2. Par contre, tout mouvement le long d'une droite de pente négative placée à 45° implique un nouveau partage entre des clientèles modales t_1 et t_2 effectuant un nombre total de déplacements constant.

On appelle *report modal* (effet de *transfert*) le passage du point 1 au point TD où les parts modales ont changé sans que le nombre total de

déplacements n'ait été modifié, et *induction modale* (ou effet de *génération*) le passage du point TD au point 2 où le nombre total de déplacements a changé sans que n'aient changé les parts modales.

On trouve aussi à la Figure 7 une identification des régions qui, suite à une baisse de prix du mode 1, correspondraient à de la complémentarité ou de la substitution en microéconomie classique. La question se pose alors de la relation entre la décomposition $\{[1 \rightarrow \text{TD}]; [\text{TD} \rightarrow 2]\}$ définie en transport et les décompositions connues de Slutsky (1915) ou de Hicks (1939).

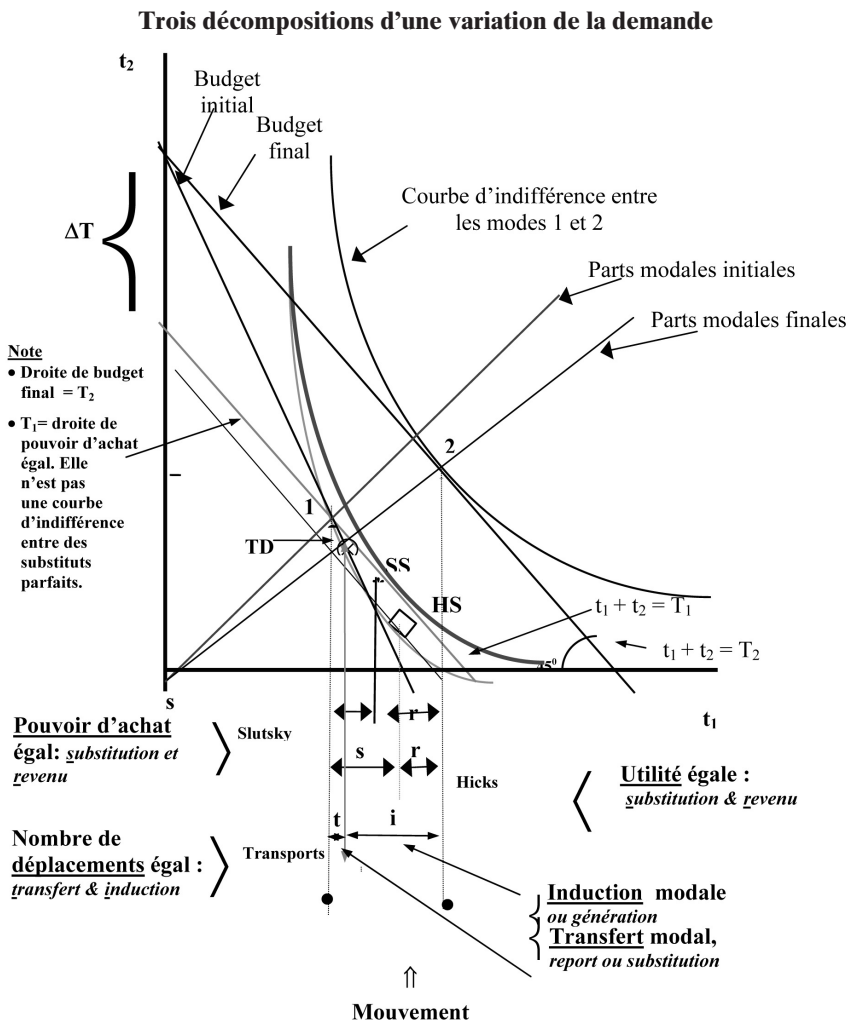
Figure 7
La décomposition transports



4.2. La décomposition transports et celles de Hicks et de Slutsky

ii) *Hypothèse simplificatrice : une droite à 45 degrés.* À la Figure 8, nous supposons que les combinaisons de biens $\underline{1}$ et $\underline{2}$ maximisent l'utilité et que, suite à la baisse du prix du mode $\underline{1}$, la droite de budget final passant par le point $\underline{2}$ est inclinée à 45° .

Figure 8



Source : Gaudry, 1998.

Cette hypothèse particulière nous permet de superposer la parallèle à cette droite de budget final, parallèle que nous faisons passer par le point 1 et une droite à déplacements totaux constants comme T_1 ou T_2 de la Figure 7. Cela revient à supposer l'identité visuelle de la droite à budget égal de Slutsky et de la droite à déplacements totaux constants (à **output total égal**) des modèles de trafic. Cette superposition de deux droites différentes permet de comparer sur le même graphique les trois décompositions qui nous intéressent.

Pour définir l'effet de substitution de Slutsky, il faut trouver le point d'utilité maximale SS sur la droite de budget hypothétique qui incor-

pore les nouveaux prix relatifs prévalant au point **2** mais qui nous permet d'acheter l'ancienne combinaison de biens **1**. Comparé au point **1**, le point **SS** n'est pas à output égal mais à **budget égal** (*pouvoir d'achat*). Les mouvements $\{[1 \rightarrow SS]; [SS \rightarrow 2]\}$ sont les effets de substitution et de revenu de Slutsky qui définissent la décomposition de Slutsky.

L'effet de substitution de Hicks requiert de trouver le point de tangence **HS** entre la courbe d'indifférence initiale qui détermine le choix du point **1** et la droite de budget minimum incorporant les nouveaux prix relatifs prévalant au point **2**. Comparé au point **1**, le point **HS** n'est pas à budget ou à output égal mais à *utilité égale*. Les mouvements $\{[1 \rightarrow HS]; [HS \rightarrow 2]\}$ sont les effets de substitution et de revenu de Hicks qui définissent la décomposition de Hicks.

La différence la plus importante entre les trois groupes d'effets sont que : (i) les effets de substitution de Slutsky et de Hicks sont nécessairement positifs, ce qui n'est pas le cas de l'effet de report, même si on s'attend à ce qu'il le soit normalement ; (ii) l'effet d'induction est nécessairement positif, ce qui n'est pas le cas des effets de revenu auxquels il ressemble, même si on s'attend à ce qu'ils le soient normalement.

4.3. *Expression de la décomposition transports sous forme d'élasticités*

Niveaux absolus et relatifs des effets. Pour comparer les propriétés économiques et de pertinence entre modèles de trafic de facture (*i.e.* structure, données, méthode de quantification des paramètres, ...) très différente, on peut bien sûr s'intéresser dans chaque cas au calcul des *niveaux des effets* de report et d'induction ; mais il est préférable d'exprimer leurs comportements et résultats économiques sous forme d'*élasticités* par mode, ce qui permet d'ignorer les ordres de grandeur très variés, tant des variables expliquées que des variables explicatives, en cause dans tout groupe de modèles choisis un peu au hasard.

L'une des lames du rasoir empirique. Les élasticités sont un résumé simple et intuitif des propriétés qu'il est utile d'appeler « *de pertinence* » d'un modèle, pour les distinguer de ses propriétés *numériques* et *statistiques* : elles résument lisiblement *la force* des relations entre les variables, relations dont le signe, la taille et la signification statistique constituent les sorties principales, souvent intuitivement illisibles, d'un modèle. Le caractère raisonnable des élasticités est un critère de sélection entre modèles plus clair et limpide que celui du niveau de signification statistique des variables, puisque des résultats peuvent être significatifs au sens statistique mais irrecevables parce que dérai-

sonnables (et inversement). En conséquence, beaucoup d'auteurs se gardent bien de calculer et de montrer les élasticités de leur modèle par crainte d'un rejet immédiat de leurs résultats, si « statistiquement significatifs » soient-ils. Soumettre un modèle à l'épreuve des élasticités, combinée au filtre des tests de forme fonctionnelle dont nous avons déjà parlé (voir note 30), constitue aujourd'hui à notre avis une sorte de rasoir d'Ockham empirique à double lame.

Une somme d'élasticités calculables pour tout modèle. Si l'on admet l'importance de cette compréhension intuitive des résultats d'un modèle, compréhension dont les formes d'expression très diverses seront introduites plus bas et rappelées à l'Annexe 3, l'élasticité *usuelle* de la demande par mode de transport s'avère être utilement exprimable et décomposable de manière à reprendre les termes de la décomposition transports. On exploite pour cela **deux tautologies**. En effet, comme il est vrai par définition que le nombre total de déplacements T est la somme des déplacements par mode T_m , et que chaque T_m est le produit du nombre total T par p_m , la part modale du mode m , *i.e.* :

$$T = T_1 + \dots + T_m + \dots + T_M \quad (28)$$

et

$$T_m = T \cdot p_m \equiv T \cdot \frac{T_m}{T}, \quad (29)$$

on peut alors profiter du fait que (29) est un produit et qu'en conséquence $\eta(T_m, X_k)$, l'élasticité de la demande par mode T_m par rapport à une variable X_k , est décomposable en un impact sur la part modale p_m et un impact sur la demande totale T , indépendamment du mode :

$$[\eta \text{ du Mode } m] = [\eta \text{ du Total}] + [\eta \text{ de la Part } m],$$

ou (30)

$$\eta(T_m, X_k) = \eta(T, X_k) + \eta(p_m, X_k).$$

Cette décomposition de l'élasticité modale est non seulement conforme à la structure de nombreux modèles, comme nous l'avons vu, mais donne naturellement naissance aux élasticités d'induction et de report. En fait, elle serait applicable à toute variation de la demande à laquelle on souhaiterait appliquer la décomposition transports plutôt qu'une autre décomposition **conceptuelle**, donc par définition construite et *inobservable*, du passage du point 1 au point 2 *observés*.

Cas particulier de modèles couplés. Si le modèle (29) représentatif dans sa structure même de la modélisation du trafic est en fait un modèle « couplé » au sens où l'utilité globale des modes tirée du dénominateur du modèle de choix modal influence le niveau de la demande totale tous modes comme en (22), on peut réécrire (30) plus explicitement :

$$\eta(T_m, X_k) = \eta(T, X_k) + \eta(T, U) \bullet \eta(U, X_k) + \eta(p_m, X_k) \quad (31)$$

ce qui rend explicite, dans l'élasticité d'induction elle-même, l'effet de X_k qui passe par la variable de couplage U .

Expressions de l'élasticité. Le membre droit de la formule (30) comprend une élasticité de niveaux et une élasticité de parts (ou de probabilités). D'où viennent ces notions et sont-elles les plus adéquates ?

À l'Annexe 3, nous arguons de l'importance croissante de la notion d'élasticité à mesure qu'on s'éloigne de la linéarité des modèles. Nous rappelons l'origine de la notion simple d'élasticité et son développement progressif dans le cadre des modèles de niveaux à mesure qu'on reconnaissait le caractère aléatoire de l'élasticité. Nous présentons le développement plus récent de la notion d'élasticité dans les modèles probabilistes et de parts et, au vu du caractère standardisé des variables dépendantes de ces modèles, nous suggérons de remplacer l'usage courant des « élasticités-probabilités » ou des « élasticités-parts » par la mesure plus claire des « élasticités-points ».

5. CALCUL ÉCONOMIQUE ET PRÉVISION

La situation de projet. Les modifications du réseau existant définies dans les projets sont envisagées *à la marge* : on ne s'attaque pas au problème de conception d'un réseau optimal mais seulement à celui de l'ajout ou de la modification de quelques liens nouveaux ; ces changements sont décrits dans un nombre limité de configurations envisagées, les *scénarii de réseau* qui font ensuite l'objet d'une évaluation dans le cadre d'équilibre partiel conditionnel déjà défini plus haut.

Il existe des procédures plus ambitieuses qui cherchent à élaborer une offre optimale. Elles calculent un équilibre conjoint des demandes définies par paire origine-destination et de l'offre de services définis par lien. La plus connue de ces procédures, implantée dans le logiciel FNEM (Friesz *et al*, 1986), vise à établir le réseau de service pour les marchandises.

On imagine facilement la complexité de telles formulations à deux niveaux visant à réconcilier *design optimal* de la structure logistique de l'offre des chargeurs et *demande* de fret des expéditeurs. Même en l'absence de congestion, ce qui revient à ignorer à la Figure 1 le niveau intermédiaire de la Performance, la détermination du vecteur d'offre par une optimisation posera de redoutables problèmes d'existence et d'unicité des solutions d'équilibre proposées, sans parler de la pertinence d'une définition agrégée de l'offre logistique. En pratique et à toutes fins utiles, on définit donc à la main le réseau en situation de projet.

Modèles de trafic et calcul économique. On obtient des prévisions de demande en croisant les configurations de réseaux établies par le maître d'œuvre avec des *scénarii économiques* établissant les niveaux et la répartition des activités économiques prévus en situation de projet. Ces prévisions, dont chacune est conditionnelle à un croisement de deux *scénarii* particuliers, servent au calcul de rentabilité de l'investissement requis. Du point de vue économique, trois formes de ce calcul nous intéressent : le calcul *des surplus*, le calcul de *rentabilité financière* et le calcul de *valeur ajoutée* : on utilise presque exclusivement les deux première formes, la troisième, cas particulier de la seconde, est utile pour déterminer si un investissement qui n'est pas financièrement rentable peut malgré tout dégager une valeur ajoutée positive.

Exclusions du calcul. Pour préciser la nature de *la tendance centrale* lors de l'emploi des modèles de trafic à des fins de calcul économique, nous développerons ici la question du point de vue de l'*avantage transport pur* qui est l'élément central tiré des études de trafic. Nous souhaitons donc exclure de l'exposé le rôle des nuisances et n'inclure que partiellement le rôle des divers composants du vecteur des services offerts par le réseau [sûreté ; sécurité ; temps] parce que seul le temps est à ce jour complètement incorporé aux modèles de trafic. Dans quel esprit sont faites ces exclusions et inclusions ?

Concernant les nuisances (bruit, pollution, gaz à effet de serre), toujours légitimement incorporées au calcul des avantages et des coûts et si bien étudiées surtout en France (*e.g.* Boiteux et Baumstark, 2001), nous nous contentons de dire qu'elles sont dérivées des flux calculés par les modèles de trafic et qu'il est légitime de les appeler « externalités » et de les prendre en compte dans le calcul.

Externalités et *membership* du club. Par contre, nous ne considérons pas ici la congestion et les accidents comme des externalités mais comme des composants du vecteur qui définissent le bien « déplacement » : il s'agit de niveaux de service inclus parmi les variables qui

caractérisent l'équilibre de *Performance* du système décrit à la Figure 1. En conséquence, nous faisons comme si le risque d'accident pouvait être considéré de la même manière que le niveau de congestion : s'il est compensé par les assurances, une situation qui varie⁵⁶ entre pays, il devient interne aux membres du club des usagers des modes de transport et ne concerne plus directement⁵⁷ autrui.

Les membres du club peuvent naturellement souhaiter réduire le risque d'accident ou le niveau de congestion et sont peut-être prêts à payer un prix pour le faire et obtenir un bien de qualité supérieure, mais il s'agit en principe de leurs affaires en propre. On comprendra qu'ils se demandent généralement si le niveau de service reçu [sécurité ; temps] est une bonne affaire ou pas par rapport à ce qu'il leur coûte, mais c'est tout. Par exemple, aucun d'entre eux ne contestera la présence d'autrui sur le réseau si elle obéit aux règlements du club : les mauvais chauffeurs en situation de conduite légale font partie du risque d'exposition assumé par les autres membres. La valorisation des améliorations au réseau ou à sa gestion⁵⁸ peut montrer qu'il y aurait avantage à définir dans certains cas une offre concurrente que le producteur public refuse d'autoriser, mais l'absence de choix sur le marché, les raisons de cette pauvreté de l'offre publique ou du nombre insuffisant de clubs, ne relèvent pas davantage de notre propos.

Comment le modèle de trafic est-il utilisé ? Comme un modèle produit, pour toute valeur observée y dans l'échantillon, une valeur calculée y_c^* , on pourrait penser que faire une prévision avec un modèle consiste à modifier les valeurs des variables explicatives et à produire par simulation une nouvelle valeur calculée y_c^{**} . Et cette méthode de simulation est bien la méthode de prévision habituelle⁵⁹. Mais la méthode utilisée principalement dans l'analyse des projets de transport est la *méthode du pivot* qui consiste plutôt à n'utiliser le

56. Certains pays subventionnent le système hospitalier et le traitement des handicapés lourds, d'autres peu ou prou. Nous faisons aussi comme si aucune perte sociale ou dommage moral n'existaient au-delà des pertes compensées par les assurances parce que ce débat est trop complexe pour une présentation centrée sur la modélisation du trafic. Nous ne souhaitons pas non plus discuter davantage ici de la valorisation des accidents et de leurs victimes.

57. Nous faisons abstraction des problèmes d'asymétrie d'information qui peuvent justifier l'intervention d'un régulateur.

58. Par exemple, la construction d'une autoroute qui interdirait l'accès à tout chauffeur de moins de 35 ans et de plus de 50 ans, indépendamment du sexe du chauffeur, réduirait le risque d'accident mortel de plus de la moitié et serait peut-être rentable sur un marché libre, mais de telles discriminations ne sont pas autorisées par la loi.

59. On peut par exemple traiter la valeur prédite comme un paramètre inconnu et l'estimer avec les autres paramètres du modèle, ce qui produira des écarts-types d'estimation pour ce paramètre comme pour les autres. Pour consulter la documentation complète d'un algorithme qui autorise ce genre de prévision en plus de la prévision habituelle par simulation sur un modèle de niveaux, voir Liem *et al* (2000).

modèle que pour calculer l'effet des modifications envisagées par le projet, $\Delta y_c = y_c^{**} - y_c^*$, et à appliquer cette différence Δy_c aux valeurs y observées en situation de référence ; la valeur prédite est alors : $y^f = y + \Delta y_c$.

On ne connaît pas bien les propriétés statistiques de la méthode du pivot. On sait bien sûr qu'elle évite la situation paradoxale suivante : une valeur prédite pour un lien amélioré y_c^{**} inférieure à la demande en situation de référence y . Ce résultat est possible pour certaines observations lorsque la valeur y_c^* calculée par le modèle pour ces observations est inférieure à la valeur observée y d'une quantité $y - y_c^*$ plus grande que Δy_c , celle qu'on attribue à l'effet de l'amélioration : l'ajustement d'un modèle exige de passer au plus proche, dans un sens ou l'autre, de l'ensemble des valeurs observées – ce qui ne veut pas dire que les demandes sur les liaisons ou liens qui nous intéressent soient particulièrement bien reproduites.

Dans le cas si important pour nous du modèle Logit (agrégé) de parts, l'application du pivot consiste à multiplier la part modale existante par le changement du niveau d'utilité du mode, procédure parfois appelée « Logit incrémental » et que reprend Laferrière (1994) dans son étude des propriétés de l'expression :

$$p_m^f = \frac{p_m^r \cdot \exp(\hat{V}_m^f - \hat{V}_m^r)}{\sum_q p_q^r \cdot \exp(\hat{V}_q^f - \hat{V}_q^r)} \quad \left| \begin{array}{l} \text{pour une liaison donnée } ij \end{array} \right. \quad (32)$$

où, pour une liaison donnée de i à j : (i) le membre de gauche p_m^f est la part prédite du mode m pour la période de prévision f ; (ii) le premier terme du côté droit est la part observée du même mode m durant la période de référence ; (iii) les parenthèses comprennent des différences entre l'utilité représentative estimée avec les coefficients du modèle en période de prévision et en période de référence. Ce sont les changements prévus des variables en situation de projet qui induisent ces différences.

Laferrière montre que, dans la mesure où les erreurs d'un modèle résultent d'une formulation déficiente en variables explicatives (ce qui est généralement le cas), la prévision par la méthode du pivot peut être plus précise⁶⁰ que la prévision qui serait faite à partir des valeurs calculées par le modèle, mais que la variance de cette prévision est plus

60. Il n'existe pas à notre connaissance de preuve comparable pour le Logit désagrégé.

grande. En ces sens, l'usage du pivot a des implications économétriques. Il nous faudrait donc mieux connaître les autres implications de l'usage du pivot, méthode si prépondérante dans le calcul économique.

5.1. *Calcul de l'avantage transport par la méthode du surplus*

On s'attend à ce que les investissements en transport, comme toute hausse du capital, modifient les productivités marginales de toutes les ressources : main d'oeuvre, entrepreneurship, capital industriel, sol, etc. On pourrait alors penser que ce soit une bonne idée de mesurer ces impacts sur toutes les ressources, l'une après l'autre – pas seulement sur la valeur des terrains, ce qui serait incomplet, mais de manière exhaustive. Suite à un tel investissement, par exemple, des personnes prendront un second emploi, ou même un premier, et on devrait calculer l'ensemble de ces effets.

Mais c'est une idée à la fois exhaustive et inutile, sauf pour localiser très approximativement certains effets redistributifs de l'investissement – qui favorise fatalement certaines localisations plus que d'autres. Un tel calcul est inutile parce qu'on sait qu'il est possible de calculer la totalité de l'avantage *sur le réseau lui-même*, calcul qui fournira, pour la surface située sous la courbe de demande de transport, une valeur exactement égale à la valeur nette des surplus des producteurs et des consommateurs de tous les marchés pertinents situés à l'origine ou à la destination des flux.

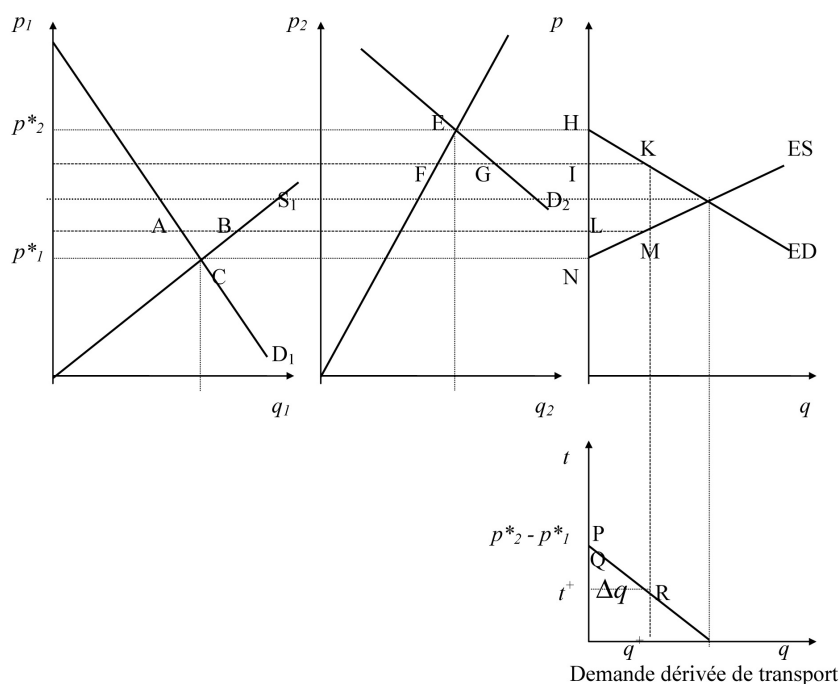
Cette voie, infiniment moins difficile que la première, permet de formuler « l'hypothèse de base » et donne le point de référence de tout calcul d'avantage par la méthode des surplus. On en recense deux formes principales, différenciées par les mots « classique » et « Logit ». La première est présentée graphiquement à la Figure 9.

i) *La méthode classique*. On sait, certainement depuis Tak et Ray (1971) qui s'inspiraient d'une démonstration antérieure de Hicks sur le calcul des surplus dans un réseau de transport de l'électricité, que, si les industries sont concurrentielles aux deux extrémités d'une route reliant deux régions, la somme des surplus occasionnés par la chute du coût de transport t calculée sur le réseau est *exactement* égale à la variation des surplus nets aux deux extrémités – au moins dans le cas linéaire illustré. Sur la Figure 9, chacune des régions produit le même bien (e.g. du blé), mais elles ont des courbes d'offre et de demande très différentes avant qu'il n'y ait commerce entre elles grâce à la baisse du coût de transport. À mesure que se poursuit cette baisse, le bien est exporté de la région **1** vers la région **2** jusqu'à l'égalité des offres et demandes excédentaires *ES* et *ED* caractérisée par des gains et des pertes de surplus très différents selon les régions.

Parmi les variantes de cette démonstration graphique, il faut compter celle de Jara-Diaz (1986) qui étudie par ailleurs l'impact du réseau sur les niveaux de concurrence sur les deux marchés (induisant des gains et pertes nouveaux) et, plus récemment, une version détaillée du même problème avec plusieurs itinéraires (Kidokoro, 2004).

Figure 9

Offre, demande et surplus des usagers du transport entre régions concurrentielles



ii) *La méthode Logit*. Small et Rosen (1981) ont montré que la différence entre les *logsum* (17) de la situation de référence et celle de la situation de projet jouait dans les Logit linéaires un rôle semblable à celle de l'aire présente sous la courbe de demande linéaire classique. Elle s'écrit pour un individu t :

$$\Delta U_t \equiv \frac{1}{\rho} \left\{ \left[\ln \left(\sum_m e^{V_m} \right) \right]_f - \left[\ln \left(\sum_m e^{V_m} \right) \right]_r \right\} \quad (33)$$

où ρ désigne l'utilité marginale du revenu et les indices f et r les évaluations en situation finale et initiale.

Kidokoro (2006) a étudié de près la différence entre ces deux méthodes ainsi qu'entre elles et la méthode de Wardrop. Par ailleurs,

Abraham (2006) compare aussi les méthodes classique et Logit et traite de leurs différences si le choix modal ou l'affectation obéit à la Loi d'Abraham. L'étude des différences entre les surplus calculés par la méthode classique et la méthode Logit est une activité en croissance. On sait en particulier, que la non linéarité de l'influence du revenu « perturbera la formule du *logsum* » (McFadden, 1998) et que la non linéarité des variables compliquera le calcul usuel des différences de surplus entre deux états du réseau puisqu'il faudra tenir compte de la courbure des fonctions de demande dans le calcul des surfaces.

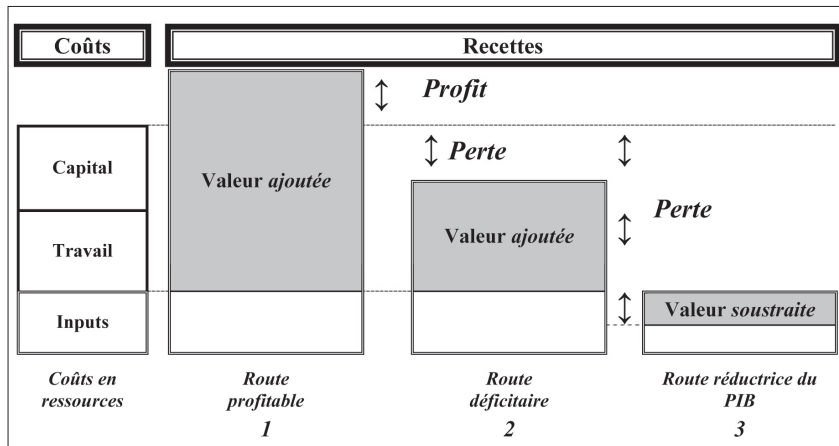
5.2. *Calcul financier et calcul de la valeur ajoutée*

Modèles de trafic et couverture des coûts directs. Un surplus des usagers ne rend pas un projet rentable. Les modèles de trafic fournissent les courbes de demande qui permettent de savoir si les droits et redevances d'usage du nouveau service seront suffisants pour justifier tout projet qui doit s'autofinancer. Le calcul de certaines des recettes spécifiques de transport est parfois simple – par exemple, le point (t_1, q_1) à la Figure 2 implique une recette de taxe spécifique sur le carburant (TIPP en France) – et nous ne souhaitons pas le détailler sauf sur un point précis de recherche sur le sens économique de la couverture des coûts.

La question récemment élaborée (Gaudry, 2001) au bénéfice d'une commission d'enquête sur ces questions est celle de savoir si un tarif de transport n'est qu'un transfert qui enrichit Pierre au détriment de Paul ou s'il peut simultanément les appauvrir tous les deux. Dans la mesure où certains niveaux tarifaires impliquent une réduction du PIB, le calcul de rentabilité financière sera plus intéressant que ce qu'on en dit d'habitude : l'enjeu d'une recette de transport irait au-delà de celui d'un simple transfert dans un jeu à somme nulle. L'argument, illustré au Tableau 4, reprend une formulation du problème, développée à l'origine (The Economist, 1991) pour comprendre comment certains secteurs des économies centralement planifiées réduisent en permanence le PIB, notamment par l'importation de biens intermédiaires qui sont ensuite transformés dans des usines et réexportés à des prix qui impliquent des recettes inférieures au coût des inputs importés pour les produire.

Pour envisager cette possibilité, il faut disposer d'une comptabilité analytique des coûts de construction et d'entretien de l'infrastructure qui nous intéresse, par exemple une route. Au lieu d'en comparer simplement les recettes et coûts totaux, il faut comparer les recettes aux coûts par catégorie comptable d'établissement de la valeur ajoutée de n'importe quelle activité économique.

Tableau 4
Route qui augmente ou diminue le PIB



Une route à valeur ajoutée négative. Dans l'exemple du Tableau 4, les coûts d'infrastructure et d'entretien d'une route, mesurés au prix des facteurs (nets des taxes et subventions), sont comparés aux recettes de transport spécifiques reçues par l'agence routière. Si, comme dans la colonne 2, les revenus ne couvrent pas les coûts du capital, du travail et des inputs (matières premières, pièces, etc.), la route encourt un déficit. Si toutefois, comme dans la colonne 3, les recettes sont inférieures à la valeur des inputs matériels, alors le sel, l'asphalte, le gravier et les gravillons, les pièces et les carburants des véhicules d'entretien et des autres inputs matériels sont transformés en services de valeur moindre que la somme de ces inputs. La route détruit alors continuellement de la valeur ajoutée parce que le coût de ses inputs est supérieur à la valeur de marché son output : elle réduit le PIB en permanence.

Il ne fait aucun doute que beaucoup de routes rencontrent ces conditions et qu'il faut utiliser des modèles de trafic pour évaluer si des baisses de qualité de service (e.g. en enlevant le pavage et en rétablissant une surface de roulement en gravillons ou en terre) impliqueraient des recettes spécifiques plus basses mais sans soustraction de valeur ajoutée. Dans beaucoup de pays, la régionalisation des routes pose des arbitrages de ce genre avec une acuité croissante.

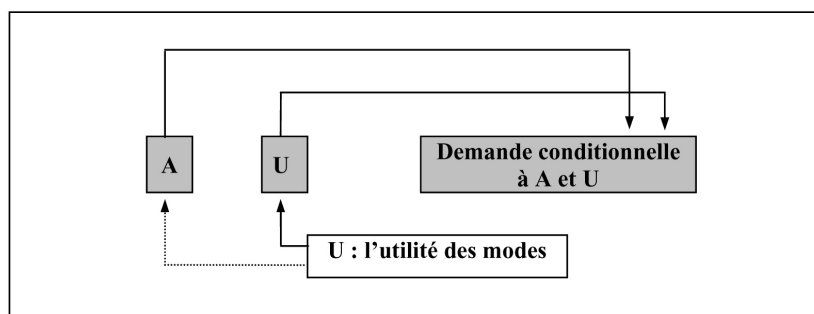
5.3. Rétroactions sur l'utilisation du sol

Une rétroaction partielle. Durant les années 1960, de nombreux chercheurs ont tenté de modéliser l'impact des modifications du réseau sur l'utilisation du sol. Un des pionniers (Lowry, 1964), travaillant sur

un modèle de la région de Pittsburg, s'est attaché à modéliser en région urbaine la relocalisation de la population et des emplois appartenant au vecteur d'activités \underline{A} indiqué à la Figure 10.

Figure 10

Rétroaction de l'utilité transport sur les niveaux d'activité et leur localisation



Anticipations rationnelles des gouvernements. Le modèle de Lowry est encore d'usage courant aujourd'hui ; on pourrait presque dire que c'est le seul modèle de relocalisation utilisé parce que, malgré des efforts importants consentis surtout avant 1980, il est extrêmement difficile de modéliser la localisation des autres activités : on les affecte manuellement dans l'espace. On constate que la maîtrise par l'Etat des fonctions de planification du sol et d'établissement du réseau de transport fait mentir l'idée que les anticipations rationnelles⁶¹ des consommateurs à la Lucas rendraient la politique gouvernementale impossible : c'est plutôt l'inverse qui se passe en ce sens que les demandes de transport réagissent à peu près comme l'anticipent les modèles et donnent *de facto* un pouvoir considérable aux aménageurs du territoire et du réseau.

Nature tendancielle des scénarii économiques. L'usage de méthodes quantitatives pour mesurer les impacts spatiaux a ses limites, car les scénarii en situation de projet sont définis de manière tendancielle, ce qui veut dire que l'impact de la présence du réseau est déjà incorporé à la tendance. Ainsi, prendre en compte certaines rétroactions reviendrait à ne pas avoir bien mesuré la portée du scénario tendanciel, car on opère alors un double comptage.

61. Cette idée consiste à dire que la politique économique sera nécessairement inefficace parce que les agents, ayant une connaissance sans faille de l'économie, anticiperont et compenseront d'avance ses effets.

6. CONCLUSION : SUR QUELQUES THÈMES À APPROFONDIR

Les taux marginaux de substitution (TMS), des résultats dérivés-clés. Le rapport entre les élasticités de la demande de transport par rapport aux prix et au temps des services de transport, exprimé sous forme de taux marginaux de substitution (ou valeurs du temps), est au centre de la modélisation du trafic qui fonde tout calcul économique : il vaut donc la peine de préciser l'importance pratique des TMS.

Les TMS révélés par les modèles et utilisés par les gestionnaires de réseaux diffèrent souvent des valorisations tutélaires de référence ou officielles qui guident le choix des investissements en infrastructures de transport. Le fait que, simultanément, la gestion puisse se fonder sur certains arbitrages réels révélés par le comportement local pertinent décrit par les modèles et que l'investissement puisse s'appuyer sur des arbitrages administratifs normatifs différents, pose un problème de cohérence dans tout calcul économique.

Si cette différence est communément constatée s'agissant des valorisations du temps de transport, elle est présente pour d'autres dimensions des services de transport où le TMS tutélaire n'en est pas moins imposé d'en haut. Par exemple, il y a des raisons de penser que les arbitrages faits au volant entre les gravités d'accidents par les conducteurs sur la route peuvent révéler une valorisation⁶² de la vie humaine supérieure à celle qu'utilisent les planificateurs au bénéfice supposé de ces mêmes usagers. Il est trop tôt pour savoir si les partenariats publics privés (PPP) comblent en partie ces écarts en rectifiant l'offre publique pour la rapprocher du marché.

D'autres sujets avancés ? Les types de données utilisées et les caractéristiques des méthodes de quantification nécessaires à l'obtention de résultats sur les divers paramètres sont des aspects sur lesquels nous n'avons guère été diserts mais dont il faut dire un mot en conclusion, ne serait-ce que pour souligner la structure de ce qui aurait pu être pris en considération au-delà d'un travail de niveau initiatique. Quels sujets spéciaux feraient partie d'un tel approfondissement ? Distinguons réconciliation des *sensibilités mesurées dans le temps et dans l'espace* de la question *des aléas*.

Évolutions temporelles et spatiales des agrégats. Notre présentation des modèles de trafic a volontairement mis l'accent sur les coupes

62. Pour un exemple canadien d'une telle différence entre les valeurs déduites du comportement de conduite et les valeurs officielles, exemple rédigé pour le Groupe de travail Boiteux II, voir Gaudry (2006).

transversales spatialisées et négligé les séries chronologiques, en particulier celles qui visent à expliquer l'évolution des grands agrégats urbains, régionaux ou nationaux du nombre total de déplacements des personnes ou des marchandises représentés par $T = 2\,175$ au Tableau 3 mais souvent conservés sous forme de **Tonnes-km** ou **Voyageurs-km** dans les chroniques. L'évolution de ces agrégats fait aussi appel à d'autres élasticités que celles qui définissent les TMS : les élasticités de la demande par rapport aux niveaux des activités économiques et du revenu, tant à court qu'à long terme.

Si ces modèles en séries chronologiques portant sur des totaux non spatialisés font à juste titre partie de la grande famille des modèles de trafic, ils sont mal reliés au calcul économique appliqué aux projets, sauf pour l'établissement nécessaire des prévisions à long terme et l'étude de leurs fluctuations. Traités comme exogènes aux *scénarii* économiques déjà donnés, ces composants du calcul mériteraient de voir leur rattachement au cadre de calcul être l'objet d'une étude spécifique qui traiterait alors naturellement des erreurs de prévision dans le temps. Quels seraient les chapitres d'un travail de ce genre ?

i) Court terme et long terme. Il est naturel que l'effort de modélisation dont notre analyse rend compte se porte avant tout sur la répartition modale parce que la vitesse d'ajustement des reports entre modes et itinéraires est élevée par rapport à celle des modifications d'activité : dans un calcul de valeur actualisée nette (VAN) où les bénéfices futurs sont escomptés à 8 ou 10%, voire davantage, les effets nets de report dominant normalement les effets nets d'induction. Ces derniers intéressent généralement plus les « politiciens régionaux » et les promoteurs de l'équité spatiale⁶³ que les prévisionnistes. En effet, les rétroactions semblent généralement imputables à des relocalisations plutôt qu'à de nouvelles activités dans les pays avancés où dominant de nos jours les théories sur l'inéluctable convergence régionale, dont les 100 dernières années ont récemment été documentées de manière exceptionnelle pour ces pays avancés (Polèse et Shearmur, 2005).

ii) Temps et espace. La réconciliation des élasticités obtenues à partir des modèles à séries longues sur les agrégats par grande région et celles qui ont été dérivées de coupes transversales spatialisées ferait naturellement partie de cette nouvelle étude des dimensions temporelles des modèles de trafic à unités diverses (tonnes, kilomètres). La

63. Un concept bizarre puisqu'il ne vise pas explicitement des personnes, d'ailleurs mobiles et libres de leur localisation, mais des espaces géographiques ou zones dans notre terminologie. On trouve donc des gens pour lesquels désenclaver les Vosges du Nord ou le Grand Nord du Canada est une question d'équité spatiale sans obligations en retour.

France est un pays favorable à cet égard car on y trouve des séries sur les marchandises-km depuis 1845 (Sauvant, 2002) et sur les voyageurs-km et l'usage des modes de communication depuis 1800 (Grübler, 1990).

Le rôle des aléas. Notre description sommaire des modèles de trafic pourrait porter à croire que toutes les variables y sont *prédéterminées*, au sens statistique de ce mot qu'on oppose à *aléatoire*⁶⁴, mais cela n'est pas le cas dans la pratique de la modélisation même si nous en avons minoré les considérations statistiques. Pour donner généralement une idée juste des modèles et rendre compte de leur comportement, en particulier en prévision, il est certes utile d'explicitier cette dimension de l'aléatoire, tant dans la pratique courante que dans les travaux plus pointus

A. L'aléatoire et les variables expliquées : erreurs d'observation et de spécification. Si nous n'avons pas signalé au passage le caractère aléatoire des élasticités⁶⁵, en présentant (30) et (31), nous avons au moins reconnu le caractère aléatoire fondamental des variables dépendantes ou *expliquées* dans les modèles de niveaux en les faisant dépendre d'une erreur de régression [e.g. (7)-(8) ou (27)] obéissant à des lois de distribution non précisées⁶⁶; et nous avons justifié la forme des modèles Logit par la nature de la distribution de l'erreur associée à l'équation de chaque fonction d'utilité représentative [e.g. (16)]. Dans les deux cas, les propriétés intrinsèques des aléas jouent un rôle doublement figuratif : les erreurs de régression⁶⁷ sont composées d'*erreurs d'observation* sur la variable dépendante (une quantité de déplacements, ou une utilité, selon le type de modèle) et d'*erreurs de spécification* du modèle. Comment cette prise en compte est-elle faite ?

i) *Formulations de distributions des résidus.* L'aléa probabilisé est conjuré et manipulable, et cela sans reste à subir ou traiter. En effet, ce n'est pas encore la pratique de partager l'aléa entre une partie *probabilisée* et un reste *non probabilisable et perçu globalement*, à la

64. En français, on utilise indifféremment la caractérisation latine (*aléatoire*), grecque (*stochastique*) ou arabe (au *hasard*) du jeu de dés pour exprimer cette idée.

65. La question est traitée à l'Annexe 3.

66. Utiliser un critère numérique d'ajustement comme les MCO qui minimisent la somme des erreurs (définies verticalement) au carré entre les observations y et les valeurs calculées y_c^* n'exige aucun postulat sur la distribution des erreurs : on appelle ce calcul une *calibration* des coefficients β ; elle autorise le calcul des mesures de la *sensibilité* numérique de l'ajustement aux valeurs calculées de ces coefficients. C'est l'ajout de l'hypothèse de normalité (ou d'une autre hypothèse statistique) des mêmes erreurs qui fait passer à l'*estimation* de paramètres et à des mesures de *fiabilité* des mêmes coefficients calculés de la même façon et par la même formule.

67. Comme nous l'a précisé Ariane Dupont-Kieffer, le flottement dans l'usage des termes « erreur » et « résidu » est associé à l'évolution de la compréhension de la régression où, ce qui était au départ une erreur d'observation, devient progressivement un résidu aléatoire au sens propre.

manière célèbre de Knight (1921) qui scindait ainsi le risque. Récemment, on a interprété cette partie non probabilisable de l'aléa comme une incertitude sur la nature de la distribution de la partie probabilisable, ou comme une « *contamination-e* » de cette distribution – comme si une famille de distributions inconnues était possible (Nishimura et Ozaki, 2004). Dans les transports, la décomposition pratiquée à ce jour consiste plutôt à supposer que l'erreur, par exemple de (12), est décomposable en une erreur sur la liaison *ij* et des erreurs associées aux origines ou aux destinations *i* ou *j* (e.g. Bolduc et Laferrière, 1992). Comme ces formes de probabilisation ont une influence immédiate sur les mesures de fiabilité comme les *t* de Student (Gosset, 1908), elles jouent un rôle de premier plan dans l'interprétation des résultats des modèles.

ii) *Interdépendance des résidus*. Dans les modèles de trafic, les chercheurs ne supposent pas toujours que les erreurs de régression sont indépendantes entre elles. Nous avons bien fait allusion à deux manières de faire intervenir la corrélation entre ces erreurs. Dans les modèles de niveaux, nous avons rendu compte de résultats obtenus qui prenaient en compte la corrélation spatiale entre les résidus des erreurs par liaison (voir note 17) ; dans les modèles de type Logit, nous avons dit que les structures de corrélation supposées entre les erreurs des fonctions d'utilité modales servaient à définir des branches et des hiérarchies. Comme la corrélation entre les résidus révèle⁶⁸ habituellement l'absence de variables pertinentes, elle est un excellent indicateur de la qualité de la formulation du modèle.

iii) *Hétéroscédasticité des résidus*. Les modèles utilisés dans le cadre classique des modèles de trafic sont souvent plus raffinés que ce que nous avons pu laisser croire. Et ces raffinements ont une grande influence sur les paramètres estimés et les statistiques qui en seront dérivés. Considérons le problème de la constance de la variance de l'erreur de régression, une condition importante d'obtention de statistiques sans biais appelée « *homo-scédasticité* », littéralement « *même variabilité*⁶⁹ ».

Ce problème est inhérent à de nombreux modèles de niveaux dont la variable dépendante, par exemple un flux de transport, est nécessairement importante pour les liaisons entre les grandes villes et faible pour les liaisons entre les petits centres d'activité. Dans ces conditions, il est impossible d'avoir une variance de l'erreur qui soit de même taille

68. Car la corrélation pure est peu probable, voire invraisemblable.

69. Le mot grec *skedasis* signifie variabilité ou dispersion. Nos formules usuelles de variance ne constituent qu'une formalisation possible de la variabilité.

pour l'ensemble de l'échantillon : la variance est hétéroscédastique et des corrections vers l'homoscédasticité sont à envisager, sans quoi l'erreur de régression sera toujours faible pour les petits flux et importante pour les grands flux. La question se corse encore davantage dans les modèles de niveaux où on utilise des transformations Box-Cox de la variable dépendante qui modifient automatiquement la variance de l'erreur⁷⁰ de régression, problème qui se corrige⁷¹ en adoptant des instruments distincts pour la forme et pour l'hétéroscédasticité (e.g. Dagenais *et al.*, 1987 ; Gaudry, 2004).

Le problème se pose aussi dans les modèles Logit. Bhat (2000) a fait remarquer qu'on doit rendre explicite le fait qu'en (15)-(16) l'erreur u_i n'est généralement homoscédastique que par hypothèse : sa variance est alors égale à $\pi^2\mu^2/6$ et constante entre les alternatives. Dans cette expression de la variance, le facteur d'échelle $1/\mu$, qui provient de la distribution sous-jacente de Gumbel $\{[f(u_i) = (1/\mu)[\exp(-u_i/\mu)][\exp(-u_i/\mu)]\}$, est habituellement supposé commun à toutes les alternatives et posé égal à 1 ; il pondère en fait les parts respectives des parties systématique et aléatoire de chaque fonction d'utilité de type (16), exactement comme dans un modèle de niveaux : un μ_i élevé réduit la part systématique de l'explication au profit de la part aléatoire, et cela différemment selon le mode considéré. C'est dire l'importance de la question dans un modèle.

Pour comprendre que l'hétéroscédasticité est inhérente au modèle Logit, il suffit d'imaginer un cas bimodal et de se rappeler que les choix sont codifiés **0** ou **1**. Cela signifie que toute probabilité \mathbf{p}_m^* calculée par le modèle produit une erreur d'ajustement de la forme $[0 - \mathbf{p}_m^*]$ ou $[1 - \mathbf{p}_m^*]$, erreur qui aura généralement deux variances. Domencich et McFadden (1975) avaient utilisé une correction simpliste qui n'a pas convaincu et qui est tombée dans l'oubli depuis.

B. L'aléatoire et les variables explicatives. La prise en compte de l'aléatoire parmi les variables explicatives des modèles de trafic se fait de trois manières principales.

i) *Biais de simultanéité.* On peut corriger le biais d'équations simultanées engendré par la présence dans la fonction de demande d'une variable endogène au système, par exemple la fréquence de service. Un exemple de formulation intuitive de ce problème pourrait être : la

70. Dans les modèles de type Logit Box-Cox, la variable dépendante n'est pas transformée, ce qui s'appelle parfois Box-Tidwell (1962) : la variance de l'erreur peut aussi y être influencée par les transformations appliquées sur les variables explicatives, mais évidemment moins que lorsque la variable dépendante elle-même est transformée comme dans les modèles de niveaux.

71. L'application d'une transformation Box-Cox peut réduire ou induire l'hétéroscédasticité. La correction doit donc être faite simultanément avec l'estimation de la forme véritable.

fréquentation d'une navette aérienne est-elle importante parce que la fréquence est élevée ou est-elle élevée parce que la fréquence est sans égale ? La prise en compte de la simultanéité a un impact important sur les élasticités⁷² et les TMS.

ii) *Erreurs d'observation berksonniennes et classiques.* Il existe d'excellents résumés du problème des erreurs d'observation en régression (e.g. Buzas *et al*, 2003) qui distinguent⁷³ bien entre les erreurs d'observation classiques, qui conduisent à des estimations biaisées des MCO et les erreurs de type Berkson (1950) qui maintiennent les propriétés sans biais de cet estimateur.

Dans les modèles de trafic, on fait grand usage des hypothèses propres à générer des erreurs de type bergsonien : on suppose par exemple que le temps d'attente véritable est égal au temps d'attente mesuré plus une erreur, ou que la distance vraie est égale à la distance routière utilisée plus une erreur (indépendante). Formellement, dans le cas berksonien, la valeur vraie X^v est distribuée autour de la valeur observée X^o , c'est-à-dire que $[X^v = X^o + u]$. Dans le cas classique par contre, la valeur mesurée est distribuée autour de la valeur vraie : la valeur observée est égale à la valeur vraie plus une erreur, c'est-à-dire que $[X^o = X^v + u]$.

iii) *Erreurs de spécification.* Un des problèmes structurants des modèles de trafic est celui du temps de transport, dont on sait qu'il est incertain mais pour lequel les observations de sa variabilité sont rarissimes : soit alors on le traite comme fixe et on prend sa valeur moyenne, ce qui est erroné et donnera des résultats biaisés parce que les mesures manquantes ne sont pas indépendantes⁷⁴ de la moyenne, soit on utilise des subterfuges qui approximent le second moment (l'écart-type ou la variance) et le troisième (l'asymétrie à gauche ou à droite) et on les insère aussi dans la régression.

La théorie économique centrée sur les deux premiers moments nous est peu utile car le troisième moment est connu comme décisif tant pour déterminer le rôle du temps de transport que pour la compréhension du comportement en général : motivation de l'épargne de précaution, marketing des jeux de hasard, description du changement climatique et sensibilisation des populations au risque encouru, pro-

72. Pour une comparaison de 6 estimateurs à information limitée ou complète et prise en compte de l'autocorrélation multiple dans un cadre urbain à trois niveaux et deux modes (VP et TEC), voir Gaudry (1980) mentionné plus haut.

73. Il existe une scholastique asymptotique qui permet de réconcilier ces deux types d'erreurs en un seul.

74. Les distributions empiriques des trois premiers moments du temps de transport observé ne sont pas orthogonales.

pension à faire la guerre, structuration des portefeuilles. En effet, on s'inquiète beaucoup moins d'une hausse ou d'une baisse de 1 % du PIB que d'une baisse de 5 % – qui serait dramatique. Un modèle d'espérance d'utilité fondé sur les deux premiers moments seuls est faux et donnera des résultats d'autant plus biaisés que le troisième moment est plus important que le second.

Des travaux récents (Lapparent, 2004) ouvrent une voie nouvelle en supposant que les usagers des itinéraires de transport qu'ils utilisent fréquemment ont en tête des distributions des temps de transport et, selon qu'ils soient optimistes ou pessimistes, « tordent » dans un sens ou l'autre⁷⁵ la différence entre leur distribution anticipée et la distribution réelle du temps de transport. Les transports sont un excellent champ d'application et d'élaboration de théories sur le risque et l'incertain.

Modélisation du trafic et théorie économique moderne. C'est bien en cherchant des solutions à des problèmes pratiques de transport que des concepts fondamentaux de théorie économique ont été élaborés : Dupuit cherchait à résoudre un problème de tarification d'un pont lorsqu'il a défini sa célèbre courbe de demande ; il a élaboré sa théorie de la concurrence dynamique comme contrepoids à des volontés politiques de réguler les chemins de fer. Ce sont les exigences de réalisme des modèles de trafic qui ont provoqué chez McFadden le développement de la théorie et de l'économétrie des choix discrets. Les besoins d'expression claire et intuitive des relations empiriques ont donné naissance sans fanfare à l'usage des élasticités à l'ENPC, bien avant leur « redécouverte » théorique autoproclamée par Marshall.

Combien de secteurs économiques ont suscité autant d'innovation en sciences économiques ?

Références

- Abraham C., « Amertume et acceptabilité des péages : 1. Les émeutes du Pont d'Arcole », *Les Cahiers Scientifiques du Transport*, 40, 2001, p. 53-66.
- Abraham C., « Calcul socio-économique et infrastructures de transport : Comment calculer le surplus des usagers, et en particulier des usagers reportés ? », Note transmise à l'auteur, 10 p., le 7 juillet 2006.
- Abraham C. et Coquand P., « La répartition du trafic entre itinéraires concurrents : réflexions sur le comportement des usagers, applica-

75. Par des fonctions puissances.

- tion au calcul des péages », *Revue générale des routes et aérodromes*, 357, 1961, p. 57-76.
- Abraham C., Baumgart T. et Blanchet J.-D., « Un modèle de prévision du trafic aérien intérieur », *Revue de l'Aviation Civile*, 1969, p. 15-28.
- Abrahamsson T., « Estimation of Origin-Destination Matrices using Traffic counts – A Literature Survey », IR-98-201, International Institute for Applied Systems Analysis, Laxenburg, May 1998.
- Anscombe F. J. et Tukey J. W., « The criticism of transformations », Paper presented before the American Statistical Association and Biometric Society, Montreal, 1954.
- Arbour D., Bégin P., Gohier J.-P., Latour G. et Thivierge J., « Le modèle de transport », Textes photocopiés présentés au Comité AMUR-ETEN, Service d'Urbanisme, Ville de Montréal, 1969.
- Arbour D., Bégin P., Gohier J.-P., Latour G. et Thivierge J., « Le modèle de transport : schéma », Service d'Urbanisme, Ville de Montréal, 1971a.
- Arbour D., Bégin P., Gohier J.-P., Latour G. et Thivierge J., « Le modèle de transport : simulations métro 1981 », Service d'Urbanisme, Ville de Montréal, 1971b.
- Arduin J. P., « Modèles de prévision de trafic utilisés par SOFRERAIL », SOFRERAIL, Paris, 1989, 17 p.
- Armington P., « A Theory of Demand for Products Distinguished by Place of Production », International Monetary Fund Staff Papers, 1969, p. 159-176.
- Barten A., « Maximum likelihood estimation of a complete system of demand equations », *European Economic Review*, 1, 1969, p. 7-73.
- Baumol W. J. et Vinod H. D., « An Inventory Theoretic Model of Freight Transport Demand », *Management Science*, 16, 7, 1970, p. 413-421.
- Beckmann M., McGuire C. B. and C. B. Winsten, *Studies in the economics of transportation*, Yale University Press, New Haven, Connecticut, 1956 ; also published as Rand-RM-1488-PR, Rand Corporation, Santa Monica, CA, May 12, 1955.
- Bendtsen P. H., « Simplification of Traffic Models », *Traffic Engineering and Control*, 15, 1974, p. 817-819.
- Berkson J., « Are There Two Regressions ? », *Journal of the American Statistical Association*, Vol. 45, 1950, p. 164-180.
- Beuthe M., Demilie L. et Jourquin B., « The International Impact of a New Road Taxation Scheme in Switzerland », Ch. 5 in Beuthe M. et Nijkamp P. (eds), *New Contribution to Transportation Analysis in Europe*, Ashgate, august, 1999, p. 97-117.

- Bhat C. R., « Flexible Models Structures for Discrete Choice analysis », Ch. 5 in Hensher, D. A. and Button K. J. (eds), *Handbook of Transport Modelling*, Elsevier Science, 2000, p. 71-90.
- Bierlaire M., « The total demand scale : a new measure of quality for static and dynamic origin-destination trip tables », *Transportation Research B*, 36, 2002, p. 837-850.
- Bierlaire M. et Toint Ph. L., « MEUSE : an origin-destination matrix estimator that exploits structure », *Transportation Research B*, 29, 1, 1995, p. 47-60.
- Blankmeyer E. C., « A Metro Model », Unpublished manuscript, Economics Department, Princeton University, 1971.
- Blum U., « Benefits and External Benefits of Transport – a Spatial View », in *The Full costs and Benefits of Transportation*, Springer Verlag, Berlin, 1997, p. 219-243.
- Boiteux M. et Baumstark L., *Transports : choix des investissements et coûts des nuisances*, Commissariat Général du Plan, La Documentation Française, Paris, juin 2001.
- Bolduc D. et Laferrière R., « Regression-Based Variance Estimators for the Error Components Model », Publication AJD-47, Agora Jules Dupuit, Université de Montréal, 19 p. June 1992. Téléchargeable de la page Méthodes Quantitatives du site Agora Jules Dupuit www.ajd.umontreal.ca.
- Box G. E. P. et Cox D. R., « An Analysis of Transformations », *Journal of the Royal Statistical Society, Series B*, 26, 1964, p.211-243.
- Box, G. E. P. et Tidwell P. W., « Transformation of the Independent Variables », *Technometrics*, 4, 4, 1962, p. 531-550.
- Boyce D. E., Mahmassani H. S. et Nagurney A., « A Retrospective on Beckmann, McGuire and Winsten's », *Studies in the Economics of Transportation*, University of Massachusetts at Amherst, 2004, 27 p.
- Braess D., « Über ein Paradoxon aus der Verkehrsplanung », *Unternehmensforschung* 12, 1968, p. 258-268.
- Braess D., Nagurney A. et Wakolbinger T., « On a paradox of traffic planning », (*Translation from the original German*), *Transportation Science*, 39, 2005, p. 446-450.
- Bröcker J., « Assessing Spatial Economic Effects of Transport by CGE Analysis : State of the Art and Possible Extensions », october 2000, 19 p. Téléchargeable de la page Transport et Économie du site Agora Jules Dupuit www.ajd.umontreal.ca.
- Brown A. et Deaton A. S., « Models of consumer behaviour », *The Economic Journal*, 82, 328, december, 1972, p. 1145-1236.

- Buzas J. S., Tosteson T. D. et Stefanski L. A., "Measurement Error", Institute of Statistics, Mimeo, Series n° 2544, North Carolina State University, april, 2000, 38 p.
- Cascetta E. et Di Gangi M., « A Multi-Regional Input-Output Model with Elastic Trade Coefficients for the Simulation of Freight Transport Demand in Italy », *Transport Planning Methods*, PTRC Education and Research Services Ltd, Volume P 404-2, 1996, 13 p.
- Cascetta E. et Nguyen S., « A Unified Framework for Estimating or Updating Origin/Destination Matrices from Traffic Counts », *Transportation Research B*, 22, 1988, p.437-455.
- CATS, Chicago Area Transportation Study. Final report in three parts. Study conducted under the sponsorship of [the] State of Illinois, Department of Public Works and Buildings, County of Cook, Board of Commissioners of Cook County. CONTENTS : v. 1, Survey findings, December, 1959 ; v. 2, Data projections, July, 1960 ; Transportation plan, April, 1962. Chicago [1959-1962].
- Chopinnet Y., « Méthodologie de prévision de trafic voyageurs-Liaison transalpine », SNCF, 1997, 18 p.
- Cirillo C., « Optimal Design for Mixed Logit Models », Groupe de recherche sur les transports, Facultés Universitaires Notre-Dame de la Paix, Namur, Association for European Transport and contributors, 2005, 11 p.
- Clavering J. et Kirby H., « Data Accuracy as an Aid to Model Development : towards a Manual of Guidance », *Proceedings of the European Transport Forum*, PTRC Education and Research Services, London, 1994, 11 p.
- COST 318, « Interactions between High-Speed Rail and Air Passenger Transport », Final Report of the Action, European Commission, Directorate General Transport, Brussels, 1998.
- Cournot A-A., *Recherches sur les Principes Mathématiques de la Théorie des Richesses*, réédition, Calmann-Lévy, Paris, 1974.
- Dafermos S., « Traffic equilibrium and variational inequality », *Transportation Science*, 17, 3, 1983, p. 301-318.
- Dagenais M., Gaudry M. et Liem T.C., « Urban Travel Demand : the Impact of Box-Cox Transformations with Nonspherical Residual Errors », *Transportation Research*, B 21, 6, 1987, p. 443-477.
- Dantzig G. B., « On the Shortest Route through a Network », *Management Science*, 6, 2, 1960, p. 187-190.
- Davidson R. et MacKinnon J. G., *Transforming the Dependent Variable*, Ch. 14 in *Estimation and Inference in Econometrics*, Oxford University Press, 1993.

- de Cromières E., « Le positionnement de l'élément transport dans l'organisation logistique », Colloque du 18 novembre 2004 : Transports et réseaux : continuités et ruptures, Conseil Général des Ponts et Chaussées, Paris, Actes du colloque, 40-41, <http://rp.equipement.gouv.fr/bicentenaire-cgpc/home.htm>.
- de Lapparent M., « Optimism, Pessimism toward Risk of Time Losses in Business Air Travel Demand, and Insurance Premium », Publication AJD-77, Agora Jules Dupuit, Université de Montréal, January, 2004, 29 p. Téléchargeable de la page Analyse de la demande du site Agora Jules Dupuit www.ajd.umontreal.ca.
- de Lapparent M., « The choice of a mode of transportation for home to work trips in the French Parisian region : application of mixed GEV models with non linear utility functions », Paper presented at the 11th International Conference on Travel Behaviour Research, Kyoto, August, 2006.
- Deaton A. S., « The analysis of consumer demand in the United Kingdom, 1900-1970 », *Econometrica*, 42, 2, 1974, p. 341-367.
- Debaille S., « Un modèle de reconstitution de matrices origine-destination en milieu urbain », Rapport I.R.T. N° 27, Institut de Recherche des Transports, décembre, 1977.
- Debaille S., « Reconstitution de matrices origine-destination en milieu urbain : le modèle NEMROD », Bulletin RECHERCHE-TRANSPORTS n° 26, Institut de Recherche des Transports, avril, 1979, p. 13-23.
- Deepack E. et Laferrière R., « Demand Model Developments to Assess High Speed Passenger Train Markets on Windsor to Quebec City Corridor », Proceedings of the 20th Annual Meeting, Canadian Transportation Research Forum/Le Groupe de recherches sur les transports au Canada, 1994, p. 598-609.
- DETR, U.K. Department of Transport, Environment and the Regions, « Appendix D. The Statistics Used within ERICA », 1998, 13 p.
- Dial R. B., « A Probabilistic Multipath Traffic Assignment Model which Obviates Path Enumeration », *Transportation Research*, 5, 1971, p. 83-111.
- Domencich T. et McFadden D., *Urban Travel Demand : A Behavioral Analysis*, North-Holland, New York, 1975.
- DOT, U.K. Department of Transport, « Trip Matrix Building and Validation Program Suite, User Documentation », Highway Engineering Computer Branch, HECB/R/27, London, 1982.
- Dupuit J., « De la mesure de l'utilité des travaux publics », *Annales des Ponts et Chaussées*, 8, 2, 1844, p. 332-75.

- Ekelund R. B. et Hébert R. F., *Secret Origins of Modern Microeconomics : Dupuit and the Engineers*, The University of Chicago Press, 1999.
- Ellis R. H. et Rassam P. R., « National Intercity Travel : Development and Implementation of a Demand Forecasting Framework », U.S. Dept. of Transportation Contract T-8 542, Modification No. 1, National Technical Information Service PB 192 455, U. S. Dept. of Commerce, March 1970.
- Erlander S. et Stewart N. F., *The Gravity Model in Transportation Analysis : Theory and Extensions*, volume 3 of Topics in Transportation, VSP, Utrecht, the Netherlands, 1990.
- Finney D. J., *Probit analysis : a statistical treatment of the sigmoid response curve*, Cambridge University Press, Cambridge, 1947.
- Finney D. J., *Probit Analysis*, Cambridge University Press, 1964.
- Florian M. et Gaudry M., « A Conceptual Framework for the Supply Side in Transportation Systems », *Transportation Research*, B 14, 1-2, 1980, p.1-8.
- Florian, M. et M. Gaudry, « Transportation Systems Analysis : Illustrations and Extensions of a Conceptual Framework », *Transportation Research*, B 17, 2, 1983, p. 147-154.
- Florian, M., Gaudry M. et Lardinois C., « A Two-Dimensional Framework for the Understanding of Transportation Planning Models », *Transportation Research*, B 22, 6, 1988, p. 411-419.
- Frank M. et Wolfe P., « An algorithm for quadratic programming », *Naval Research Logistics*, Quarterly 3, 1956, p. 95-110.
- Friesz T. L., Gottfried J. A. et Morlok E. K., « A Sequential Shipper-Carrier Network Model for Predicting Freight Flows », *Transportation Science*, 20, 2, 1986, p. 80-91.
- Gaudry M., *The Demand for Public Transit in Montreal and its Implications for Transportation Planning and Cost-Benefit Analysis*, Ph.D. Thesis, Economics Department, Princeton University, 1973, 188 p.
- Gaudry M., « A Note on the Economic Interpretation of Delay Functions in Assignment Problems », *Lecture Notes in Economics and Mathematical Systems*, 118, 1976, p. 368-381.
- Gaudry M., « Six Notions of Equilibrium and their Implications for Travel Modelling Examined in an Aggregate Direct Demand Framework », in Hensher D. et Stopher P. (eds), *Behavioural Travel Modelling*, Croom Helm, Ch. 6, 1979, p. 138-163.
- Gaudry M., « A Study of Aggregate Bi-Modal Urban Travel Supply, Demand, and Network Behavior Using Simultaneous Equations

- with Autoregressive Residuals », *Transportation Research*, B 14, 1-2, 1980, p. 29-58.
- Gaudry M., « Key Substitution-Complementarity Features of Travel Demand Models, with Reference to Studies of High Speed Rail Interactions with Air Services », in COST 318, Interactions between High-Speed Rail and Air Passenger Transport, Final Report of the Action, European Commission, Directorate General Transport, Brussels, 1998, p. 204-238.
- Gaudry M., « The Four Approaches to Origin-Destination Matrix Estimation : Some Considerations for the MYSTIC Research Consortium », Report for the MYSTIC Project ST-97-SC.2101 of the Commission of the European Communities, Directorate General Transport, Working Paper N° 2000-10, Bureau d'économie théorique et appliquée (BETA), Université Louis Pasteur, septembre, 1999, 23 p. Téléchargeable de la page Analyse de la demande du site Agora Jules Dupuit www.ajd.umontreal.ca.
- Gaudry M., « Government transport policy accountability under the Canadian Transportation Act of 1996 and the Kyoto protocol of 1997 », Déposé le 4 mai 2001 au Comité d'examen de la loi sur les transports au Canada/*Canadian Transportation ACT Review*, Ottawa. Publication AJD-6, Agora Jules Dupuit, Université de Montréal, mai, 2001, 45 p. Téléchargeable de la page Comptes-réseaux du site Agora Jules Dupuit www.ajd.umontreal.ca.
- Gaudry M., « The Robust Gravity Form in Transport and Trade Models », *Zeitschrift für Kanada-Studien* 24, 1, Band 44, 2004, p. 144-161. Téléchargeable de la page Analyse de la demande du site Agora Jules Dupuit www.ajd.umontreal.ca.
- Gaudry M., « Coefficients, taux marginaux de substitution et élasticités dans les modèles de régression classique et logistique : applicabilité aux estimations disponibles pour le projet Seine-Nord Europe », Comité économique, Mission Seine-Nord Europe, Voies Navigables de France, Béthune, 8 novembre 2005, 11 p.
- Gaudry M., « Life, Limb and Bumper Trade-Offs Calculable from Road Accident Models : a Note on a Multimoment Portfolio Analysis and Life Asset Pricing Model (LAPM) », *Transport Reviews*, 26, 4, 2006, p. 501-520.
- Gaudry M., Briand A, Paulmyer I. et Tran C-L., « Choix modal transpyrénéen ferroviaire, intermodal et routier : un modèle Logit Universel de forme Box-Cox Généralisée », Publication AJD-88, Agora Jules Dupuit, Université de Montréal, août 2006, 84 p.

- Gaudry M. et Kowalski J. S., « Demand-Performance-Supply Equilibria in Centrally Planned Economies », *Communist Economies*, 2, 2, 1990, p. 205-221.
- Gaudry M. et Lamarre L., « Estimating Origin-Destination Matrices from Traffic Counts : A Simple Linear Intercity Model for Quebec », *The Logistics and Transportation Review*, 15, 5, 1979, p. 631-642.
- Gaudry M. et Wills M. J., « Estimating the Functional Form of Travel Demand Models », *Transportation Research*, 12, 4, 1978, p. 257-289.
- Gaudry M., Heinitz F., Last J. et Mandel B., « Methodological Developments within the Quasi-Direct Format Demand Structure : the Multicountry Application for Passengers MAP-1 », October, 1998, 67 p., Report for the Strategic European Multimodal Modelling (STEMM) Project, Contract No. : ST-96-SC.301 of the Commission of the European Communities, Directorate General Transport, 1998.
- Gaudry M., Jara-Diaz S. R. et de D. Ortuzar J., « Value of Time Sensitivity to Model Specification », *Transportation Research*, B 23, 2, 1989, p.151-158.
- Gaudry M., Mandel B. et Rothengatter W., « Introducing Spatial Competition through an Autoregressive Contiguous Distributed (AR-C-D) Process in Intercity Generation-Distribution Models within a Quasi-Direct Format (QDF) », Publication CRT-971, Centre de recherche sur les transports, Université de Montréal, June, 1994.
- Gauss C. F., *Theoria Combinationis Observationum Erroribus Minimis Obnoxiae*, Göttingen, Dieterich, 1823.
- Gosset, W. S. ["Student"], « The probable error of a mean », *Biometrika*, 6, 1, 1908, p.1-25.
- Groenewegen P., « Supply and Demand », *The New Palgrave Dictionary of Economics* 4, 1987, p. 553-556.
- Grübler A., *The Rise and Decline of Infrastructures. Dynamics of Evolution and Technological Change in Transport*, Physica-Verlag, Heidelberg, 1990.
- Hall R. E. et McFadden D., *A Disaggregated Behavioral Model of Urban Travel Demand*, Charles River Associates, Cambridge, Massachusetts, 1972
- Hensher D. A. et Greene W. H., « The Mixed Logit Model : The State of Practice », *Transportation*, 30, 2, 2003, p. 133-176.
- Hicks J. R., *Value and Capital*, Second edition, Clarendon Press, Oxford, 1939.

- Jara-Diaz S., « On the Relations Between User' Benefits and the Economics of Transportation Activities », *Journal of Regional Science*, 26, No. 2, 1986, p. 379-391.
- Jentsen T. et Nielsen S. K., « Calibrating a gravity model and estimating its parameters using traffic volume counts ». Paper presented to the 5th Conference of Universities' Transport Study Groups, University College, London, 1973.
- Johnson L. W., « Stochastic parameter regression : an annotated bibliography », *International Statistical Review*, 45, 1977, p. 257-272.
- Johnson L. W., « Regression with random coefficients », *Omega*, 6, 1978, p. 71-81.
- Johnson L. W., « Generalised functional form and random coefficients in transportation research », Australian Transport Research Forum, Sydney, 1979a.
- Johnson L. W., « An Introduction to Generalized Functional Form and Random Coefficients in Transport Modelling », *Environment and Planning, A*, 11, 9, 1979b, p. 1017-1037.
- Jourquin B., *Un outil d'analyse économique des transports de marchandises sur des réseaux multi-modaux et multi-produit. Le réseau virtuel, concept, méthodes et applications*, Ph.D. thesis, Facultés Universitaires Catholiques de Mons, 1995.
- Jourquin B. et Beuthe M., « A decade of freight transport modelling with virtual networks : Acquired experiences and new challenges », GTM-paper, April 2004, in A. Reggiani and P Nijkamp (ed), *Spatial Dynamics, Networks and Modelling*, Edward Elgar, 2006.
- Kau J. B. et Sirmans C. F., « The Functional Form of the Gravity Model », *International Regional Science Review*, 4, 1979, p. 127-136.
- Kidokoro Y., « Cost-Benefit Analysis for Transport Networks », *Journal of Transport Economics and Policy*, 38, 2004, p. 275-307.
- Kidokoro Y., « Benefit estimation of transport projects – a representative consumer approach », *Transportation Research*, B 40, 2006, p. 521-542.
- Knight F. H., *Risk, Uncertainty and Profit*, Houghton, Mifflin, 1921.
- Kornai J., *Growth, Shortage and Efficiency*, University of California Press, Berkeley and Los Angeles, 1982.
- Kornai J. et Weibull J. W., « The Normal State of the Market in a Shortage Economy : a Queue Model », Discussion paper, The Institute for International Economic Studies, Stockholm University, 1977.
- Kurihara M. et Sullivan E. C., « Estimating Origin-Destination Matrices from Roadside Interview Surveys », *Transportation Research*, B, 21, 1987, p. 233-248.

- Laferrière R., « Une Agrégation Nouvelle des Itinéraires de Transport Aérien (ANITA) », Publication CRT-574, Centre de recherche sur les transports, Université de Montréal, février, 1988, 106 p. Téléchargeable de la page Analyse de la demande du site Agora Jules Dupuit www.ajd.umontreal.ca.
- Laferrière R., « A Travel Demand Forecasting Dilemma : To Pivot or Not ? », *The Logistics and Transportation Review*, 30, 1, 1994, p. 21-29. Téléchargeable de la page Analyse de la demande du site Agora Jules Dupuit www.ajd.umontreal.ca.
- Lancaster K. J., « A new approach to consumer theory », *Journal of Political Economy* LXXXLV, 1966, p. 132-157.
- Lancaster K. J., *Consumer Demand : A New Approach*, Columbia University Press, 1971.
- Leonardi G., « The Structure of Random Utility Models in the Light of the Asymptotic Theory of Extremes », WP-82-91, International Institute for Applied Systems Analysis, september, 1982, 45 p.
- Leontief W., *The Structure of the American Economy, 1919-1929*, Cambridge, Mass, 1941.
- Leurent F., « Calcul de deux modèles d'affectation », Rapport de convention SETRA-INRETS, Référence INRETS : 18-98 SETRA, février, 1999, 41 p.
- Levitt S. D. et Dubner S. J., *Freakonomics*, William Morrow, New York, 2005.
- Liem T. C., Gaudry M., Dagenais M. et Blum U., « LEVEL : The L-1.5 program for BC-GAUHESEQ (Box-Cox Generalized AUtoregressive HEteroskedastic Single EQUation) regression and multi-moment analysis », in Gaudry M. et Lassarre S., eds, *Structural Road Accident Models : The International DRAG Family*, Elsevier Science Publishers, Oxford, Ch. 12, 2000, p. 263-324.
- Lill E., « Das Reisegesetz und seine Anwendung auf den Eisenbahnverkehr », Wien 1891.
- Low D., « A New Approach to Transportation System Modelling », *Traffic Quarterly*, 26, 1972, p. 391-404.
- Lowry I. S., « A Model of Metropolis », RAND Memorandum 4025-RC, 1964.
- Mandel B., Gaudry M. et Rothengatter W., « A Disaggregate Box-Cox Logit Mode Choice of Intercity Passenger Travel in Germany and its Implications for High Speed Rail Demand Forecasts », *The Annals of Regional Science*, 31, 2, 1997, p. 99-120.
- Marche R., « Pour mieux comprendre les déplacements interrégionaux de voyageurs : un modèle multimodal de demande : 1^{re} partie : Les objectifs retenus et la conception générale du modèle –

- 2^e partie : Description du modèle et premiers résultats », *Les Cahiers Scientifiques de la Revue Transports*, n° 2., 1980.
- McFadden D., « On independence, structure, and simultaneity in transportation demand analysis », Working paper 7511, Institute of Transportation and Traffic Engineering, University of California, Berkeley, CA, 1975.
- McFadden D., « Modelling the Choice of Residential Location », in Karlqvist A., Lundqvist A., Snickars F. et Weibull J. (eds), *Spatial Interaction Theory and Planning Models*, North-Holland, 1978, p. 75-96.
- McFadden D., « Measuring Willingness-To-Pay for Transportation Improvements », Ch 15 in Gärling T., Laitila T. et Westin K. (eds), *Theoretical Foundations of Travel Choice Modelling*, Pergamon, 1998, p. 339-364.
- McFadden D. et Train K., « Mixed MNL models for discrete response », *Journal of Applied Econometrics*, 15, 2000, p. 447-470.
- McLynn J. M., Goldman A. J., Meyers P. R. et Watkins R. H., « Analysis of a Market Split Model », *Journal of Research of the National Bureau of Standards*, 72-B, #1, 1968.
- Mercenier J., Mérette M., Harvey S. et Karangwa E., « GET-1, A Transport-Oriented Computable General Equilibrium Model of the Canadian Economy », November, 2001, 27 p. Téléchargeable de la page Transport et Économie du site Agora Jules Dupuit www.ajd.umontreal.ca.
- Meyer J. R. et Straszheim M. H., *Pricing and Project Evaluation*, Brookings Institution, Washington, 1971.
- Moore E. F., « The shortest path through a maze ». International Symposium on the Theory of Switching, 1957. Proceedings, Part II, Harvard University, Cambridge, MA. ; also in *Annals, Computation Laboratory* Vol. 30, 1959.
- Morellet O. et Marchal P., « M.A.T.I.S.S.E., Un modèle intégrant étroitement contexte socio-économique et offre de transport », Rapport 203, Institut National de la Recherche sur les Transports et leur Sécurité, décembre, 1995, 97 p.
- Morgenstern O., « Vorläufiger Bericht über einen Hauptteil des Fragenkomplexes bezüglich des Wettbewerbes zwischen Eisenbahn und Auto in Oesterreich », Oesterreichischen Institutes für Konjunkturforschung, Vienne, le 28 août 1936, 102 p.
- Moses L., « The Stability of Interregional Trading Patterns and Input-Output Analysis », *American Economic Review*, 1955.
- Netter M., « Equilibrium and marginal cost pricing on a road network with several flow types », in : *Traffic Flow and Transportation*, Pro-

- ceedings of the Fifth International Symposium on the Theory of Traffic Flow and Transportation, Elsevier, New York, 1972, p. 155-163.
- Nguyen S., « An algorithm for the traffic assignment problem », *Transportation Science*, 8, 1974, p. 203-216.
- Nguyen S., « Estimating an O-D Matrix from Network Data : A Network Equilibrium Approach », Publication # 39, Centre de recherche sur les transports, Université de Montréal, 1977.
- Newton I., *De Motu Corporum in Gyrum*, 1684.
- Nishimura K. G. et Ozaki H., « Search and Knightian Uncertainty », *Journal of Economic Theory*, 119, 2004, p. 299-333.
- Nördling W. von, « Le prix de revient des transports par chemin de fer et la question des voies navigables en France, en Prusse et en Autriche », *Annales des Ponts et Chaussées : Mémoires et Documents*, 6^e Série, 11, 1, 1886, p. 292-303.
- Orro A., Novales M. et Benitez F. G., « Nonlinearity and Taste Heterogeneity Influence on Discrete Choice Model Forecasts », Association for European Transport and contributors, 2005, 18 p.
- Ortúzar J. de D. et Willumsen L. G., *Modelling Transport*, Wiley, 2001.
- Overgaard P., « Development of a simplified traffic model for the city of Silkeborg », O.E.C.D. Seminar in Copenhagen.
- Podkaminer L., « Macroeconomic Disequilibria in Centrally Planned Economies : Identifiability of Econometric Models Based on Theory of Households Behavior Under Quantity Constraints », *Zakład Badań Statystyczno-Ekonomicznych Zeszyt 30*, Warsaw 1989.
- Polèse M. et Shearmur R., *Économie urbaine et régionale. Introduction à la géographie économique*, Paris, Economica, 2005.
- Portes R., Quandt R., Winter D. et Yeo S., « Macroeconomic Planning and Disequilibrium Estimates for Poland, 1955-1980 », *Econometrica*, 55, 1, 1987, p. 19-41.
- Potts R. B. et Oliver R. M., *Flows in Transportation Networks*, Academic Press, 1972.
- Quandt R. E. et Baumol W. J., « « The Demand for Abstract Transport Modes : Theory and Measurement », *Journal of Regional Science*, 6, 2, 1966, p. 13-26.
- Quesnay F., *Tableau Économique*, Versailles, 3^e édition, 1759.
- Quinet E., *Principes d'économie des transports*, Economica, Paris, 1998.
- Rassam P. R., Ellis R. H. et Bennett J. C., « The N-Dimensional Logit Model : Development and Application », Peat Marwick and Mit-

- chell & Co, Washington, D.C., 1970, 42 p. Also in Highway Research Record, 369, 135-147, 1971.
- Rea J. C., Platts J. et Wills M. J., « The Potential for Rail Passenger Services in the Windsor-Quebec City Corridor », RTAC Newsletter, Summer, 1977, p. 9-15.
- Ricardo D., *The Principles of Political Economy and Taxation*, London, 1817.
- Rosen S., « Hedonic prices and implicit markets », *Journal of Political Economy*, 82, 1974, p. 34-55.
- Sauvant A., « Volume et partage modal du transport des marchandises de 1845 à nos jours », Notes de Synthèse du SES, mars-avril, 2002, p. 19-26.
- Sévène, L., « Notes prises par les élèves au cours de chemin de fer, session 1876-1877 », École Nationale des Ponts et Chaussées, Paris, 1877.
- Slutsky E. E. , « Sulla teoria del bilancio del consumatore », *Giornale degli Economisti*, 51, july, 1915, p. 1-26.
- Small K. A. et Rosen H. S. « Applied Welfare Economics with Discrete Choice Models », *Econometrica*, 49, No. 3, January, 1981.
- Spieß H., « A Maximum Likelihood Model for Estimating Origin-Destination Matrices », *Transportation Research*, B, 21, 5, 1987, p. 395-412.
- Steuart-Denham J., *An Inquiry into the Principles of Political Economy*, 1767.
- Tavernier R., « Note sur les principes de tarification et d'exploitation du trafic voyageurs », *Annales des Ponts et Chaussées*, Mémoires et Documents, 6ème Série, 17, 2 , 1889, p. 559-654.
- The Economist, « Economic Focus : The value-subtractors of Eastern Europe », January 5 1991, p. 51.
- Theil H., « The Theory of Rational Random Behaviour and its Application to Demand Analysis », *European Economic Review*, 6, 1975, p. 217-226.
- Transport Canada, « SOMPS : An Outline of PERAM Forecasts for Domestic Travel by Mode », Draft preliminary working paper, Ottawa, Transport Canada, Strategic Planning Group, 1979.
- Truong T. P., *A Theoretical Analysis of Choice Demand and Time Allocation : with Reference to Travel*, Ph. D. Thesis, School of Economics and Financial Studies, Macquarie University, December, 1981, 282 p.
- Tukey J. W., « On the comparative anatomy of transformations », *Annals of Mathematical Statistics*, 28, 1957, p. 602-632.

- Van der Tak H. G. et Ray A., « The Economic Benefits of Road Transport Projects », World Bank Staff Occasional Paper 13, Baltimore, 1971, 42 p.
- Wardrop, J. G., « Some theoretical aspects of road traffic research », Proceedings, Institute of Civil Engineers, PART II, Vol.1, 1952, p. 325-378.
- Warner S. L., *Stochastic Choice of Mode in Urban Travel : A Study in Binary Choice*, Northwestern University Press, Evanston, Ill., 1962.
- Weibull W., « A statistical theory of the strength of materials », Ingeniörs Ventenskaps Akademiens Handlingar, No 151, Stockholm, 1939.
- Weibull W., « A statistical distribution function of wide applicability », *Journal of Applied Mechanics*, 18, 1951, p. 293-297.
- Williams H. C. W. L., « On the Formation of Travel Demand Models and Economic Evaluation Measures of User Benefit », *Environment and Planning*, 9a(3), 1977, p. 285-344.
- Wills M. J., *Development of the Highway Network, Traffic Flow and the Growth of Settlements in Interior British Columbia*, M.A. Thesis, Department of Geography, The University of British Columbia, Vancouver, 1971.
- Wills M. J., *Linear and Nonlinear Estimators of the O-D Matrix*, Ph. D. Thesis, Department of Geography, The University of British Columbia, Vancouver, 1978, 342 p.
- Willumsen L., « Estimation of an O-D matrix from traffic counts : a review », Working paper No 99, Institute for Transport Studies, Leeds University, 1978.
- Wilson A. G., « A Statistical Theory of Spatial Distribution Models », *Transportation Research*, 1, 1967, p. 253-269.
- Wilson A. G., *Entropy in Urban and Regional Planning*, Pion, London, 1970.
- Working E. J., « What do statistical "demand curves show ? », *Quarterly Journal of Economics*, 41, 212-35, 1927.
- Zuylen H. J. van, « A method to estimate a trip matrix from observed link volume counts », PTRC Annual Meeting, 1978.

Annexe 1. Les opérations classiques de l'analyse input-output

Pour reproduire dans notre cadre la séquence classique dans un cas simple, considérons la matrice suivante fondée sur une matrice O-D symétrique :

Tableau 5
Matrice comptable avec flux O-D symétriques

<i>Matrice III</i>		Destination : réseau					Destination : extérieur				Total	
O/D	B	T	M	J	NL	1	2	3	4	FL	TL	
Origine : réseau	B	200			200			200		200	400	
	T	200		400	600	1200	400			400	1600	
	M		400			400		400		400	800	
	J		600			600			600	600	1200	
	NE	200	1200	400	600	2400					1600	4000
Origine : extérieur	1		400					100	300			
	2			400				100	300			
	3	200					100	100				
	4				600		300	300				
	FE	200	400	400	600	1600						
Total	TE	400	1600	800	1200	4000						

et intéressons-nous à la demande sortante (de même structure que la demande entrante). On définit alors ainsi A, la matrice des coefficients « directs » :

$$A = N \cdot B^{-1}, \tag{1.1}$$

où B est la matrice diagonale formée des éléments du vecteur de demande totale sortante TL :

$$B \equiv \begin{vmatrix} \mathbf{TL}_1 & & & \\ & \mathbf{TL}_2 & & \\ & & \mathbf{TL}_3 & \\ & & & \mathbf{TL}_4 \end{vmatrix}$$

en sorte que :
$$NL = A \cdot TL . \tag{1.2}$$

Dans la matrice A, un coefficient a_{ij} a le même sens qu'en analyse input-output : par exemple le coefficient $a_{21} = 200/400$ tiré de la matrice A pertinente au cadre matriciel IV suivant

<i>IV</i>	B	T	M	J	
B	0	200/1600	0	0	≡ A
T	200/400	0	400/800	600/1200	
M	0	400/1600	0	0	
J	0	600/1600	0	0	

signifie que 5 unités de flux passent par Talon pour 10 unités de flux passant par Bourassa de toutes origines.

Et on peut par ailleurs aussi dériver la matrice $(I - A)^{-1}$ des coefficients « indirects » à partir de (2) où l'on substitue (1.2) :

$$TL = A \bullet TL + FL \quad (1.3)$$

d'où l'on tire bien :

$$TL = (I - A)^{-1} FL \quad (1.4)$$

On constate que la matrice $C = (I - A)^{-1}$ des coefficients indirects c_{ij} conserve son sens : par exemple, $c_{34} = 0,125$ signifie que pour 1000 unités passant par Jardin (de toutes origines), 125 unités de plus passent par Montigny même s'il n'existe dans le cadre matriciel IV aucun flux qui va directement de Montigny à Jardin, comme l'indiquent les cases foncées de A et C.

$$\begin{array}{|c|c|c|c|} \hline 1,0 & -0,125 & 0 & 0 \\ \hline -0,5 & 1,0 & -0,5 & -0,5 \\ \hline 0 & -0,25 & 1,0 & 0 \\ \hline 0 & -0,375 & 0 & 1,0 \\ \hline \end{array}
 \quad -1 \quad
 \begin{array}{|c|c|c|c|} \hline 0,6675 & 0,125 & 0,0625 & 0,0625 \\ \hline 0,5000 & 1,000 & 0,5000 & 0,5000 \\ \hline 0,1250 & 0,250 & 0,2500 & 0,1250 \\ \hline 0,1875 & 0,375 & 0,3750 & 0,0125 \\ \hline \end{array}
 \quad \equiv C$$

Il s'ensuit que, pour faire une prévision des flux intermédiaires N^* à partir d'une hausse prévue des éléments des demandes finales sortantes, FL^* , on calcule successivement :

$$TL^* = (I - A)^{-1} FL^* \quad (1.5)$$

et

$$N^* = A \bullet B^* \quad (1.6)$$

Annexe 2. Sur une interprétation économique de la courbe débit-temps

Routes et fonctions de production. Notre point de vue, exprimé ailleurs (Gaudry, 1976) et résumé ici de manière partielle et vulgarisée

pour en tirer des implications de politique économique, est qu'il faut comprendre chaque lien a d'un réseau comme une fonction de production de déplacements Q_a qui utilise comme intrants les (véhicules)-distance D_a et les (véhicules)-temps T_a .

Paradoxalement, ces inputs caractérisent aussi l'unité d'output parce qu'ils définissent le vecteur des composants physiques constitutifs du bien q_a nommé « déplacement par a » : ce bien n'a pas d'autre consistance empirique⁷⁶ que celle des caractéristiques mesurables du vecteur.

Et si c'est en effet un vecteur de *caractéristiques physiques par déplacement* qui est produit par cette sorte d'usine qu'est toute infrastructure⁷⁷ correctement définie (avec sa technique de production, comme on l'a expliqué à la Figure 1), alors chaque déplacement individuel q_a produit sur le lien a est complètement défini par son vecteur d'éléments mesurés en dimensions physiques :

$$q_a \equiv [d_a, t_a], \quad (2.1)$$

où d_a désigne la longueur constante⁷⁸ du lien a , $t_a = t_a(v_a)$ désigne la courbe de débit-temps de parcours qui est fonction de v_a , le débit d'usage $v_a = v_a(Q_a)$ sur ce lien⁷⁹. Ce débit v_a est exprimé par exemple en véhicules par heure et, dans le cas simple du lien unique discuté ici, réfère *de facto* à la quantité de déplacements produite Q_a elle-même parce que l'itinéraire et le lien sont confondus.

La conséquence de la double face de la distance et du temps, à la fois intrants et extrants, est que deux déplacements de même longueur d effectués à des vitesses différentes, *i.e.* **qui ne prennent pas le même temps** sur a , ne forment que deux des biens distincts appartenant à l'ensemble $\{B\}$ que peut produire l'usine a à partir de véhicules-intrants :

$$\{q_{a1} \equiv [d_a, t_{a1}], \dots, q_{a2} \equiv [d_a, t_{a2}], \dots, q_{ab} \equiv [d_a, t_{ab}], \dots, q_{aB} \equiv [d_a, t_{aB}]\}, \quad (2.2)$$

ensemble de toutes les combinaisons accessibles sur et au-dessus de la courbe $t_a = t_a(v_a)$ qui est bien la seule combinaison précisément définie par le temps *minimal* requis⁸⁰ par le flux de débit v_a .

76. Les « lignes de désir » des années 1960 sont en fait les matrices O-D.

77. Nous négligeons le coût du véhicule, tenu pour nul.

78. La formulation est plus générale dans l'étude citée parce que le déplacement y peut exiger un parcours produit sur plusieurs liens connexes reliant une paire O-D.

79. Il s'agit naturellement du débit maximal pour les niveaux donnés des inputs.

80. On peut bien sûr choisir une vitesse inférieure à la vitesse maximum réalisable, comme le font les entreprises de transport en commun.

Fonctions de production et courbes de débit-temps. Il y a deux façons d'obtenir les courbes débit-temps $t_q(v_a)$ à partir de la fonction de production.

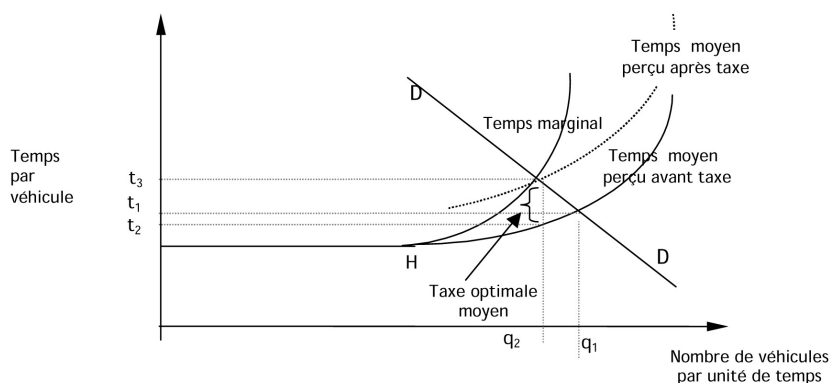
i) *Une projection physique.* La première consiste à transformer la fonction de production à deux arguments en ignorant la distance. Cette projection particulière de la fonction permet de définir pour l'autre dimension un temps moyen et un temps marginal par unité produite, comme à la Figure 11.

Un problème d'interprétation se pose alors parce que l'abscisse désigne bien le taux d'output de déplacements, mais l'ordonnée ne fait que montrer la variation minimale de leur qualité t_a lorsque ce taux croît : naturellement, la qualité de t_a diminue, comme s'étirole avec la foule la qualité « réelle » du bain en piscine publique.

Mais une projection de la fonction de production conserve ses dimensions de nature physique. En conséquence, si l'on propose qu'une taxe égale à la distance entre les deux courbes au point q_2 de la Figure 11 soit imposée pour déplacer, verticalement et d'une hauteur égale à cette distance, la courbe dite de « temps *ressenti ou perçu* » $t_q(v_a)$ jusqu'au point précis d'intersection indiqué par la courbe en pointillés, on demande quelque chose qui n'a pas de sens. On **confond deux univers** de dimensions différentes : il faudrait une taxe définie et payée en temps.

Figure 11

Courbe débit-temps de parcours et taxe optimale



Une telle proposition surprend par ailleurs le lecteur : car la réaction normale à la baisse de la qualité du produit dégradé « output d'équilibre » $q_a^* \equiv [d, t_a^*]$ devrait être de susciter une proposition de baisse de son prix plutôt que celle d'une hausse par une taxe. Dans le même ordre d'idées, qui songerait à gérer une piscine en haussant le tarif payé

par tous à chaque fois qu'arrive un nouveau baigneur ? Qui songerait à imposer à tous ceux qui sont déjà dans la queue d'un guichet bancaire automatisé un tarif variable, qui augmenterait bien sûr chaque fois que la queue s'allonge, et payé à leur arrivée au guichet ?

Ce qui se passe réellement sur le marché est plutôt un ajustement de la nature du vecteur produit et l'élaboration d'une tarification en conséquence. Chaque banque (ou piscine privée) définit une certaine qualité de service (une distribution de probabilités) et fixe un tarif par niveau de service. Sur un marché concurrentiel de biens à qualité variable et dépendante de l'usage⁸¹, les tarifs les plus élevés correspondent normalement à des espérances d'utilité des services plus élevées (à des équivalent-certains plus élevés). Si on paie plus cher, on obtient par exemple une queue en moyenne plus courte ou moins incertaine ; le vecteur de niveaux de service acheté est payé en argent à chaque transaction au guichet.

Alors, si une taxe « optimale » en argent n'a pas grand sens sur une fonction de production, une taxe en nature lui serait-elle préférable ? Peut-on atteindre le résultat visé par la taxe, par exemple par l'ajout de chicanes, rehaussements et rétrécissements à une rue, pour obtenir le niveau et la vitesse qui seraient optimaux en l'absence de ces mesures d'apaisement ? Malheureusement, ces mesures déplacent le *niveau* de la fonction de production vers le haut contrairement à une taxe en argent qui cherche en principe à déplacer non pas la fonction elle-même mais seulement sa *valeur perçue*.

Une chicane modifie la fonction de production et réduit l'output tout en engageant des ressources réelles pour atteindre son but. Une chicane n'est pas une prébende : elle diminue la composante temporelle de l'output produit au nouvel équilibre présumé et exige pour son implantation la saisie de ressources réelles : une double perte de ressources, indépendamment des transferts impliqués. Une chicane fait de la redistribution entre Pierre et Paul tout en les appauvrissant⁸².

Si les relations de demande et de capacité sont exprimées en unités physiques, la taxe n'a pas de sens. Peut-on alors passer en unités économiques ?

ii) *Une projection économique*. Après un rappel de l'objectif d'efficacité visé, examinons successivement les fonctions de demande et de coût, puisqu'il faut des unités communes.

81. La piscine, le guichet bancaire et la route ne sont en rien des biens publics au sens de Samuelson. Ce sont des biens rivaux à qualité variable et dont la consommation d'une unité supplémentaire impose un coût mesurable.

82. Ce qui est compatible avec une amélioration absolue de la situation de l'un ou de l'autre au détriment du premier.

C'est dans un contexte d'omniscience particulier que les économistes⁸³ insistent envers et contre tous pour recommander la taxe vertueuse qui comblera l'écart, naturellement connu de manière exacte à tout moment pour l'ensemble des liens du réseau, entre le coût (moyen) privé valorisé *supposé* et le coût (marginal) social *évident* par unité. La belle taxe déplacera enfin la courbe de coût perçu parallèlement vers le haut jusqu'au nirvana de l'optimum social.

La relation de coût. On rencontre de nombreuses variantes de la courbe débit-temps transformée en courbe de *débit-coût du temps par déplacement*, expression qui s'abrège par « débit-coût temps ».

La dérivation la plus simple à partir de la fonction de production suppose que le prix de la distance est nul et fixe à 1 le prix de référence du temps pour garder intacte l'échelle de l'ordonnée de la Figure 11, rebaptisée « *coût du temps par véhicule* ». Dans ce résultat, le coût moyen du temps par véhicule augmente toujours au-delà du même point *H* de congestion naissante, parce que le composant temps du vecteur (2.1) change suite aux modifications de la nature physique du déplacement.

Ce résultat pose à notre avis un problème non pas tant à cause de la hausse du coût considéré comme tel mais parce que l'output croissant qui provoque cette hausse n'est pas de qualité constante. Le résultat exprime donc en fait un coût par unité *ajusté pour la qualité* (congestion) changeante, et qui devrait être désigné par c^* pour le distinguer du coût à qualité constante c .

Explicitons un ajustement pour la qualité par un exemple simple⁸⁴ de forme multiplicative⁸⁵ :

$$c^* = c \bullet t_a^a \quad (2.3)$$

où t_a^a approxime ici la courbure de la fonction dans le domaine situé au-delà du point de congestion *H*. Par voie de conséquence, le coût marginal est aussi un coût marginal ajusté pour la qualité, si cette notion a un sens. Comment donc interpréter cette fonction de coût (2.3) alors que les fonctions de coût classiques en microéconomie supposent que la *qualité du bien produit est constante* ?

83. Et pas des moindres puisque la liste inclut le canadien William Vickrey qui a contribué à en établir l'argumentaire.

84. Le lecteur intéressé trouvera l'exemple plus général construit avec deux caractéristiques de temps, mais le même genre de fonction multiplicative, dans le rapport COST 318 ou dans Gaudry (1998).

85. Nous ne cherchons pas à reproduire la forme exacte des fonctions de congestion, mais à illustrer le problème.

Pour comprendre qu'une fonction de coût ajustée pour la qualité pose une question de fond, imaginons la courbe de coût moyen de production d'un bien ordinaire, la saucisse, dont la qualité t_s se modifierait à mesure qu'augmenterait le débit de l'usine. Si l'abscisse est définie en kilogrammes de viande, la courbe de coût de production devient intéressante. On pourrait en dire ceci : à bas débit (près de l'origine), la viande produite⁸⁶ n'est en fait que de la *saucisse de Toulouse* qui, à débit moyen, se transforme en *saucisse de Strasbourg* pour enfin devenir, dans les hauts débits, du *Weisswurst*. On voit alors mal comment définir une courbe de coût unique pour cette viande de qualité variable en fonction de la quantité produite, pour ne rien dire de la courbe d'offre. C'est pourtant le sens de la courbe de coût temps obtenue par un procédé analogue : si l'input temps est homogène, l'output déplacement ne l'est pas.

La courbe de coût ajustée pour la qualité peut-elle alors être utile ? La minimisation du coût temps du transport a un sens assez clair pour une firme qui, dotée d'un contrôle centralisé de sa flotte de véhicules, dispose de l'usage exclusif du réseau. Connaissant les prix de ses inputs propres, elle peut en principe minimiser son coût total de transport : (a) soit par une affectation des véhicules optimale au sens *système*, ce qui lui permettra de calculer par lien une taxe virtuelle⁸⁷ implicite ; (b) soit en transformant cette taxe virtuelle en péage réel imposé à ses véhicules, récupéré par elle comme autorité organisatrice, et en supprimant le contrôle centralisé de ses chauffeurs.

Mais le problème se complique pour les usagers de la voiture particulière pour lesquels, contrairement aux taxis et aux transports de fret, les prix de référence des inputs sont davantage des valeurs d'option qu'autre chose, ce qui fait de la courbe de coût temps moyen par véhicule une construction sensible aux valorisations préconisées. Mais admettons qu'il existe pour ce marché des valeurs empiriques applicables au conducteur représentatif de l'offre de conduite *pour toute*⁸⁸ *paire origine-destination*. Le fait demeure que la courbe de coût temps ainsi dérivée correspondra toujours à une quantité dont la qualité est mouvante.

86. À l'évidence une théorie mortadelle digne de *Freakonomics* (Levitt & Dugner, 2005) et qui répond à la question : « Pourquoi les automobilistes détestent-ils la tarification routière au coût marginal social ? »

87. Qui définit une sorte de prix d'ordre (*shadow price*) de l'usage du lien.

88. Ce qui est manifestement faux dans la vraie vie. Pour une part, c'est justement parce qu'ils ont des valeurs du temps différentes que les consommateurs choisissent des moments et des modes de transport variés.

Si donc on s'acharne à vouloir une taxe, il faudra que la courbe de demande **DD** soit aussi exprimée dans des unités conformes à celles de la courbe de coût. À la Figure 11, elle n'est pas ajustée pour la qualité et l'y mettre n'est pas simple.

La relation de demande. La courbe de demande déjà illustrée est la dérivée partielle de la fonction de demande du bien par rapport au composant temps de trajet : elle décrit comment le nombre de déplacements demandé varie lorsque la qualité du service baisse, à longueur et prix par lien donnés. Sa localisation et sa forme dépendent de coefficients tirés de la fonction estimée, par exemple ceux de (C) pour le cas multiplicatif présenté en (2.5).

Pour l'exprimer dans les mêmes unités que la courbe de coût compensé pour la qualité, il faudra partir de la forme estimée et trouver sa nouvelle position sur le graphe quand on l'exprime en prix compensé pour la qualité.

Supposons alors, conformément à (2.3), que la quantité ajustée pour la qualité s'écrive :

$$q^* = q \cdot t_a^{-\alpha} \quad (2.4)$$

et que la fonction de demande exprimée en unités compensées pour la qualité soit de forme multiplicative et s'écrive (A) en (2.5). Comme la forme estimée (C) est obtenue après substitution en (B) du prix compensé par sa valeur (2.4), il nous faut reparamétriser la courbe de demande empirique (C) en termes des paramètres de (2.4). Cela est réalisable dans notre exemple simple dont les opérations :

$$\left. \begin{array}{l} \text{(A) } q^* = \beta_0 \cdot p^{*\beta_1} \cdot y^{\beta_2} \\ \text{(B) } q = \beta_0 \cdot p^{\beta_1} \cdot t^{-2\alpha_1\beta_1} \cdot y^{\beta_2} \\ \text{(C) } q = \beta_0 \cdot p^{\beta_1} \cdot t^{\gamma_1} \cdot y^{\beta_2} \end{array} \right\} \quad (2.5)$$

impliquent qu'il est possible d'exprimer (C) ou sa dérivée partielle en valeurs ajustées car, en (C), $\alpha_1 = -\gamma_1/2\beta_1$.

Il n'est donc pas trivial, même dans un cas multiplicatif simple, de relocaliser la courbe de demande **DD** en fonction du prix compensé. Cet exercice terminé, on pourra établir une taxe dans la nouvelle

version de la Figure 11 dont l'ordonnée montrera des coûts et des prix compensés.

Si on accepte notre interprétation de la fonction de production présentée plus haut, les propositions des économistes impliqueraient donc de tenter l'égalisation d'un coût et d'un prix *ajustés pour la qualité* et dont le niveau d'équilibre est établi par une courbe de demande aux unités conformes à celle des fonctions de coût. Il est possible que leurs propositions s'accompagnent d'un manuel pratique utilisable pour des fonctions de demande et d'ajustement qualité plus générales que la forme multiplicative si particulière, mais alors ce tour de force est mal connu.

Politique de tarification routière. On peut alors comprendre la méfiance de l'électeur médian. Coincé entre les fonctions de production exprimées en unités physiques (où la taxe n'a pas de sens) et les fonctions de coût exprimées en unités ajustées pour la qualité (où la taxe par lien, utilisé par un public O-D aux valeurs du temps plurielles, n'est guère calculable), peut-être a-t-il l'intuition, reflétée dans le discours politique, qu'il y a quelque part un problème de définition du bien.

On le comprendrait de ne pas être encore très ouvert à une discussion sur le « coût marginal » ou le « coût social » de son déplacement, discussion qui lui semble inachevée si la courbe débit-temps sur laquelle repose l'argumentaire n'est pas vraiment une courbe d'offre pour un bien homogène de qualité constante. Aussi préfère-t-il parfois aux péages optimaux recherchés des péages approximatifs qui reflètent très grossièrement les variations normales de la valeur des ressources dans le temps, comme le fait d'ailleurs toute tarification de pointe.

Son interrogation sur la nature des unités des services de transport a des liens de parenté avec l'objet de la querelle, qui a duré de 1891 jusqu'en 1920 et au-delà, entre F. W. Taussing et A. C. Pigou sur la nature de l'output en transport.

Nous pensons qu'il y a *quelque chose* de la nature vectorielle des services de transport et de l'hétérogénéité des valeurs du temps des usagers qui exige encore d'être pris en compte dans cette littérature sur la courbe de débit-temps et les équilibres de production ou de consommation recherchés.

Annexe 3. Élasticités fixes et stochastiques : un rappel

Deux modèles représentatifs. Les deux formes typiques des modèles de trafic sont les modèles classiques de niveaux (e.g. (26)) et les

modèles Logit probabilistes ou de parts (*e.g.* (15)) dont nous souhaitons discuter des élasticités en prenant en compte l'usage des transformations de Box-Cox.

En nous inspirant de Gaudry (2005), nous nous intéressons donc au modèle de régression classique linéaire et à sa généralisation Box-Cox, respectivement :

$$y_t = \sum_{k=1}^K \beta_k \cdot X_{kt} + u_t. \quad (3.1-A)$$

et

$$(\lambda_y) y_t = \sum_{k=1}^K \beta_k \cdot X_{kt}^{(\lambda_{X_k})} + u_t. \quad (3.1-B)$$

où la variable dépendante y est supposée strictement positive, l'indice k est associé aux variables explicatives et les indices t des observations ont été conservés.

Nous nous intéressons aussi au modèle Logit Linéaire explicatif de la probabilité de choix ou de la part de marché de l'alternative i , et à sa généralisation Box-Cox Logit Standard, respectivement, comme aux équations (15) et suivantes :

$$p(i) = \frac{\exp(V_i)}{\sum_{j \in C} \exp(V_j)} \quad (3.2-A1)$$

$$V_i = \beta_{i0} + \sum_n \beta_m^i X_n^i + \sum_s \beta_{is} X_s + u_i \quad (3.2-A2)$$

où les variables explicatives utilisées dans la fonction d'utilité représentative d'une alternative V_i sont partagées entre des variables X_n descriptives de celle-ci et des variables X_s communes aux alternatives, et :

$$V_i = \beta_{i0} + \sum_n \beta_m^i X_n^{(\lambda_{im}^i)} + \sum_s \beta_{is} X_s^{(\lambda_{is})} \quad (3.2-B)$$

où nous avons négligé les indices t des observations pour alléger l'écriture.

Rappel intuitif de la transformation de Box et Cox. Dans ces deux formulations (3.1-B) et (3.2-B), la transformation Box-Cox de toute variable dépendante ou explicative Var_v

$$Var_v^{(\lambda)} \equiv \begin{cases} \frac{(Var_v)^\lambda - 1}{\lambda} & , \text{ if } \lambda \neq 0, \\ \ln(Var_v) & , \text{ if } \lambda \rightarrow 0. \end{cases} \quad (3.3)$$

est bien définie de façon usuelle⁸⁹. On remarque en passant que, dans cette transformation :

(i) le rôle principal du terme $-1/\lambda$ est d'assurer la continuité de la transformation à 0 : une simple transformation puissance impliquerait que la variable transformée tende vers 1, plutôt que vers son logarithme naturel, lorsque son λ propre tend⁹⁰ vers 0. À toutes autres fins utiles, l'interprétation du λ est la même que celle d'une simple puissance ;

(ii) dans (3.1-B), ou d'ailleurs dans (3.2-B), la dérivée partielle de la variable dépendante par rapport à un X_k explicatif est de forme $\beta_k X_k^{\lambda-1}$, ce qui implique que le signe de λ n'a pas d'effet sur le signe anticipé du b_k parce que $X_k^{\lambda-1}$ est toujours positif.

Variables explicatives continues et notion simple ou « échantillonnale ». Si l'on s'intéresse à une variable explicative continue X_k , la notion simple de l'élasticité est le ratio de la variation en pourcentage de la variable dépendante y et de la variation en pourcentage de la variable explicative d'intérêt, c'est-à-dire :

Élasticité de y par rapport à $X_k \equiv$	$\frac{(\text{Variation de } y / \text{niveau de référence de } y) 100}{(\text{Variation de } X_k / \text{niveau de référence de } X_k) 100}$	(3.4)
--	---	-------

Plus précisément, pour une fonction quelconque :

$$y \leftarrow f(X_1, \dots, X_k, \dots, X_K) + e, \quad (3.5)$$

89. Nous ignorons donc ici le coefficient m défini dans l'article de Box et Cox (1964) et nous en tenons à la définition commune. Par ailleurs, lorsque la variable dépendante y n'est pas transformée mais que les variables explicatives le sont, on dit parfois Box-Tidwell (1962) plutôt que Box-Cox, mais nous négligerons cette pratique.

90. La démonstration de la limite lorsque l tend vers 0 est faite en utilisant la règle de l'Hospital.

on écrit :

$$\frac{\% \Delta y}{\% \Delta X_k} \equiv \frac{\frac{\Delta y}{y^r} \times 100}{\frac{\Delta X_k}{X_k^r} \times 100} = \frac{\Delta y}{\Delta X_k} \cdot \frac{X_k^r}{y^r}, \quad (3.6)$$

où y^r et X_k^r sont les niveaux de référence de y et de X_k . Mais on préfère généralement à cette mesure d'arc la mesure en un point :

$$\eta (y , X_k) = \left. \frac{\partial y}{\partial X_k} \cdot \frac{X_k}{y} \right|_{X_k^r, y^r, X_\ell^r}, \quad (3.7)$$

où la ligne verticale signifie que la dérivée est « évaluée au point $y = y^r$ et $X_k = X_k^r$ » pour les variables d'intérêt y et X_k et « au point $X_\ell = X_\ell^r$ » pour les autres variables appartenant à (3.5). En effet, même si cette fonction est linéaire et que la dérivée partielle en (3.7) est égale au coefficient β_k , l'élasticité variera selon le ratio des valeurs de référence choisies.

Cette notion simple est souvent attribuée⁹¹ dans la littérature à Marshall qui l'aurait découverte en 1882 lors d'un séjour à Palerme mais ne l'aurait rendue publique qu'en 1890 dans la première édition des *Principles of Economics*. Elle a l'avantage d'être un nombre pur (sans unités) et d'être intuitive : on voit immédiatement si une élasticité est convaincante.

Dans le cas où l'on estime une fonction, le rapport des valeurs de référence peut être un point particulier de l'échantillon ou un ensemble de points, par exemple le point représentatif des valeurs moyennes des variables. Sauf exception, par exemple logarithmique, l'élasticité dans les modèles de régression classique et logistique varie selon le point d'évaluation. On l'évalue souvent au point moyen de l'échantillon, ce qui a pour avantage que la valeur moyenne de y correspond à l'espérance mathématique de cette variable si la fonction (3.5) est linéaire.

91. La notion d'élasticité était commune dans les cours à l'ENPC (Sévène, 1877 ; Tavernier, 1889) et utilisée dans sa formulation mathématique par un élève de l'ENPC qui publiait sur les coûts ferroviaires dans les *Annales des Ponts et Chaussées* (Nördling, 1886) comme l'ont vérifié Ekelund et Hébert (1999) qui laissent entendre qu'elle aurait pu être empruntée par Marshall qui lisait le français et citait même Dupuit sur d'autres points. *Caveat indigator !*

En fait, il est sage de reconnaître d'emblée que l'évaluation peut généralement être faite en n'importe quel point t de l'échantillon et d'incorporer cette information à la formule (3.7) qui devient alors :

$$\eta_t (y, X_k) = \frac{\partial y}{\partial X_k} \cdot \frac{X_k}{y} \Bigg|_{X_{kt}, y_t, X_{lt}} \quad (3.8)$$

En conséquence, il ne faudra pas confondre l'évaluation au point moyen d'un échantillon et la moyenne des évaluations faites chacune en un point de l'échantillon. L'élasticité « à la moyenne » n'est pas l'élasticité moyenne calculée en faisant la moyenne des élasticités ponctuelles.

Reconnaissance du caractère aléatoire de y . En fait l'expression (3.8) traite implicitement la fonction (3.5) sans reconnaître le caractère aléatoire de y . Puisque cette variable est bien aléatoire, on peut s'intéresser rigoureusement à tous ses moments. Pour les deux premiers moments, on écrirait, respectivement pour le premier et le second :

$$\eta_t (E(y), X_k) = \frac{\partial E(y)}{\partial X_k} \cdot \frac{X_k}{E(y)} \Bigg|_{X_{kt}, E(y_t), X_{lt}} \quad (3.9)$$

$$\eta_t (\sigma(y), X_k) = \frac{\partial \sigma(y)}{\partial X_k} \cdot \frac{X_k}{\sigma(y)} \Bigg|_{X_{kt}, \sigma(y_t), X_{lt}} \quad (3.10)$$

Plus généralement, on peut s'intéresser aussi à la médiane de y ou à d'autres moments de y , comme le troisième, mais ces applications sont rarissimes même si, par exemple, l'impact de X_k sur le point médian est souvent d'intérêt pratique en économie. En fait, la mesure elle-même de l'élasticité devient aléatoire puisqu'elle incorpore des variables aléatoires. On peut alors définir des écarts-types pour l'élasticité, mais nous ne développerons pas ici ces idées.

Remarquons en passant que la probabilité de choix dans le modèle Logit est aussi traitée comme une variable déterministe, même si on peut relier la forme (3.2) à des aléas de l'utilité individuelle, dans la

tradition dite « de l'utilité aléatoire ». En effet, personne ne s'intéresse à $E(p_i)$, l'espérance de la probabilité de choix, ou à d'autres moments de cette probabilité comme sa variance⁹² ou son asymétrie. À notre connaissance, on ne trouve dans la littérature aucune formulation qui traite des moments de la probabilité ou de la part du modèle (3.2) de manière analogue à (3.9) ou (3.10).

Notion simple dans le cas du modèle Logit. L'application mécanique de la notion non aléatoire simple à la probabilité ou à la part de marché produit une expression correspondant exactement à (3.8) :

$$\eta_t(p_m, X_k) = \frac{\partial p_m}{\partial X_k} \cdot \frac{X_k}{p_m} \Bigg|_{X_{kt}, p_{mt}, X_{lt}} \quad (3.11)$$

qui est bien la formule « élasticité-parts » utilisée le plus souvent avec le modèle Logit, sauf pour de rares évaluations de la forme « élasticité-points de pourcentage » dont nous parlerons maintenant.

La notion de points de probabilité ou de pourcentage. Comme une part ou une probabilité est déjà un pourcentage, il serait préférable, pour retrouver l'intuition première exprimée en (3.4), d'utiliser la notion de « points de pourcentage » (en agrégé) ou de « points de probabilité » (en désagrégé) plutôt que la mesure (3.11) d'élasticité de la part ou de la probabilité. La nouvelle mesure est obtenue en multipliant (3.11) par la part de référence (au point t) :

$$\pi_t(p_m, X_k) = \frac{\partial p_m}{\partial X_k} \cdot \frac{X_k}{1} \Bigg|_{X_{kt}, p_{mt}, X_{lt}} \quad (3.12)$$

et elle correspond bien à ce que l'on fait quand on discute de valeurs obtenues en appliquant (3.11) puisqu'il est alors nécessaire de se rappeler du niveau de référence de la part ou de la probabilité pour l'interpréter (3.11) : gagner 5 points donne une forte élasticité si la part de marché du mode considéré représente 10 % du marché mais donne une élasticité faible si elle est de 90%. Il est plus simple et plus clair de calculer le nombre de points, comme le fait (3.12).

92. Ici comme plus haut en (3.1) ou (3.5), il ne faut pas confondre l'intérêt porté à la variance de l'erreur de régression (le problème de l'hétéroscélasticité) avec celui de la variance de la variable dépendante.

Les variables catégoriques dans les deux types de modèles. Même si la dérivée d'une variable auxiliaire (catégorique ou binaire) X_d n'existe pas et qu'il est impossible de calculer (3.7), il est toutefois possible de calculer l'élasticité arc (3.6) d'une variable catégorique en utilisant son coefficient de régression (dans (3.8), (3.11) ou (3.12)) et X_d^* , la valeur moyenne des observations sur les valeurs positives de X_d , comme cela est démontré dans Dagenais *et al* (1987) pour le cas classique (3.7). Cette méthode donne pour le cas de l'élasticité-points propre (3.12) :

$$\beta_c \cdot X_d^* \cdot (1 - p_m) \cdot p_m \quad (3.13)$$

Dans le cas ordinaire d'une variable continue affectée d'une transformation Box-Cox, on calculerait simplement pour (3.13) :

$$\beta_v \cdot X_v^{\lambda_v} \cdot (1 - p_m) \cdot p_m \quad (3.14)$$

où X_v doit être évalué quelque part, par exemple à la valeur moyenne⁹³ de cette variable.

On peut donc, grâce à (3.13) et (3.14), calculer l'élasticité-point directe (propre) pour toute variable explicative d'un modèle Logit, que cette variable soit continue ou catégorique, comme on peut le faire pour un modèle classique. Cette possibilité de calcul des élasticités de toutes les variables d'un modèle fait de ce filtre un outil complet d'évaluation intuitive des résultats des modèles.

93. En principe, les parts p_m à utiliser pour ce calcul sont les parts estimées p_m^* ; si les parts observées sont utilisées pour évaluer la fonction, une erreur systématique égale à $(p_m - p_m^2 - p_m^* + p_m^{*2})$ est introduite. Pour les cas d'évaluation à la moyenne, l'usage des parts observées plutôt que des parts calculées introduira une erreur proportionnelle commune aux élasticités calculées pour toutes les variables.

CHAPITRE 2

ANALYSE COÛT-BÉNÉFICE DANS UN CONTEXTE DE CONCURRENCE INTERMODALE ET INTRAMODALE

Marc Ivaldi¹ et Catherine Vibes¹

Ce chapitre propose une approche pour évaluer le trafic d'une nouvelle infrastructure et son impact sur le bien-être social en tenant compte explicitement de ses effets sur les concurrences intermodale et intramodale entre les opérateurs, cela bien sûr dans le cas où cette concurrence est imparfaite. Une nouvelle infrastructure a pour effet de modifier le niveau de qualité de service proposé sur le marché et donc de changer les conditions de concurrence sur ce marché. Autrement dit, il s'agit de décrire le nouvel équilibre économique, ce que ne font pas habituellement les modèles de prévision de trafic qui ne tiennent pas compte des réactions des acteurs aux actions de leurs concurrents. Les modèles de trafic usuels supposent les prix de transport donnés de façon exogène et en général égaux aux coûts. Cette hypothèse n'est plus vraie dans les cas de plus en plus nombreux où les opérateurs fonctionnent sur un marché amont oligopolistique. Il faut alors modéliser non seulement le comportement des usagers mais aussi celui des opérateurs, les deux étant liés. Non seulement les modèles de trafic en sont affectés mais aussi les calculs de surplus. Nous pouvons donc associer à cet équilibre en termes de trafic une mesure du bien-être social et ainsi évaluer différents projets, les comparer, en tenant compte des interactions stratégiques entre les firmes sur le marché concerné. La méthodologie présentée ici est très générale, mais sa mise en œuvre peut exiger beaucoup de données ; aussi nous proposons dans l'application particulière présentée une procédure parcimonieuse en données, ce qui la rend opérationnelle dans de nombreux cas. Elle évite la lourde tâche que représenterait une collecte d'informations

1. École d'économie de Toulouse (Toulouse School of Economics, TSE)

précises sur le marché et ses différents acteurs tout en permettant de traiter les enjeux essentiels d'une analyse coût-bénéfice d'une nouvelle infrastructure.

Considérons le marché du transport des passagers à longue distance en Europe. Les processus de libéralisation achevés ou amorcés ces dernières décennies dans les industries de transport aérien et ferroviaire ont eu respectivement pour effet une modification du réseau aérien en étoile entraînant l'apparition des compagnies à bas coûts sur les liaisons secondaires et le développement de la technologie des trains à grande vitesse. Ainsi non seulement les compagnies aériennes et les opérateurs ferroviaires font face séparément à une concurrence intramodale, mais ces changements structurels et technologiques introduisent une concurrence intermodale sérieuse entre compagnies aériennes et compagnies de chemin de fer. Il s'agit alors pour nous de considérer le transport entre une origine et une destination comme un marché avec des produits différenciés en concurrence les uns avec les autres.

Dans ce marché de produits différenciés, nous modélisons l'offre et la demande afin de retrouver les résultats de l'équilibre. La demande de marché provient d'une classe générale de modèles de choix discrets du comportement du consommateur. Les voyageurs peuvent tout d'abord choisir le mode par lequel ils désirent voyager, puis pour ce mode, décider d'un service de transport en particulier. Pour cette structure de choix discret, une spécification appropriée pour une analyse empirique est le modèle logit emboîté. Du côté de l'offre, les entreprises déterminent les prix et qualités de leurs services. Les caractéristiques du marché sont dérivées d'un contexte d'équilibre de Nash. Les entreprises se comportent de manière stratégique : chacune d'entre elles détermine ses prix sachant que ses concurrents font de même.

Une fois l'équilibre de marché établi, nous sommes à même d'évaluer l'impact de la mise en place, par exemple d'une nouvelle infrastructure ferroviaire sur le parcours. Celle-ci est alors considérée comme un nouveau concurrent sur le marché proposant un service d'un niveau de qualité différent de ceux des alternatives déjà en place. Les entreprises intègrent cette nouvelle configuration dans les services proposés aux voyageurs et réagissent en modifiant leur prix. Nous pouvons donc mesurer les effets de ce changement structurel en termes de variations des prix, calculer les parts de marché qui leur sont associées et finalement déterminer le surplus du consommateur. Ces mesures des coûts et bénéfices pour les différents acteurs économiques sur

le marché permettent alors d'évaluer et de classer différents projets d'aménagement des transports dans une zone géographique donnée.

Pour faciliter l'exposition de notre approche, nous utilisons le marché du transport de passagers entre Cologne et Berlin comme illustration. Sur ce marché particulier s'exercent des concurrences intermodale et intramodale qui correspondent bien au contexte dans lequel nous nous situons.

Ce chapitre est organisé de la manière suivante. La section 1 décrit en détail les données nécessaires pour appliquer notre méthodologie sur l'exemple du lien entre Cologne et Berlin. La section 2 présente le modèle offre et demande sur ce marché et examine les particularités de l'équilibre. La section 3 fournit les résultats de la simulation de la mise en place de la nouvelle infrastructure destinés à permettre la prise de décision quant à la réalisation d'un tel projet. Pour finir, une conclusion est présentée dans la section 4.

1. DONNÉES SUR UN EXEMPLE GÉNÉRIQUE

Nous appliquons notre modélisation à un exemple spécifique : la concurrence modale dans le trafic de passagers sur le lien Cologne-Berlin en Allemagne. L'évaluation de la construction d'une infrastructure ferroviaire supplémentaire est particulièrement pertinente sur cette liaison pour plusieurs raisons. Tout d'abord, il s'agit d'un trajet de 600 kilomètres connectant deux des plus importantes villes d'Allemagne, d'ouest en est, et représentant une part significative du trafic total des passagers longue distance en Allemagne. Deuxièmement le transport ferroviaire a tout à fait sa place dans ce couple origine-destination pour lequel la diversité des voyageurs (affaires, loisirs...) explique probablement la différenciation des services proposés : une compagnie aérienne nationale Lufthansa (LH), trois compagnies à bas coûts, l'opérateur ferroviaire en place Deutsche Bahn AG (DB), la menace d'entrée d'autres compagnies de chemins de fer, et la route. Enfin le découpage du territoire en différentes autorités régionales (Länder) contraint l'opérateur DB à desservir un certain nombre de gares présentes sur le trajet. La technologie des trains à grande vitesse ne peut alors pas être utilisée de façon optimale sur cette liaison dont la distance permettrait une réelle concurrence en termes de temps entre les transports ferroviaire et aérien. Ainsi, une nouvelle infrastructure destinée à un service express entre Cologne et Berlin permettrait de proposer une option attractive et adaptée à la concurrence croissante des autres modes, tout en conservant la desserte nécessaire de certaines gares sur le parcours grâce à l'installation initiale.

La plupart des données de trafic, de parts de marché, des caractéristiques et prix des alternatives a été fournie par Deutsche Bahn AG pour l'année 2003. Pour chaque mode de transport et les deux catégories principales de passagers (affaires et loisirs), le prix mesuré en euros correspond au prix le plus fréquemment payé par un consommateur représentatif pour un trajet entre Cologne et Berlin. En ce qui concerne les caractéristiques des services, nous utilisons la vitesse mesurée en kilomètres par minute (km/mn), la fréquence quotidienne, et la capacité (sièges disponibles par voyage). Certaines valeurs de coûts marginaux sont aussi fournies par DB. Des statistiques descriptives sont présentées dans le Tableau 1.

Nous considérons que cet ensemble de données correspond au minimum d'information nécessaire pour calibrer ce modèle. Il ne permet pas d'aller plus loin dans l'analyse et notamment il ne permet pas de tester la robustesse des hypothèses, tâche qui nécessite une quantité plus grande et plus variée d'information.

Tableau 1
Données sur la liaison Cologne-Berlin

		Parts de Marché %	Parts des Alternatives %		Prix Euros		Vitesse Km/H	Fréquence Voyages/ Jour	Capacité Sièges/ Voyage	Coût Marginaux*	
			Affaires	Loisirs	Affaires	Loisirs				Affaires	Loisirs
Rail	DB	31.1	15.7	51.4	90.0	60.0	138.5	16	736	na	13
Air	DBA	14.8	19.2	9.1	169.0	51.2	404.6	8	136	na	26
	HLX	13.8	17.8	8.4	169.0	46.7	404.6	5	169	na	26
	GW	8.2	10.6	5.0	169.0	46.1	404.6	4	111	na	26
	LH	14.3	18.5	8.7	240.0	53.4	404.6	9	123	na	na
Route	Voiture	17.9	18.2	17.4	110.0	80.0	114.5	Infini	5	na	na

* Estimations fournies par DB.

2. MODÈLE ÉCONOMIQUE : SPÉCIFICATION ET CALIBRATION

Dans cette section, nous présentons le cadre qui nous permet de décrire la concurrence entre les différentes alternatives de transport dans l'exemple de la liaison entre Cologne et Berlin. Dans un tel contexte de produits différenciés et de structure oligopolistique, les entreprises se livrent à une concurrence à la fois en termes de prix et de caractéristiques des produits. Ici nous nous concentrons sur la concurrence en prix, en supposant qu'à court terme, les caractéristiques des

produits sont fixes. Nous décrivons d'abord le côté de la demande puis nous dérivons les stratégies de tarification des entreprises impliquées dans une concurrence à la Bertrand.

2.1. La demande

Nous considérons les choix effectués par les consommateurs potentiels qui souhaitent voyager de Cologne à Berlin, en termes de mode et de service proposés. Trois modes de transport (rail, air, route) sont présents sur le lien, avec une diversification sur le segment du trafic aérien : trois compagnies à bas coûts ainsi que la firme en place Lufthansa. Les alternatives sont caractérisées par un paramètre de qualité et un prix.

Nous supposons aussi l'existence d'une alternative dite alternative extérieure : au lieu de choisir un des services offerts pour voyager entre Cologne et Berlin, les consommateurs peuvent décider de ne pas effectuer le voyage. Ainsi la dimension totale du marché est définie comme le nombre de consommateurs qui seraient potentiellement intéressés par ce voyage.

Les voyageurs éventuels pour un aller simple de Cologne à Berlin sont hétérogènes. Nous considérons deux catégories principales dans la population des passagers : les voyageurs d'affaires et de loisirs. En conséquence, nous analysons deux marchés séparés qui sont traités de manière symétrique.

Supposons que chaque consommateur fasse son choix de manière séquentielle : il décide du mode de transport puis du service pour voyager². Un choix, c'est-à-dire une alternative, un produit, ou encore un service de transport, est alors une combinaison entre un mode de transport (par exemple, air) et un service fourni par un opérateur de transport (par exemple, Lufthansa). Notons que pour la route, l'opérateur de transport est le conducteur de la voiture. Il y a J alternatives classées en $G + 1$ groupes. Ici nous avons 4 groupes $g = 0, 1, \dots, G$ où le groupe 0 correspond à celui de l'alternative extérieure, les autres groupes correspondent à ceux des trois modes, c'est-à-dire rail, air, route.

2. Intuitivement, le choix du mode entraîne des décisions de plus long terme, car il implique l'achat d'un équipement (le véhicule particulier par exemple) ou l'achat d'un abonnement (au train, à l'avion).

L'utilité du consommateur i dépend de la qualité ψ_j et des prix p_j de chaque alternative, i.e. de chaque service de transport. La fonction d'utilité associée aux alternatives j ($j = 1, 2, \dots, J$) est la suivante :

$$U_{ij} = V_j + \varepsilon_{ij}, \quad (1)$$

où V_j représente le niveau d'utilité moyen commun à tous les passagers (partie déterministe) et ε_{ij} (partie aléatoire) correspond à la déviation du consommateur i du niveau d'utilité commun en ce qui concerne le produit j , c'est-à-dire que l'aléa ε_{ij} correspond à l'appréciation inconnue (pour l'analyste) du consommateur i pour le produit j . Le niveau d'utilité moyen peut se décomposer de la manière suivante :

$$V_j = \psi_j - hp_j, \quad (2)$$

où h représente la sensibilité de l'utilité par rapport au prix, ou encore l'utilité marginale du revenu³.

Les préférences aléatoires sont décrites selon un modèle de type « logit emboîté ». Dans ce contexte, les produits au sein d'un même groupe sont des substituts plus proches que des produits de groupes différents. Afin de permettre cette dépendance et cette corrélation entre les utilités des alternatives appartenant à un même groupe, la partie aléatoire peut être spécifiée comme la somme pondérée de variables inobservables qui représentent le goût du consommateur i pour une alternative appartenant au groupe g :

$$\varepsilon_{ij} = \sigma v_{ig} + (1 - \sigma) v_{ij}, \quad \forall i = 1, \dots, N \quad (3)$$

Le paramètre σ est appelé coefficient de corrélation intragroupe et prend des valeurs comprises entre 0 et 1. Il donne une mesure du degré de corrélation entre les alternatives, ici des services de transport, appartenant à un même groupe, ici un mode de transport : plus σ est élevé, plus la corrélation entre alternatives d'un même groupe est élevée. Quand il est égal à 1, les préférences des consommateurs entre alternatives sont parfaitement corrélées et les alternatives sont alors des substituts parfaits ; quand il est égal à 0, les préférences ne sont pas corrélées et les consommateurs mettent toutes les alternatives sur le même plan. Les parties aléatoires v_{ig} et v_{ij} sont supposées être distribuées selon la distribution de Gumbel (ou des valeurs extrêmes). De ce fait il en est de même de ε_{ij} .

3. Cette fonction d'utilité indirecte n'est pas seulement commode pour les calculs. Elle permet de dériver une demande cohérente avec notre propos. Il est aisé de voir que la demande marshallienne associée à cette fonction d'utilité est exactement égale à un. Or nous traitons ici du problème de choisir d'effectuer ou de ne pas effectuer un déplacement.

Le consommateur choisit l'alternative j appartenant au groupe g qui maximise son utilité et qui satisfait :

$$U_{ij} \geq U_{ij'}, \quad \forall j' \neq j \quad (4)$$

Nous pouvons alors calculer la probabilité de choisir l'alternative j à partir des probabilités de choisir le groupe g et de choisir l'alternative j conditionnelle au choix du groupe g . En appliquant la méthodologie proposée par Berry (1994), qui consiste en l'expression du niveau d'utilité moyen comme fonction des parts de marché observées qui sont des mesures des probabilités de choix, nous dérivons la spécification des demandes associées à chaque alternative :

$$\ln(s_j) - \ln(s_0) = \psi_j - hp_j + \sigma \ln(s_{j/g}), \quad (5)$$

où les parts de marché s_j and $s_{j/g}$ sont définies comme $s_j = \frac{q_j}{N}$, et $s_{j/g} = \frac{q_j}{N_g}$, $j = 0, 1 \dots J$, respectivement, où de plus q_j est la quantité de

produit j sur le marché (affaires ou loisirs), N est la taille de ce marché et N_g est la taille du segment de marché g . Nous pouvons dériver de cette spécification les expressions des élasticités simples et croisées de la demande de transport par rapport aux prix. L'élasticité-prix s'écrit :

$$\eta_j = \frac{dq_j}{dp_j} \times \frac{p_j}{q_j} = hp_j \left(s_j - \frac{1}{1-\sigma} + \frac{\sigma}{1-\sigma} s_{j/g} \right) \quad \forall j \in g \quad (6)$$

Les élasticités croisées pour des produits du même groupe ou pour des produits de groupes différents s'écrivent :

$$\eta_{j,k} = \frac{dq_j}{dp_k} \times \frac{p_k}{q_j} = hp_k s_k \left(\frac{\sigma}{1-\sigma} \times \frac{s_{k/g}}{s_k} + 1 \right) \quad \text{if } j \neq k \quad j, k \in g \quad (7)$$

$$\eta_{j,k} = \frac{dq_j}{dp_k} \times \frac{p_k}{q_j} = hp_k s_k \quad \text{if } j \neq k \quad k \notin g, j \in g$$

2.2. L'offre

Dans notre spécification, chaque entreprise fournit un produit, c'est-à-dire un service de transport sur le lien « origine-destination ». Premièrement, nous considérons le service voiture comme proposé par

une entreprise en concurrence avec les autres alternatives⁴. Deuxièmement, pour tous les autres services de transport offerts sur la liaison en question, les entreprises choisissent les prix de transport qui maximisent leur profit, sachant que leurs concurrents font de même. Ils entrent dans une concurrence à la Bertrand⁵. Comme dans Ivaldi et Verboven (2005), le résultat de cet équilibre de Nash est défini par un ensemble de J conditions nécessaires de premier ordre :

$$p_j = c_j + \frac{1-\sigma}{h(1-\sigma s_{j/g} - (1-\sigma)s_j)} \quad (8)$$

Le prix du produit j est égal au coût marginal du produit j plus une marge.

La solution qui détermine l'équilibre est obtenue en résolvant le système d'équations (5) et (8). Il est ensuite possible d'établir une mesure du surplus du consommateur, noté CS. Celle-ci correspond à la valeur monétaire espérée du maximum des utilités et ne prend pas en compte les effets externes. Elle s'écrit :

$$CS = \frac{1}{h} \ln \left[1 + \sum_{g=1}^G \left(\sum_{i \in g} \exp \left(\frac{V_i}{1-\sigma} \right) \right)^{1-\sigma} \right]. \quad (9)$$

2.3. Calibration

Dans la mesure où l'ensemble des informations nécessaires à l'estimation statistique de notre modèle n'est pas disponible, nous proposons de le calibrer à l'aide de données sur les prix, parts de marché, caractéristiques des différents services et quelques coûts marginaux. Nous devons aussi définir la taille du marché qui est donné par la quantité potentielle de déplacements, autrement dit qui prend en compte les décisions de ne pas se déplacer. La calibration de cette taille est délicate et est discutée en détails dans Ivaldi et Vibes (2007).

Notre méthodologie consiste en la résolution d'un système d'équations issues de notre modèle nous permettant de retrouver les paramè-

4. Nous avons alternativement fait l'hypothèse que cette entreprise établit le prix de manière à couvrir le coût moyen. Quand nous estimons ce modèle, nous n'obtenons pas de différence significative avec la situation d'une entreprise qui maximise son profit.

5. L'hypothèse de concurrence à la Bertrand paraît adaptée à l'exemple Cologne-Berlin que nous considérons ici car l'offre de transport y est variée et importante. Toutefois il serait intéressant de tester l'hypothèse de concurrence à la Cournot, ce qui serait possible si nous disposions de plus de données.

tres décrivant l'équilibre, tels que h , σ , ψ_j , ou encore les coûts marginaux manquants.

Très schématiquement, la procédure consiste en trois étapes. Tout d'abord, il s'agit de retrouver les valeurs de h et σ à l'aide des expressions des élasticités dérivées du modèle logit emboîté (Équation 6) et des parts de marché que nous observons. Les valeurs de ces paramètres, ainsi que celles des parts de marché, sont alors introduites dans les équations de demande (Équation 5), que nous inversons afin de retrouver les valeurs des ψ_j pour chaque alternative. Enfin, connaissant les valeurs de h , σ , des prix et des parts de marché, nous utilisons les équations de prix (Équation 8) pour retrouver les coûts marginaux. Techniquement, ne possédant pas de mesure des élasticités pour ce marché très spécifique, et connaissant certaines valeurs des coûts marginaux, nous générons aléatoirement un vecteur d'élasticités et le modèle, c'est-à-dire les équations de demande, d'élasticités et de prix simultanément, est calibré sur les valeurs de coûts marginaux que nous connaissons. Autrement dit, le vecteur d'élasticité retenu est celui qui nous permet de retrouver ces coûts marginaux *in fine*. Les élasticités sont tirées dans des lois normales en utilisant comme moyenne et variance des valeurs cohérentes avec les élasticités du marché loisir. Celles-ci peuvent être complètement déterminées puisque, pour ce marché, nous disposons d'informations supplémentaires sous la forme de coûts marginaux.⁶ Nous obtenons ainsi les valeurs de h et σ associés au vecteur d'élasticités retenu, puis ψ_j , et enfin les coûts marginaux manquants.

L'indice de qualité associé au produit j , ψ_j , peut ensuite être exprimé comme la somme pondérée de différentes composantes de cette qualité (essentiellement vitesse, fréquence et capacité). Comme nous disposons d'un nombre suffisant de produits et donc de valeurs de l'indice de qualité et puisque nous observons les valeurs des caractéristiques des différents services de transport, nous pouvons, par inversion d'un système linéaire, calibrer les poids de ces caractéristiques dans les indices de qualité des alternatives⁷. Nous pouvons donc mesurer l'impact de chaque caractéristique (fréquence, capacité, vitesse) sur l'indice de qualité de chaque service de transport. Comme une nouvelle infrastructure est associée à un vecteur particulier des caractéris-

6. En réalité la procédure pour estimer ces élasticités fait appel à un argument bayésien développé dans notre article principal. Voir Ivaldi et Vibes (2006). On le voit ici, la calibration résulte d'un jeu subtil entre informations sur les coûts marginaux et sur les élasticités avec le modèle structurel comme arbitre.

7. Nous tenons compte également des caractéristiques de transport non observées en introduisant dans ces équations des constantes relatives aux modes de transports.

tiques, nous avons donc la possibilité d'évaluer son indice de qualité associé.

Une fois tous ces paramètres calibrés, il est donc possible de simuler une nouvelle configuration du marché après un changement comme la mise en place d'une nouvelle infrastructure. Celle-ci se traduit, comme on vient de le dire, par un nouvel indice de qualité. Le modèle de concurrence est alors utilisé pour calculer les nouveaux prix et les nouvelles parts de marché de tous les services de transport, y compris celui associé à la nouvelle infrastructure de transport. À ce nouvel équilibre correspond une mesure du bien-être social, par exemple le surplus du consommateur⁸. Ainsi il est possible d'associer à chaque équilibre résultant de la mise en œuvre d'une nouvelle infrastructure une mesure du bien-être social et il est donc possible de classer ces projets d'infrastructure selon cette mesure de bien-être social⁹.

3. SIMULATION

Une fois les paramètres du modèle calibrés, nous sommes capables d'évaluer les effets de changements structurels sur les prix et parts de marché des concurrents, ainsi que sur le surplus du consommateur. En particulier nous simulons ici l'impact d'une nouvelle installation ferroviaire entre Cologne et Berlin destinée au passage d'un train à grande vitesse de type ICE dont la spécificité serait de n'effectuer aucun arrêt sur le parcours et donc d'opérer à une vitesse moyenne de 220 kilomètres par heure. Le temps de trajet est alors significativement réduit et ce train permet d'effectuer un trajet centre ville à centre ville en un temps comparable à celui d'un vol et des connexions nécessaires depuis l'aéroport. Ce nouveau train express présente les mêmes fréquences, capacités et coûts marginaux que le service en place. Les résultats de cette simulation sont présentés dans le Tableau 2, et confirment notre intuition. Nous constatons que cette alternative, attractive pour les passagers d'affaires car plus rapide et moins chère (même si ce critère est secondaire pour eux), l'est encore plus pour les passagers de loisirs, plus sensibles au niveau de qualité du mode train, pour lequel ils sont prêts à payer un peu plus cher maintenant que pour les autres modes.

8. On pourrait aussi choisir le bien-être total, somme du surplus du consommateur et des profits des entreprises de transport.

9. Une description détaillée du modèle ainsi que de la méthode de calibration est donnée dans Ivaldi et Vibes (2006).

Les paramètres estimés à l'équilibre ainsi que les résultats de la simulation nous permettent de décrire les comportements des entreprises et des consommateurs. Dans un premier temps nous présentons les grandeurs d'équilibre pertinentes qui ont été calibrées. Conformément à nos attentes, l'utilité marginale du revenu des passagers de loisirs ($h \approx 0.032$) est supérieure à celle des passagers d'affaires ($h \approx 0.016$) : ces derniers sont moins sensibles à des variations de prix des billets. Ensuite, le coefficient de corrélation intragroupe prend une valeur de 0.26 sur le marché de loisirs et de 0.21 sur le marché des passagers d'affaires. Ce paramètre, différent de 0 est quand même plus proche de 0 que de 1, indique qu'une corrélation intramodale est bien présente, mais que le degré de substituabilité entre alternatives d'un même groupe n'est pas si élevé, laissant place à une concurrence certaine entre des options de transport n'appartenant pas au même mode, c'est-à-dire une concurrence intermodale. Enfin, les valeurs d'élasticités calculées à l'équilibre sont plus élevées que celles que l'on peut trouver dans la littérature (Oum, Waters et Young, 1990) car le degré de concurrence est particulièrement élevé sur la liaison que nous étudions.

Tableau 2

Introduction d'une nouvelle infrastructure

			AFFAIRES		LOISIRS	
Part de l'alternative extérieure			15 %		30 %	
Variations : valeurs, %			Valeur	Variation %	Valeur	Variation %
Prix	Rail	DB	70.4	-21.8 %	42.9	-28.5 %
		Train Express	96.2		62.3	
	Air	DBA	166.7	-1.4 %	50.5	-1.5 %
		HLX	166.9	-1.2 %	46.0	-1.5 %
		GW	167.8	-0.7 %	45.7	-0.9 %
		LH	237.8	-0.9 %	52.6	-1.3 %
	Route	Voiture	106.6	-3.12 %	77.7	-2.9 %

		AFFAIRES		LOISIRS		
Part de l'alternative extérieure		15 %		30 %		
Variations : valeurs, %		Valeur	Variation %	Valeur	Variation %	
Parts de Marché %	Rail	DB	8.8	-33.9 %	21.7	-39.6 %
		Train Express	28.2		48.4	
	Air	DBA	11.9	-26.8 %	3.0	-53.2 %
		HLX	11.0	-27.1 %	2.7	-53.4 %
		GW	6.4	-28.5 %	1.6	-53.9 %
		LH	11.5	-26.9 %	2.8	-53.3 %
	Route	Voiture	11.5	-25.6 %	6.0	-50.8 %
Alternative extérieure		10.6	-29.4 %	13.7	-54.4 %	
Surplus du Consommateur		145.6	18.4 %	61.1	65 %	

Nous pouvons alors expliquer avec plus de précision les résultats de notre simulation. Nous observons tout d'abord que, tous marchés confondus, les concurrents réagissent en baissant leurs prix, mais de façon peu marquée pour les compagnies aériennes qui ne peuvent réduire leurs marges davantage et se sentent moins menacées que l'opérateur ferroviaire en place. Ensuite, cette nouvelle option de voyage attire beaucoup de passagers auparavant usagers des autres alternatives de transport, ainsi qu'une partie des consommateurs qui préféreraient ne pas effectuer le trajet, comme l'indique la diminution de la part de marché de l'alternative extérieure. En effet nous constatons une réduction des parts de marché des services en place, et particulièrement prononcée sur le marché des passagers loisirs. Les passagers d'affaires ayant une préférence marquée pour la qualité du transport aérien (Ivaldi et Vibes, 2007) et ne possédant qu'une faible sensibilité à une variation des prix présentent un engouement moins marqué, bien que significatif, que les passagers de loisirs pour ce train. Les voyageurs de loisirs observent une nette amélioration du niveau de qualité de leur mode favori, ce qui suffit, malgré le prix, à attirer les quelques passagers qui voyageaient en avion pour des raisons de tarifs.

Par ailleurs, la part de marché de l'alternative extérieure ayant diminué sur les deux marchés, nous constatons que cette nouvelle infrastructure aurait pour effet non seulement de répartir le trafic entre les modes de manière différente, mais aussi d'induire du trafic sur cette liaison.

Enfin, nous observons une augmentation du surplus du consommateur sur les deux marchés suite à la mise en place de cette nouvelle infrastructure. Celle-ci représente une alternative de transport d'une part plus appréciée des voyageurs d'affaires et de loisirs, et d'autre part plus attractive pour une partie de la population auparavant non décidée à effectuer le trajet.

Nous pouvons aussi calculer le profit généré par les opérateurs et notamment l'opérateur de la nouvelle infrastructure. De ce fait on pourrait évaluer la part des coûts fixes qui serait couverte, à condition de tenir compte du fait que le modèle est statique et défini pour une année donnée.

4. CONCLUSION

L'évaluation économique d'un projet tel que la construction d'une voie de chemin de fer supplémentaire sur un trajet particulier passe nécessairement par une analyse de la concurrence sur ce marché. Cette infrastructure constitue un nouveau concurrent parmi les opérateurs de transport et il est indispensable de pouvoir mesurer les réactions de chacun suite à ce changement structurel. Ces transformations liées à la qualité du service font l'objet d'une simulation qui nous permet de déterminer les variations au niveau des prix et des parts de marchés des différentes entreprises sur le marché. Nous sommes ensuite capables de calculer l'impact sur le bien-être social d'un tel aménagement, ou plus particulièrement la variation de surplus du consommateur qui lui est associée, sans tenir compte des implications externes. Cette méthodologie peut également être utilisée afin de dériver les caractéristiques d'un service qui permettraient d'atteindre un objectif de surplus souhaité et prédéterminé.

Dans l'exemple de la liaison entre Cologne et Berlin, notre méthodologie de calibration puis de simulation nous permet de constater que cette nouvelle infrastructure constituerait un concurrent sérieux sur le marché puisqu'elle permettrait un trajet rapide et se positionnerait en concurrence directe avec les compagnies aériennes. Cet aménagement a également pour effet d'induire du trafic et d'augmenter le surplus du consommateur, en particulier sur le marché des passagers de loisirs.

Références

Antes J., Friebel G., Niffka M. et Rompf D., *Entry of Low-Cost Airlines in Germany. Some Lessons for the Economics of Railroads and*

- Inter-Modal Competition*, mimeo, Deutsche Bahn et Université de Toulouse, 2004.
- Berry S. T., « Estimating Discrete-Choice Models of Product Differentiation », *The RAND Journal of Economics*, vol. 25, Issue 2, 1994, p. 242-262.
- European Commission, *European Union Energy and Transport in Figures*, Directorate General for Energy and Transport, en coopération avec Eurostat, Bruxelles, 2003.
- Glass A., « An Assessment of the Desirability of On-Track Competition : The Ipswich-London route », Institute of Transport Studies, Université de Leeds, 2003.
- Holvad T., Huang B. et Preston J., « Review of Introduction of Competition in railways in Europe », Transport Studies Unit, Université d'Oxford, 2003.
- IDEI, « The Economics of Passenger Rail Transport : A Survey », 2003. http://www.idei.fr/doc/wp/2003/rapport_db_1.pdf.
- IDEI, « Entry in the Passenger Rail Industry : A Theoretical Investigation », 2003. http://www.idei.fr/doc/wp/2003/rapport2_db_2.pdf.
- Ivaldi M. et Verboven F. (2004), « Quantifying the Effects from Horizontal Mergers in European Competition Policy », *International Journal of Industrial Organization*, à venir.
- Ivaldi M. et Vibes C., *Price Competition in the Intercity Passenger Transport Market: a Simulation Model*, Working Paper, Institut d'Économie Industrielle, 2007.
- Johnson D. et Whelan G., « Modelling the Impact of Alternative Fare Structures », Institute of Transport Studies, University de Leeds, 2003.
- McFadden D., « The Measurement of Urban Travel Demand », *Journal of Public Economics*, vol. 3, 1974, p. 303-325.
- Nash C. et Rivera-Trujillo C., « Rail Regulatory Reform in Europe, Principles and Practice », Institute of Transport Studies, University de Leeds, 2004.
- Oum T. H., Waters W.G. et Jon Say Yong, « A Survey of Recent Estimates of Price Elasticities of Demand for Transport », Infrastructure and Urban Development Department, the Worldbank, 1990.
- Preston J., Wardman M. et Whelan G., « An Analysis of the Potential for On-Track Competition in the British Passenger Rail Industry », *Journal of Transport Economics and Policy*, vol. 33, part 1, 1999, p. 77-94.

CHAPITRE 3

ACTUALISATION : PRISE EN COMPTE DU TEMPS DANS UN ENVIRONNEMENT RISQUÉ

Christian Gollier¹

En janvier 2005, le Commissariat Général du Plan remettait un rapport préconisant la baisse du taux (réel) d'actualisation de 8 à 4 %, et même 2 % pour des horizons temporels supérieurs à 30 ans. À cette occasion, une critique récurrente s'éleva pour prédire une augmentation massive des projets d'investissements publics franchissant le test de la valeur actualisée nette (VAN) positive, en particulier pour les projets aux bénéfices s'étalant sur le très long terme (lutte contre l'effet de serre, biodiversité, gestion des déchets). Or, l'État français ne prévoit ni d'augmenter la pression fiscale, ni de creuser son déficit pour faire face au financement de ces investissements. Le risque serait donc de voir le politique choisir arbitrairement ceux de ces projets qu'il mettrait en œuvre ; un échec évident du calcul économique.

Le choix d'un taux de 8 % en 1985 avait été justifié en partie par la nécessité d'intégrer une prime de risque dans le calcul économique public. Évidemment, cette méthode consistant à réduire les VAN de tous les projets en augmentant uniformément le taux d'actualisation est critiquable à plus d'un point de vue. Elle ne peut être justifiée que si tous les projets d'investissement publics en concurrence de financement ont des risques comparables autant en intensité qu'en corrélation avec le risque macroéconomique, et qu'en étalement dans le temps. Dans le cas contraire, le choix d'un taux uniforme de 8 % pénalise injustement les projets les moins risqués, ou ceux dont les incertitudes sont les plus éloignées dans le temps. Ce sont les raisons pour lesquelles le rapport de janvier 2005 rejette cette méthode et propose d'intégrer le risque dans l'évaluation de cash-flows équivalent

1. Université de Toulouse (IDEI et LERNA)

certains plutôt que dans le choix du taux d'actualisation, c'est-à-dire qu'il propose d'intégrer le risque dans le numérateur plutôt que dans le dénominateur de la VAN.

Cette méthode a plusieurs avantages. Avant tout, elle rend au taux d'actualisation sa vraie fonction, celle d'un taux de change entre consommation future certaine et consommation immédiate. En conséquence, elle réaffirme la règle indispensable de l'unicité du taux d'actualisation.

Mais si le risque n'est pas imputé au taux d'actualisation, comment celui-ci doit-il être intégré au calcul ? Force est de constater qu'une faiblesse du rapport est de ne pas décrire avec précision la manière dont ceci doit être fait. Pourtant, tout indique que cette question est critique dans le processus d'évaluation. Ainsi, le capital sans risque a été rémunéré à un taux réel proche de zéro pourcent durant le vingtième siècle, alors que, globalement, le capital risqué a été rémunéré à un taux bien supérieur durant la même période. Rappelons en effet que le rendement d'un panier représentatif des actions françaises a été rémunéré à un taux réel annuel moyen de 4 % durant le siècle écoulé, et que ce rendement a atteint 6,9 % outre Atlantique (Dimson *et al*, 2000). Si on rajoute une telle prime de risque au nouveau taux d'actualisation sans risque du Plan, on retrouve globalement le taux de 8 %. Néanmoins, comme indiqué ci-dessus, ce calcul ne peut être qu'indicatif, pour le cas très spécifique d'un projet d'investissement public dont le risque dupliquerait le risque du marché des actions. Ceci nous donne malgré tout une information importante : si le risque est convenablement intégré au calcul économique dans le numérateur, il ne faut pas s'attendre *a priori* à une augmentation massive des projets d'investissement franchissant le test de la VAN positive. Seuls les projets à *faible risque* dont le taux de rendement interne est compris entre 4 % et 8 % franchissent maintenant ce test, alors qu'ils ne les franchissaient pas avant la réforme.

Ceci pose à nouveau le problème de l'évaluation du risque dans le calcul économique. Les objectifs de ce chapitre consistent à rappeler les quelques règles simples de la théorie moderne de la finance qui régissent cette évaluation.

1. JUSTIFICATION DE LA VAN EN INCERTITUDE

On considère une économie « à la Lucas (1978) » en temps discret, avec un agent représentatif éternel. L'existence d'un agent représentatif dans une économie hétérogène est démontrée par exemple par Constantinides (1982). Gollier (2001a) montre comment prendre en

compte l'inégalité des richesses dans le calcul économique. L'espérance de vie infinie de l'agent représentatif signifie implicitement que les consommateurs intègrent les préférences de leurs descendants comme si c'étaient les leurs. Cet agent représente les générations présentes et futures.

Dans ce modèle, je ne cherche pas à expliquer la croissance économique. Elle prend la forme d'un vecteur exogène de variables aléatoires (c_0, c_1, c_2, \dots) de dimension infinie, où c_t représente la consommation par habitant à la date t . On suppose connue la distribution de probabilité de ce vecteur, conditionnellement à toutes les informations disponibles aujourd'hui ($t = 0$). Cette distribution caractérise le risque macroéconomique et son évolution dans le temps. Le bien-être intertemporel de l'agent représentatif est mesuré par la valeur actuelle de son flux d'espérance d'utilité future :

$$V_0 = \sum_{t=0}^{\infty} e^{-\delta t} Eu(c_t).$$

Le paramètre de préférence pure pour le présent est noté δ . La fonction d'utilité u est supposée croissante et concave.

On considère un projet d'investissement caractérisé par un vecteur de cash-flows aléatoires (X_0, X_1, X_2, \dots) , où X_t est le bénéfice net des coûts à la date t généré par l'investissement. On suppose connue la distribution de probabilité de ces cash-flows, ainsi que leur corrélation avec le risque macroéconomique.

Les bénéfices et les coûts du projet sont équitablement partagés par les consommateurs. Soit ε la part des cash-flows nets perçus par chaque consommateur. Si le projet d'investissement est réalisé, l'agent représentatif obtiendra un niveau de bien-être égal à

$$V_1 = \sum_{t=0}^{\infty} e^{-\delta t} Eu(c_t + \varepsilon X_t).$$

Évidemment, le projet est socialement désirable si l'agent représentatif voit son bien-être intertemporel augmenté grâce à la réalisation de l'investissement, c'est-à-dire si V_1 est supérieur à V_0 . Comme on suppose une population de grande taille, la part ε est très petite. En conséquence, V_1 est supérieur à V_0 si

$$\frac{\partial}{\partial \varepsilon} \left[\sum_{t=0}^{\infty} e^{-\delta t} Eu(c_t + \varepsilon X_t) \right]_{\varepsilon=0} > 0,$$

ou encore si

$$\sum_{t=0} e^{-\delta t} EX_t u'(c_t) > 0.$$

Comme X_0 et c_0 sont certains, on peut réécrire cette condition nécessaire et suffisante comme

$$VAN = X_0 + \sum_{t=1} e^{-r_t t} B_t > 0, \quad (1)$$

avec

$$e^{-r_t t} = e^{-\delta t} \frac{Eu'(c_t)}{u'(c_0)} \quad (2)$$

et

$$B_t = \frac{EX_t u'(c_t)}{Eu'(c_t)}. \quad (3)$$

On voit que le projet d'investissement est socialement désirable si sa VAN exprimée par (1) est positive. Le taux d'actualisation r_t défini par (2) est indépendant du projet considéré. Il est donc unique et universel, mais peut varier en fonction de la maturité t du cash-flow. Par contre, le cash-flow équivalent certain B_t dépend à la fois des incertitudes sur c_t et sur X_t . Ainsi, on voit que nous avons effectivement séparé les problématiques de choix de taux d'actualisation et de prise en compte du risque du projet. La formule (1) nous montre donc la procédure à suivre pour mener à bien l'évaluation des projets d'investissement.

Proposition : Lorsque les cash-flows (X_0, X_1, \dots) d'un projet d'investissement ainsi que les anticipations de croissance économique (c_0, c_1, \dots) sont incertains, ce projet est socialement efficace si sa valeur actualisée nette est positive. Cette VAN est évaluée en deux étapes : Pour chaque maturité t ,

1) On calcule le bénéfice équivalent certain B_t à partir de la formule (3) ;

2) On actualise ce bénéfice au taux r_t défini par la formule (2).

Enfin, la VAN est obtenue en sommant ces bénéfices équivalents certains actualisés.

2. CHOIX DU TAUX D'ACTUALISATION

Le taux d'actualisation socialement efficace dérivé de la formule (2) peut se réécrire comme

$$r_t = \delta - \frac{1}{t} \ln \left(\frac{Eu'(c_t)}{u'(c_0)} \right). \quad (4)$$

Cette formule se retrouve dans tous les livres de référence en finance, tel Cochrane (2001). Elle fonde la théorie moderne de la structure par terme des taux d'intérêt initiée par Vasicek (1977) et Cox, Ingersoll et Ross (1985). Elle est généralement traduite dans un cas très particulier, qui combine deux hypothèses supplémentaires. La première suppose que l'utilité marginale est une fonction puissance : $u'(c) = c^{-\gamma}$, où γ est l'indice relatif d'aversion au risque. La seconde hypothèse caractérise le processus de croissance exogène. Supposons pour l'instant que le logarithme de la consommation suive un mouvement Brownien de tendance connue μ et de volatilité σ . Ceci implique que $\ln c_t$ est normalement distribué d'espérance $c_0 e^{\mu t}$ et de variance $\sigma^2 t$. En combinant ces deux spécifications, on obtient que

$$\frac{Eu'(c_t)}{u'(c_0)} = E \exp(-\gamma(\log c_t - \log c_0)) = \exp(-\gamma(\mu t - 0.5\gamma\sigma^2 t)). \quad (5)$$

Dans la seconde égalité, j'ai utilisé la propriété bien connue selon laquelle l'approximation d'Arrow-Pratt est exacte dans le cas de risque normalement distribué et fonction exponentielle². En combinant les équations (4) et (5), on obtient que

$$r_t = \delta + \gamma\mu - \frac{1}{2}\gamma^2\sigma^2. \quad (6)$$

Dans Gollier (2002a) et Gollier (2005), je donne une intuition à cette formule qui permet de déduire le taux d'actualisation efficace à partir des anticipations de croissance de l'économie et des préférences des agents. En bref, le taux d'actualisation socialement efficace a trois composantes, comme indiqué dans le membre de droite de l'équation (6). La première composante est le taux de préférence

2. Voir par exemple (Gollier, 2001b, page 57).

pure pour le présent, δ . La seconde composante est un effet richesse. Parce que l'agent représentatif anticipe une hausse de sa consommation ($\mu > 0$), un euro supplémentaire à l'avenir a un effet sur son utilité plus faible qu'un euro supplémentaire immédiat, puisque son utilité marginale est décroissante avec la consommation ($\gamma > 0$). Cet effet richesse est positif sur le taux d'actualisation, et incite à réduire nos efforts pour améliorer notre avenir. Pourquoi faire des sacrifices pour un avenir de toute façon plus riche que le présent ? Cet effet richesse est d'autant plus élevé que les anticipations de croissance sont optimistes, et que l'utilité marginale décroît rapidement. La troisième composante décrit un effet de précaution. Comme l'utilité marginale est convexe avec la consommation, une augmentation de l'incertitude sur la consommation future augmente la valeur d'un euro supplémentaire en t , telle que mesurée par l'espérance de l'utilité marginale. L'incertitude réduit donc le taux d'actualisation. Exactement comme les ménages augmentent leur épargne lorsque leurs revenus futurs deviennent plus aléatoires, au niveau collectif, il est efficace d'accroître les investissements en situation d'incertitude macroéconomique. Cet effet précaution est croissant avec la volatilité σ de la croissance du PIB par habitant.

Calibrons l'équation (6). Durant le 20^e siècle, la croissance réelle du PIB par habitant a été en moyenne de $\mu = 2$ % par an,³ tandis que sa volatilité est estimée à $\sigma = 2$ %. Pour des raisons d'éthique intergénérationnelle, fixons $\delta = 0$. Finalement, quelle valeur de γ retenir pour cette calibration ? Ce choix revêt effectivement une importance considérable. Les études autant expérimentales qu'économétriques sont très nombreuses, mais apportent une réponse contrastée à cette question. Je ne chercherai pas ici à synthétiser ces travaux. Je propose donc de faire une analyse par introspection. Supposons que votre richesse soit sujette à un risque de gain ou de perte de α %. Quel pourcentage π de votre richesse êtes-vous prêt à payer pour éliminer ce risque ? Le Tableau 1 lie votre réponse π à cette question à votre indice relatif γ de votre aversion au risque. Au vu de ce tableau, il semble raisonnable de choisir une valeur de γ comprise entre 1 et 4. Suivant Hall (1988), je choisis $\gamma = 2$.

3. Il est nécessaire de rappeler ici que μ désigne l'espérance de croissance du logarithme de la consommation, qui est différent de l'espérance $\hat{\mu}$ du taux de croissance de la consommation. En fait, par le Lemme d'Ito, on a que $\hat{\mu} = \mu + 0.5\sigma^2$. Comme σ est petit, la différence est faible.

Tableau 1
Prime de risque π et aversion relative au risque γ

	$\alpha = 10 \%$	$\alpha = 30 \%$
$\gamma = 0,5$	$\pi = 0,3 \%$	$\pi = 2,3 \%$
$\gamma = 1,0$	$\pi = 0,5 \%$	$\pi = 4,6 \%$
$\gamma = 4,0$	$\pi = 2,0 \%$	$\pi = 16,0 \%$
$\gamma = 10$	$\pi = 4,4 \%$	$\pi = 24,4 \%$
$\gamma = 40$	$\pi = 8,4 \%$	$\pi = 28,7 \%$

Avec de telles anticipations de croissance et de telles préférences intertemporelles, il est socialement efficace de choisir un taux d'actualisation de 3,92 %. L'effet richesse $\gamma\mu$ à lui seul conduit à sélectionner un taux d'actualisation de 4 %, tandis que l'effet précaution $0.5\gamma^2\sigma^2$ réduit ce taux de seulement 0,08 %. À court terme, l'incertitude est si faible qu'elle n'affecte pratiquement pas le taux d'actualisation socialement efficace.

Penchons-nous maintenant sur une question cruciale pour la problématique du développement durable : faut-il choisir un taux d'actualisation plus faible pour actualiser des cash-flow plus éloignés dans le temps ? D'un point de vue théorique, rien n'interdit *a priori* que ce taux r_t décroisse avec t . Néanmoins, la formule (6) nous montre que, sous la spécification étudiée ci-dessus, le taux d'actualisation est indépendant de l'horizon temporel. Il faut comprendre que les effets richesse et de précaution jouent en sens opposés lorsqu'on modifie la maturité étudiée. Plus on s'éloigne dans le temps, plus l'espérance de c_t est grande, ce qui doit nous inciter à choisir un taux d'actualisation croissant avec l'horizon temporel. Par contre, plus on s'éloigne dans le temps, plus l'incertitude sur c_t est importante, ce qui doit nous inciter à choisir un taux d'actualisation décroissant avec l'horizon temporel. Dans le cas où $u'(c) = c^{-\gamma}$ et où la croissance du logarithme de la consommation suit un Brownien constant, l'équation (6) montre que ces deux effets s'annihilent l'un l'autre.

Gollier (2002b) relâche l'hypothèse $u'(c) = c^{-\gamma}$, tandis que Weitzman (2004) et Gollier (2004) relâchent l'hypothèse de mouvement Brownien constant. Deux justifications émergent qui peuvent justifier un taux d'actualisation décroissant. La première porte sur l'hypothèse d'un décrochage du trend de croissance μ d'un niveau élevé durant les T premières périodes à un trend de croissance μ' plus faible au-delà. Dans ce cas, les taux d'actualisation socialement efficace de court terme et de long terme correspondront à la formule (6) utilisée respectivement avec μ et μ' . La structure par terme du taux d'actuali-

sation est donc décroissante dans ce cas, exactement comme la « yield curve » sur les marchés financiers peut être inversée lorsque l'on anticipe un retournement conjoncturel en phase haute du cycle macroéconomique.

La deuxième justification d'un taux d'actualisation décroissant est basée sur l'existence d'une relation convexe entre la variance de c_t et t , alors qu'elle était linéaire dans le cas Brownien constant. Il s'agit donc d'une situation où le risque est relativement plus important à long terme qu'à court terme, ceci par rapport au cas Brownien constant. Dans une telle situation, l'effet précaution va dominer à long terme, ce qui nous incite à choisir un taux d'actualisation plus petit à long terme qu'à court terme. Une telle situation se présente par exemple lorsqu'il existe une incertitude sur le paramètre μ du trend de croissance. Il semble en effet irréaliste de supposer que l'économie croîtra pour toujours autour d'un trend de 2 %. Supposons alternativement que ce trend puisse prendre une valeur parmi n valeurs possibles μ_i , $i = 1, \dots, n$, respectivement avec probabilité p_1, \dots, p_n . Dans ce cas, on obtient aisément une généralisation de la formule (5) :

$$\frac{Eu'(c_t)}{u'(c_0)} = \sum_{i=1}^n p_i \exp(-\gamma(\mu_i t - 0.5\gamma\sigma^2 t))$$

En utilisant la formule (4), on obtient la proposition suivante.

Proposition (Gollier, 2004) : *Lorsqu'il existe une incertitude sur le trend de croissance μ du logarithme de la consommation par habitant, le taux d'actualisation r_t socialement efficace s'écrit comme suit :*

$$r_t = \delta - 0.5\gamma^2\sigma^2 - \frac{1}{t} \ln \left(\sum_{i=1}^n p_i \exp(-\gamma\mu_i t) \right). \quad (7)$$

Contrairement au cas Brownien constant décrit par la formule (6), le taux d'actualisation est ici une fonction de l'horizon temporel t . On peut vérifier que cette fonction est décroissante en t , avec un taux à très court terme égal à

$$r_{t \rightarrow 0} = \delta + \gamma \left(\sum_{i=1}^n p_i \mu_i \right) - \frac{1}{2} \gamma^2 \sigma^2,$$

et un taux à long terme qui tend vers

$$r_{t \rightarrow \infty} = \delta + \gamma \min_i \mu_i - \frac{1}{2} \gamma^2 \sigma^2.$$

Ainsi, si on pense qu'il y a autant de chance que le trend de croissance soit de 1 % ou 3 %, le taux d'actualisation socialement efficace est égal à 3,92 % à court terme, mais est égal à seulement 1,92 % à long terme. En fait, dans une telle configuration, il est facile de vérifier que l'utilisation de la formule (7) dans le calcul économique revient à calculer la VAN deux fois, une fois avec un taux constant de 5,92 % et une fois avec un taux constant de 1,92 %, et de prendre comme VAN effective la moyenne de ces deux valeurs actualisées. Mathématiquement, ceci revient au même que de calculer la VAN une seule fois, mais avec le taux d'actualisation décroissant (7). La méthode basée sur la moyenne des VAN est évidemment singulièrement plus facile à mettre en œuvre.

3. PRISE EN COMPTE DU RISQUE DU PROJET

Revenons à la formule (1) qui définit la VAN du projet, et intéressons-nous maintenant au risque du projet lui-même. On peut interpréter B_t dans la formule (1) comme le « bénéfice équivalent certain » à imputer au projet à la date t , et donc à actualiser au taux r_t . La formule (3) caractérise la manière dont ce bénéfice équivalent certain doit être calculé. Le cas le plus simple correspond à la situation où le risque du projet est indépendant du risque macroéconomique, c'est-à-dire lorsque X_t et c_t sont deux variables indépendantes. Dans ces circonstances, on a que $EX_t u'(c_t)$ est égal à $EX_t Eu'(c_t)$, ce qui implique par (3) que $B_t = EX_t$. Ceci nous donne le résultat de Arrow et Lind (1970), que l'on peut résumer de la façon suivante.

Proposition (Arrow et Lind, 1970) : Lorsque le risque du projet est non corrélé au risque macroéconomique, l'évaluation de ce projet doit se faire en neutralité au risque, c'est-à-dire que $B_t = EX_t$.

Le problème est plus délicat lorsque le risque du projet est corrélé au risque macroéconomique, ceci malgré l'hypothèse de grande taille de la population. En fait, on peut réécrire la formule (3) comme

$$B_t = EX_t + \text{cov} \left(X_t, \frac{u'(c_t)}{Eu'(c_t)} \right). \quad (8)$$

Il est utile de rappeler ici que l'utilité marginale de la consommation est décroissante. En conséquence, cette formule indique que le bénéfice équivalent certain B_t est plus petit que le bénéfice espéré EX_t , si les bénéfices du projet sont positivement corrélés avec la croissance économique, comme on peut le supposer dans la plupart des cas.

Proposition (MEDAF) : *Lorsque le risque du projet est positivement corrélé au risque macroéconomique, l'évaluation de ce projet doit intégrer une prime de risque qui réduit la valorisation du bénéfice futur à actualiser, ceci malgré la dissémination du risque du projet dans une population très large.*

On a l'habitude d'opérationnaliser cette théorie en utilisant une approximation de la formule (8). Soit C_t l'espérance de c_t . L'approximation de Taylor du premier degré de $u'(c_t)$ autour de C donne $u'(C_t) + (c_t - C_t)u''(C_t)$. Approximons par ailleurs $Eu'(c_t)$ par $u'(C_t)$. On peut donc approximer B_t par

$$B_t = EX_t - \gamma \frac{\text{cov}(X_t, c_t)}{Ec_t} \quad (9)$$

où γ est le coefficient d'aversion relatif pour le risque $\gamma = -C_t u''(C_t)/u'(C_t)$. Ainsi, la « prime de risque » est croissante en l'aversion au risque du consommateur représentatif et en la covariance entre le rendement du projet et le PIB/hbt, résultat classique du Modèle d'Évaluation Des Actifs Financiers (MEDAF).

Jusqu'à maintenant, j'ai supposé que la décision consistait soit à mettre en œuvre le projet d'investissement immédiatement, soit à l'abandonner définitivement. En réalité, dans la plupart des cas, il est possible de ne pas investir immédiatement, mais de conserver l'option d'investir ultérieurement. Or, reporter un investissement dans le temps, c'est rendre possible l'acquisition d'informations supplémentaires sur sa rentabilité sociale. Ainsi, même si un projet a une VAN positive, il peut être socialement efficace de reporter cet investissement dans l'attente de ces informations.

Pour illustrer, considérons un projet d'investissement de coût initial I irréversible supporté à la date de mise en œuvre de l'investissement, et qui génère un bénéfice unique R aléatoire à la date suivant cette mise en œuvre. Supposons néanmoins que cet aléa soit indépendant du risque macroéconomique, de manière à ce que nous puissions utiliser une évaluation neutre au risque. Finalement, on sait qu'en date $t = I$, une information s non corrélée avec le risque macroéconomique sera disponible, ce qui permettra au planificateur de réviser la distribution du bénéfice R . Comparons deux stratégies. La première stratégie consiste à investir immédiatement, ce qui génère une VAN égale à

$$H_0 = -I + e^{-rI} ER.$$

La deuxième stratégie consiste à reporter la décision d'investissement à la date $t = I$. Evidemment, on utilise à cette date l'information s disponible, et on investit que si

$$-I + e^{-r} E[R|s]$$

est positif. En conséquence, la stratégie de report génère en $t = 0$ une espérance de VAN égale à

$$H_1 = e^{-r} E \max[0, -I + e^{-r} E[R|s]] > 0.$$

On voit qu'il faut investir immédiatement non pas si la VAN espérée H_0 est positive, mais plutôt si H_0 est plus grand que H_1 . H_1 est la « valeur d'option » de report. Dans certaines applications, elle peut être très élevée. Son calcul nécessite souvent la mise en œuvre de techniques d'optimisation dynamique stochastique relativement complexes. Elles sont utilisées dans un certain nombre de grandes entreprises privées, en particulier dans le domaine de la production minière et de la pharmacie, mais leur utilisation reste embryonnaire dans le secteur public malgré leur importance cruciale dans certains domaines comme l'énergie, les transports ou les télécommunications. De très importants développements, à la fois théorique et empiriques, ont été réalisés dans ce domaine depuis les travaux pionniers d'Henry (1974) et Arrow et Fischer (1974). Les lecteurs intéressés pourront se reporter sur le livre récent de Smit et Trigeorgis (2004) pour de plus amples analyses. Devezeaux et Gollier (2001) ont développé une application dans le domaine de la valorisation de la réversibilité du site de stockage de déchets nucléaires.

4. UN POINT DE VUE PERSONNEL SUR LA PRATIQUE DE L'ACTUALISATION, LA PRESSION DES LOBBIES ET LA POLITIQUE D'INVESTISSEMENTS PUBLICS

Dans une économie française où l'État s'endette à un taux réel proche de zéro pourcent, et où les ménages acceptent de sacrifier une partie de leur consommation immédiate par l'épargne dans des livrets A de rendement réel lui aussi pratiquement nul, c'est une réforme de bon sens que de réduire le taux d'actualisation de 8 à 4 %. Il n'est en fait pas même nécessaire de faire les développements ci-dessus pour comprendre que la mise en œuvre d'un projet d'investissement public dont la rentabilité réelle certaine serait, disons, de 6 % par an va augmenter le bien-être du citoyen. En effet,

- En réalisant ce projet, l'État organise la réallocation des ressources rares de la société de projets à rentabilité réelle à 0 % vers des projets à rentabilité réelle à 6 %.

- Si les ménages acceptent de sacrifier une partie de leur consommation courante pour investir à taux nul, ils seront *a fortiori* prêts à faire de même pour une rentabilité de 6 %.

Si l'État décide l'abandon de ce projet, cela constitue donc au mieux un manque de clairvoyance et d'incompétence de l'État, une gestion de mauvais père de famille. C'est particulièrement vrai dans le cas de grands projets d'infrastructure où la perte sociale exprimée en euros plutôt qu'en pourcents est potentiellement considérable.

Pourtant, la baisse du taux d'actualisation en France a suscité une réaction unanime pour proclamer que seul un petit sous-ensemble de projets à rentabilité interne réelle certaine supérieure à 4 % pourra être réalisé. Ma première réaction fut de considérer cette attitude inadmissible de la part de hauts fonctionnaires de l'État dont la mission est de promouvoir le bien collectif.

À bien y réfléchir, et en écoutant les différents points de vue, je dois conclure sur une attitude plus nuancée, qui milite en faveur d'un travail approfondi de pédagogie en faveur d'une meilleure utilisation de l'analyse coût-bénéfice dans notre pays. Discutons ces points de vue.

L'argument standard consiste à dire qu'il existe tant de projets publics dont le rendement réel équivalent certain est supérieur à 4 % que la réalisation de ces projets créerait un très important déficit public. Cet argument ne me satisfait pas, et ceci pour plusieurs raisons. D'abord, je n'ai trouvé personne disposant d'information fiable sur l'élasticité de la demande de fonds publics à la baisse du taux d'actualisation. Existe-t-il tant de projets dont la rentabilité est comprise entre 4 et 8 % ? Ensuite, faire appel à une contrainte budgétaire exogène n'est intellectuellement pas très satisfaisant. S'il existe réellement beaucoup de projets publics à 6 %, une plateforme électorale devrait être aisément trouvée pour renverser un gouvernement qui s'opposerait à leur réalisation. Et si l'État est contraint par des engagements internationaux (Maastricht), de tels projets devraient être facilement délégués et financés par le secteur privé.

Mais en fait, je pense que le vrai problème ne se trouve pas là. Le non-dit, c'est qu'un taux d'actualisation élevé a été imposé par le passé en France pour faire face à la surestimation systématique de la rentabilité réelle des projets par les lobbies désireux de voir leur réalisation par l'État. En clair, puisque les projets sont annoncés avec une rentabilité supérieure, disons de 4 %, à leur rentabilité réelle par les lobbies,

l'État affiche un taux d'actualisation de 8% plutôt que de 4 %. Dans cette perspective, la proposition du Commissariat Général du Plan serait une catastrophe pour la gestion des fonds publics si cette réforme n'était pas combinée à un très gros effort de renforcement des capacités d'analyse coût-bénéfice de l'État, de manière à contrer l'excès d'optimisme stratégique des lobbies. Cet optimisme stratégique est d'autant plus difficile à combattre que la quantification des risques est délicate. L'accroissement de cette capacité d'expertise nécessite donc non seulement des hommes, mais aussi des données. L'État ne semble pas s'apprêter à offrir l'un ou/et l'autre.

Plus pernicieuse est l'idée qu'il faudrait ne faire que le bilan financier, plutôt que le bilan socio-économique, des projets d'investissement. Dans un bilan financier, on ne tient compte que des flux purement monétaires, c'est-à-dire que l'on ne tient compte que des euros qui rentrent et sortent des caisses publiques. C'est faire abstraction d'un grand nombre de coûts et bénéfices des actions publiques que l'État est incapable (ou qu'il refuse) de faire financer par l'usager : décès évités, pollution éliminée, temps gagné, patrimoine préservé, ... Quelle déchéance de l'analyse économique que de ne considérer parmi les coûts et bénéfices que ceux qui correspondent à des flux financiers. Le planificateur public se doit en effet tenir compte de l'ensemble de l'impact de ses actions sur le bien-être de ses administrés. Pour l'intérêt commun, il est nécessaire de combattre cette tentative d'affaiblissement du calcul économique. Si dans une perspective de partenariat public/privé, certains opérateurs privés désirent calculer leur intérêt financier à leur contribution au projet, rien ne leur interdit de calculer un bilan financier privé. Mais l'État s'intéressera seulement et entièrement au bilan socio-économique.

Mais cette réticence à faire le bilan socio-économique ne provient-elle pas d'une tension diffuse sur les valeurs tutélaires ? Prenons l'exemple de deux projets d'infrastructure de transport, A et B. Les deux projets ont une rentabilité réelle certaine de 6 %. Le bénéfice du projet A est purement financier. Le bénéfice du projet B provient pour l'essentiel d'un gain de temps pour les usagers et d'un certain nombre de vies gagnées. Le taux de 6 % est obtenu dans ce cas en utilisant les valeurs tutélaires du temps et de la vie (1 million d'euros). Il m'a semblé qu'il existe un biais en faveur du projet A sur le projet B. Le taux de 6 % du projet B, parce qu'il ne proviendrait pas d'un flux financier réel, semble prêter à caution. Si tel est le cas, c'est que les valeurs tutélaires ont été mal choisies. Celles-ci devraient alors faire l'objet soit d'une réévaluation, soit d'un effort de pédagogie.

Finalement, il ressort que les services de l'État se trouvent dans l'incapacité actuellement d'évaluer le risque des projets d'investissement dont il dispose. Or, il est entendu que la plupart de ceux-ci sont risqués, et certains sont même très risqués. De plus, ces risques sont en général fortement corrélés avec l'activité économique générale, de telle sorte que le théorème d'Arrow-Lind ne s'applique pas. La logique du calcul économique nécessite dès lors de déduire des flux socio-économiques espérés des primes de risque calculées selon les principes du Modèle d'Évaluation des Actifs Financiers. Ces primes de risque viennent évidemment amputer l'attractivité socio-économique des projets publics, et la rentabilité réelle équivalent certaine. Il me semble que la prise en compte du risque devrait fortement réduire le nombre de projets publics passant le test de la VAN équivalent-certain positive avec un taux d'actualisation de 4 %. Ce sentiment provient de l'analogie avec le secteur privé. Sur le marché américain tout au long du 20^e siècle, la rentabilité réelle des actifs sans risque a été de 1 % par an. Par contre, la rentabilité réelle moyenne des actions a été de 6,5 % par an. Cela nous donne une prime exigée par les agents économiques pour compenser le risque de 5,5 % par an (paradoxe de la prime de risque) ! L'incapacité de l'État à évaluer les risques pose alors un réel challenge à l'utilisation efficace des fonds publics. Cette incapacité est d'autant plus surprenante que des institutions de bien plus petites tailles telles que les banques, ont pour la plupart depuis 40 ans (Markowitz) développé des outils opérationnels pour mesurer et évaluer les risques. Si une telle incapacité des services de l'État devait persister, la solution de second-rang consisterait à accroître le taux d'actualisation. Il s'agirait d'une solution très inefficace, parce que l'hétérogénéité portant sur l'intensité des risques des différents projets est *a priori* élevée.

L'ensemble de ces commentaires converge pour demander soit une hausse des moyens d'évaluation de l'État français, soit un abandon de la politique *a priori* efficace d'un taux d'actualisation faible à 4 %.

5. CONCLUSION

Le risque intervient dans deux dimensions très différentes du calcul économique. Il intervient dans le choix du taux d'actualisation, car l'incertitude macroéconomique doit avoir un impact sur le niveau général de nos efforts pour améliorer notre bien-être futur ainsi que celui des générations qui nous succéderont. Le risque intervient aussi dans l'évaluation des bénéfices des investissements considérés. Seuls les investissements dont les risques sont indépendants du risque

macroéconomique et sur lesquels on n'anticipe pas d'information nouvelle sur leur rentabilité sociale escomptée doivent être évalués en supposant la neutralité au risque. Si leur rentabilité sociale est corrélée positivement au risque macroéconomique, il est nécessaire de réduire les cash-flows à actualiser d'une prime de risque, en utilisant les formules du MEDAF. Si des informations nouvelles sont anticipées à l'avenir sur cette rentabilité sociale, il faut intégrer dans la VAN du projet une valeur d'option de report de cet investissement.

L'abandon de la règle consistant à imposer un taux d'actualisation élevé pour tenir compte du risque a été rendu nécessaire pour des raisons à la fois de transparence et d'efficacité, étant donné la diversité des incertitudes des différents projets d'investissement public. Par contre, cet abandon complique singulièrement la tâche des évaluateurs. Il est maintenant nécessaire dans le calcul économique d'estimer les relations statistiques entre le rendement social du projet et le risque macroéconomique. En outre, il peut être nécessaire de modéliser la résolution dans le temps de l'incertitude sur ce rendement. Face à cette complexité et aux enjeux du calcul économique public, il est probablement nécessaire aujourd'hui de réfléchir à la constitution d'une capacité d'expertise et de contre-expertise dans ce domaine.

Références

- Arrow K.J. et Lind R.C., « Uncertainty and the evaluation of public investment decision », *American Economic Review*, 60, 1970, p. 364-378.
- Arrow K.J. et Fischer A.C., « Environmental preservation, uncertainty and irreversibility », *Quarterly Journal of Economics*, 88, 1974, p. 312-319.
- Cochran J., *Asset Pricing*, Princeton University Press, 2001.
- Constantinides G. M., « Intertemporal asset pricing with heterogenous consumers, and without demand aggregation », *Journal of Business*, 55, 1982, p. 253-267.
- Cox J., Ingersoll J. et Ross S., « A theory of the term structure of interest rates », *Econometrica*, 53, 1985, p. 385-403.
- Devezeaux de Lavergne J.G. et Gollier C., « Analyse quantitative de la réversibilité du stockage des déchets nucléaires : Valorisation des déchets », *Économie et Prévision*, 149, 2001.
- Dimson E., Marsh P. et Staunton M., *The Millenium book : A century of investment returns*, ABN-AMRO, Londres, 2000, <http://www.abnamro.com>.
- Gollier C., « Wealth inequality and asset pricing », *The Review of Economic Studies*, 68, 2001a, p. 181-203.

- Gollier C., *The economics of risk and time*, MIT Press, Cambridge, MA, 2001b.
- Gollier C., « Discounting an uncertain future », *Journal of Public Economics*, 85, 2002a, p. 149-166.
- Gollier C., « Time horizon and the discount rate », *Journal of Economic Theory*, 107, 2002b, p. 463-473.
- Hall R.E., « Intertemporal substitution of consumption », *Journal of Political Economy*, 96, 1988, p. 221-273.
- Henry C., « Investment decisions under uncertainty : the irreversibility effect », *American Economic Review*, 64, 1974, p. 1006-1012.
- Lucas R., « Asset prices in an exchange economy », *Econometrica*, 46, 1978, p. 1429-46.
- Smit H.T.J. et Trigeorgis L., *Strategic investment : Real options and games*, Princeton University Press, 2004.
- Vasicek O., « An equilibrium characterization of the term structure », *Journal of Financial Economics*, 5, 1977, p. 177-188.
- Weitzman M.L., « Statistical discounting of an uncertain distant future », mimeo, Harvard University, 2004.

CHAPITRE 4

INTÉGRATION DU RISQUE ET DE L'INCERTITUDE DANS LA CONSTRUCTION DU CALCUL ÉCONOMIQUE¹

L. Andrieu, CERMICS, ENPC

A. de Palma, THEMA,

Université de Cergy-Pontoise, ENPC et IUF

N. Picard, THEMA,

Université de Cergy-Pontoise et INED

Résumé. Ce document constitue l'une des douze contributions du réseau de recherche sur les dimensions critiques du calcul économique. Nous nous intéressons ici à la question du risque et de l'incertitude dans la construction du calcul économique et nous essayons plus particulièrement de distinguer ce qui relève de l'analyse (comment évaluer l'aversion au risque par exemple), et ce qui pourrait relever des recommandations. Dans ce document, nous commençons par rappeler les idées de base de l'analyse de projets, à savoir l'analyse des coûts et des bénéfices, et la façon dont les notions de risque et d'incertitude sont actuellement intégrées dans cette analyse. Nous rappelons ensuite les notions de base en économie et en finance permettant de comparer des variables aléatoires (MEDAF et ordre stochastique), puis nous présentons différentes mesures de risque développées essentiellement dans le monde financier. Ces mesures, intégrées dans une analyse de projet permettraient de mieux contrôler les risques

1. Nous remercions S. Proost (K.U. Leuven, Belgique), C. Gollier (IDEI, Université de Toulouse), E. Dumas (ENPC), J.-L. Prigent (THEMA, Université de Cergy-Pontoise), R. Lindsey (University of Alberta, Canada) et W. Rothengatter (Université de Karlsruhe, Allemagne) pour leurs suggestions et commentaires. Nous remercions aussi J. Maurice (ENPC) et E. Quinet (ENPC) pour leurs commentaires et remarques détaillées. Cet article a été présenté lors de la conférence « First International Conference on Funding Transportation Infrastructure » (Banff, Canada, août 2006).

associés à la réalisation d'un projet. Nous distinguons en particulier deux classes de mesures de risque : la première classe, qui regroupe des mesures telles que *l'expected shortfall*, permet de contrôler les risques « en l'absence d'évènements exceptionnels » ; la seconde, qui comprend par exemple la *Conditional Value-at-Risk*, permet la prise en compte des risques extrêmes. Enfin, nous introduisons un ratio d'efficacité permettant d'évaluer la performance d'un projet par rapport à un projet de référence.

1. INTRODUCTION

L'objectif de ce document est de proposer un guide des principaux critères permettant de prendre en compte le risque dans le calcul économique appliqué au choix des projets d'investissement dans le domaine des infrastructures des transports.

Dans l'analyse coûts-bénéfices (ACB) classique, la valeur d'un projet est égale à la somme des bénéfices actualisés, diminuée de la somme des coûts actualisés. Cette approche suppose la plupart du temps (et presque toujours en pratique) que ces flux financiers sont déterministes. Cette hypothèse est rarement vérifiée. Les aléas (sur l'offre et la demande) affectent non seulement les paramètres inclus dans l'ACB (par le truchement de la prime de risque, par exemple), mais aussi la manière dont les flux financiers sont pris en compte lors de l'évaluation de projets. Cette évaluation nécessite dans ce cas de prendre en compte les préférences du preneur de décision. Ces préférences sont censées agréger d'une certaine manière les préférences des citoyens affectés par le projet (soit directement, par ses effets perçus, soit indirectement, par la fiscalité inhérente). Ces préférences sont traduites par une fonction d'utilité. La manière dont ces aléas affectent la prise de décision dépend de la façon d'introduire l'attitude face au risque et ses liens avec l'actualisation.

La méthode la plus couramment utilisée dans la théorie du consommateur consiste à introduire une attitude face au risque – à chaque période. L'avantage de cette méthode consiste à ne pas supposer une compensation parfaite des fluctuations d'une période à l'autre. Son inconvénient majeur est de ne pas pouvoir prendre en compte les corrélations entre aléas aux différentes périodes. Cette approche, qui est bien adaptée à la théorie du consommateur, est davantage problématique dans l'analyse des flux financiers.

La méthode proposée ici nous semble plus adaptée à l'évaluation de projets par le décideur public. En effet, celui-ci est plus intéressé par la valeur actualisée des flux financiers de chaque projet dans chaque

état de la nature que par une agrégation des risques à chaque période. Ceci ne signifie pas que le preneur de décision public soit neutre au risque. De fait, nous supposons qu'il est indifférent à la répartition temporelle du risque (moyennant une fonction d'actualisation des coûts et des bénéfices aléatoires). L'attitude du décideur public face au risque est ici modélisée sur le résultat de ce calcul d'actualisation. À l'inverse, la méthode classique évoquée dans le paragraphe précédent actualise les utilités espérées. Ces dernières modélisent séparément les attitudes face aux risques des différentes périodes. La méthode classique revient alors implicitement à supposer ces aléas indépendants entre eux. Cette hypothèse nous paraît très difficilement défendable dans le cadre de l'évaluation de projets.

L'instruction cadre préconise un taux d'actualisation de 3,5 % net par an pour une période à 50 ans et un taux de 3 % net au-delà de 30 à de 50 ans. D'après [39], ce taux doit être modifié pour prendre en compte les différents risques inhérents à chaque projet. Nous pensons que la prise en compte du risque par le truchement d'une modification du taux (préconisée dans le document [39]) doit être complétée par une étude plus fine (et davantage transparente) qui prenne explicitement en compte l'ensemble des causes de risque internes et externes. Cette étude sera l'objet du présent article.

Ce document s'organise en quatre sections. La section 2 propose un bref rappel des pratiques courantes permettant d'apprécier un projet d'investissement : nous discutons d'une part de l'ACB, et d'autre part de la façon d'introduire la notion de risque dans l'évaluation de projets. Il s'agit là d'une étape fondamentale de l'évaluation d'un projet, car le risque est responsable historiquement de la plupart des échecs et faillites. Nous rappelons dans la section 3 quelques notions de base permettant de comparer des actifs financiers, puis nous proposons deux critères d'évaluation de projets. Ces critères, très répandus en finance et dans le domaine de l'assurance depuis déjà quelques années, permettent de mieux contrôler le risque associé à la réalisation du projet. Bien que cette approche ne nécessite pas la connaissance d'une fonction d'utilité, nous sommes capables de la relier à l'approche en termes d'utilité espérée. Enfin, des conclusions succinctes sont présentées dans la section 4.

2. L'ANALYSE DES COÛTS ET DES BÉNÉFICES

2.1. Les enjeux

L'analyse des coûts et des bénéfices² constitue la plus importante des informations requises pour les décisions d'investissement. Il s'agit d'une procédure permettant d'évaluer l'utilité d'un projet à travers le surplus qu'il dégage (exprimé en termes monétaires). Les idées de base de l'ACB remontent à 1844 ([23]), mais c'est entre 1939 et 1960 que cette analyse connaît son véritable essor, notamment grâce à M. Allais, P. Massé et M. Boiteux (voir par exemple [2], [37] ou [8]).

Au niveau européen, cette analyse est aujourd'hui explicitement exigée par les nouveaux règlements régissant les fonds structurels pour l'appréciation des grands projets : elle sert d'instrument de prise de décisions collectives entre la Commission européenne, les états membres, les auteurs du projet, *etc.* (cf. [29], [30] ou [34]). Dans [30], les auteurs recommandent que l'ACB énonce les points suivants :

- l'objectif de la nouvelle infrastructure,
- les raisons qui motivent sa réalisation,
- une description de l'infrastructure actuelle que remplacera l'infrastructure nouvelle,
 - un aperçu des problèmes et des risques que comporte l'infrastructure actuelle,
 - une description du fonctionnement de la nouvelle infrastructure,
 - une énumération des bénéfices espérés,
 - les problèmes ou risques qu'implique la nouvelle infrastructure,
 - une estimation des coûts que pourrait comporter un possible échec de l'infrastructure,
 - un état des coûts de l'infrastructure actuelle,
 - les bénéfices qui seront réalisés,
 - un aperçu de sa durée de vie possible.

Concrètement, cette analyse comporte deux types d'évaluations distinctes : l'analyse socio-économique et l'analyse financière. Bien que ces deux analyses reposent sur des éléments communs, elles sont souvent effectuées de manière assez autonome.

2. Nous utilisons ce terme à la suite de J.-J. Laffont, là où d'autres emploient le terme d'analyse des coûts et des avantages.

2.2. Les pratiques

2.2.1. L'analyse socio-économique

L'analyse socio-économique évalue la contribution du projet au bien-être économique de la région ou du pays. Elle est réalisée au nom de l'ensemble de la société (et pas au nom du gestionnaire de l'infrastructure, comme dans l'analyse financière). Dans le domaine des transports, les objectifs socio-économiques sont généralement liés à l'amélioration des conditions de voyage pour les marchandises et les passagers, ainsi qu'à l'amélioration de la qualité de l'environnement et du bien-être de la population bénéficiaire.

Le secteur des transports étant souvent caractérisé par l'existence de « tarifs administratifs » (lorsque les modes de transports collectifs sont subventionnés par exemple) et de coûts « externes » (les coûts environnementaux tels que la pollution, le bruit ou les accidents), l'analyse économique permet donc d'éclairer certains aspects spécifiques. Elle comporte trois phases :

1. *la correction des taxes/subventions et autres transferts* : on déduit des flux de l'analyse financière les paiements qui n'ont pas de contrepartie réelle en ressources, comme les subventions et taxes indirectes sur les intrants et les extrants,

2. *la correction des externalités* : on inclut dans les sorties et les entrées les coûts et avantages externes pour lesquels il n'existe pas de cash-flow (comme le temps épargné par l'investissement dans les transports par exemple),

3. *la conversion des prix du marché en prix fictifs* : on inclut les coûts et avantages sociaux.

Les résultats de cette analyse peuvent être exprimés de différentes manières, notamment par le *taux de rentabilité économique* (TRE qui est un indicateur de la profitabilité socio-économique d'un projet défini comme le rapport entre l'excédent brut d'exploitation et le capital).

Pour une présentation plus détaillée de l'analyse socio-économique, notamment dans le secteur des transports, on pourra consulter [29] ou [30].

2.2.2. L'analyse financière

L'objet de l'analyse financière est d'utiliser les prévisions du cash-flow (flux de trésorerie) du projet afin de calculer des taux de rendement appropriés : la *valeur actuelle nette* (VAN) et le *taux de rentabilité interne* (TRI). Dans le secteur des transports, cette analyse sera géné-

ralement conduite du point de vue du gestionnaire public ou privé de l'infrastructure.

Le critère fondamental de l'analyse financière est la VAN, qui mesure la création de valeur engendrée par l'investissement. Elle est donnée, pour l'horizon n (durée de l'amortissement ou durée de la concession), par la formule suivante :

$$\text{VAN}(S, r) = \sum_{t=0}^n \frac{S_t}{(1+r)^t} + \frac{R_n}{(1+r)^n}, \quad (1)$$

où S_t est le flux financier³ net pour l'année t , $S = (S_0, \dots, S_n)$, r est le taux d'actualisation choisi dont la valeur est fixée par le Commissariat Général du Plan ([39]) et R_n est la valeur résiduelle de l'investissement en période n . Si la VAN de l'investissement est positive, cela signifie que la recette totale actualisée est supérieure aux dépenses actualisées : l'investissement est rentable au taux d'actualisation choisi. Le critère d'acceptation est donc que la VAN soit positive en prenant comme taux d'actualisation le taux du marché.

Il existe dans l'analyse financière un critère auxiliaire, le TRI. Le TRI d'un projet est le taux qui annule la VAN. L'investissement est rentable si le TRI obtenu est supérieur ou égal à un taux d'actualisation fixé. Bien que très utilisé, ce critère peut dans certains cas être un faux ami. En effet, lorsqu'il y a plusieurs changements de signe dans le flux de trésorerie d'un projet, ce projet peut présenter plusieurs taux de rentabilité interne. Dans ce type de situation, le critère à retenir est alors celui de la VAN, mais une étude de sensibilité de la VAN au taux d'intérêt r est conseillée. Plaçons-nous par exemple en temps continu (T désigne l'horizon de temps considéré), on a alors :

$$\text{VAN}(S, r) = \int_0^T S_t e^{-rt} dt,$$

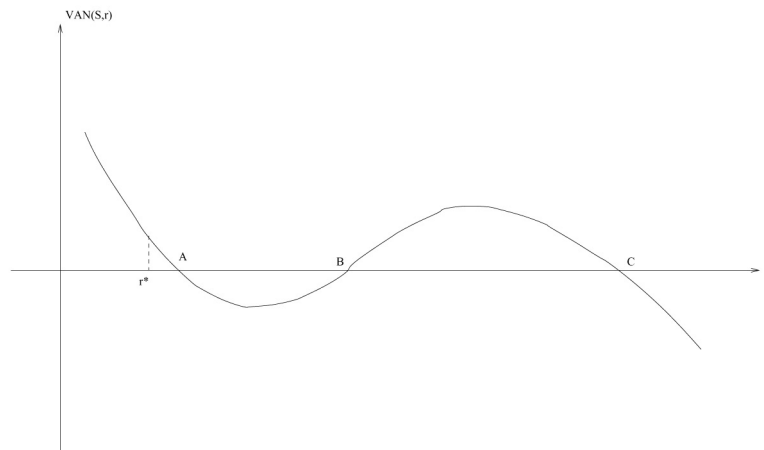
et donc

$$\frac{\partial \text{VAN}(S, r)}{\partial r} = - \int_0^T t S_t e^{-rt} dt.$$

3. Le flux financier net S , dépend de l'excédent brut d'exploitation ainsi que des investissements effectués.

En règle générale, $\frac{\partial VAN}{\partial r} < 0$, mais on peut avoir $\frac{\partial VAN}{\partial r} > 0$ dans certains cas où S_t est négatif pendant certaines périodes. Cette situation est représentée sur la Figure 1.

Figure 1
Sensibilité de la VAN au taux d'intérêt



Sur ce schéma, A , B et C sont les différents TRI du projet. On note r^* le taux du marché. On observe alors que le critère de la VAN va être très sensible au paramètre r : en sous-estimant le taux d'intérêt on aura une VAN positive, alors qu'en le surestimant légèrement (si $r^* > A$ mais $r^* < B$), on aura une VAN négative.

Le critère de la VAN suppose que les revenus dégagés sont réinvestis au taux d'actualisation alors que le critère du TRI suppose un réinvestissement au taux trouvé : le critère du TRI peut ainsi être très optimiste. Il semble donc souhaitable de privilégier le critère de la VAN. Pour plus de détails sur l'analyse financière, on pourra se reporter à [21].

Dans l'ACB, l'appréciation d'un projet d'investissement se fait toujours par rapport à un projet d'investissement alternatif (statut quo ou solutions alternatives). On peut pour cela utiliser l'un des critères de décision usuels en avenir incertain non probabilisable, à savoir le critère de Savage (ou critère « Min-Max » de regret). Ce critère consiste à identifier pour chacun des scénarios le projet le plus favorable, puis à évaluer le manque à gagner que représenterait par rapport à ce projet l'adoption de chacun des autres projets, et enfin à retenir le projet conduisant au plus petit des regrets maximums. Ce critère de décision traduit clairement une attitude de prudence de la part de l'investisseur.

Supposons par exemple qu'un décideur ait à choisir un projet d'infrastructures parmi trois projets A (solution de référence), B et C ayant comme VAN respectives (en millions d'euros) dans les scénarios ξ_1 et ξ_2 :

Table 1
VAN associées au projet

	ξ_1	ξ_2
A	19.8	19
B	18.2	20
C	21.3	18.4

On commence tout d'abord par déterminer pour chacun des scénarios le projet ayant la plus grande VAN : il s'agit pour ξ_1 du projet C , et pour ξ_2 du projet B . On en déduit la matrice des regrets suivante :

Table 2
Matrice des regrets associée au projet

	ξ_1	ξ_2	regret maximum
A	1.5	1	1.5
B	3.1	0	3.1
C	0	1.6	1.6

On recherche alors le minimum des regrets maximaux, qui correspond ici au projet A . Ce critère conduit donc à retenir le seul projet dont la VAN n'est maximale dans aucun scénario.

Concernant les investissements en infrastructures par exemple, il est souhaitable d'inclure des comparaisons avec la situation antérieure (A dans l'exemple précédent), avec les itinéraires alternatifs, et enfin avec les systèmes de transport alternatifs éventuels dans le cadre de projets routiers. La solution de référence correspond généralement à une décision de ne rien faire. Toutefois, dans certains cas, cela peut créer un problème dans le secteur des transports. En effet, si la solution de référence était « catastrophique », *i.e* si la décision de ne pas investir entraînait une paralysie du trafic et donc un coût social élevé, tout projet entraînerait des avantages importants, quel que soit son prix. Dans le cas d'un phénomène de congestion grave, pour éviter de déformer les résultats de l'analyse, il est donc nécessaire de configurer une solution de référence intégrant des interventions minimales.

Les idées générales de l'ACB ayant été rappelées, intéressons-nous maintenant à la façon dont le risque est intégré dans le choix de projets.

2.3. *Risque, incertitude et choix de projets*

Comme nous l'avons dit dans l'introduction, l'un des objectifs principaux de ce document est de réfléchir à la prise en compte du risque et de l'incertitude dans l'évaluation de projets d'investissements. Il est donc important de bien distinguer ces deux notions : nous parlerons de risque lorsque les résultats futurs (revenus, coûts) ne sont pas connus de manière certaine, mais sont entachés d'aléas, auxquels le modélisateur peut associer une loi de probabilité. Si cette opération n'est pas possible, on parle d'incertitude, et c'est alors à chaque individu (modélisateur ou preneur de décision) de se représenter les lois de probabilités gouvernant les événements aléatoires.

La distribution des aléas peut être estimée à partir de données historiques. Notons néanmoins que lorsque les événements sont rares (par exemple, catastrophes naturelles, instabilité politique, etc.), le modélisateur sera obligé de postuler des lois de probabilité. La notation correspondante dans ce cas est celle de probabilités subjectives, telle qu'introduite par Savage. Lorsque le modélisateur ignore totalement les probabilités des événements rares, l'incertitude est totale et des critères de type « Min-Max » sont recommandés.

Le risque et l'incertitude restent encore largement absents de l'ACB. La prise en compte de l'incertitude est plus difficile si la période d'actualisation est plus longue car les aléas ont dans ce cas plus de chance de se réaliser, se combiner, et avoir des conséquences plus significatives. Dans beaucoup d'études pratiques, les travaux permettant d'estimer la demande et les coûts se contentent de prendre des taux de croissance constants, qui sont de plus en plus discutables lorsque la période considérée est plus longue. Non seulement les taux de croissance ne sont pas constants, mais de plus ils sont entachés d'aléas. Pour prendre en compte l'aléa, il est nécessaire de caractériser les lois de probabilité associées à l'évolution de la demande et l'évolution des coûts. Une approche simple consiste à spécifier la manière dont la variabilité des prédictions (sur la demande et les coûts) augmente au cours du temps. De cette manière, on associe l'économétrie à l'ACB. Cette démarche est nouvelle et reste à être développée.

2.3.1. *Risque et incertitude*

Lorsque l'avenir est probabilisable (cf [31]), on peut continuer à utiliser des critères de type de celui du bénéfice actualisé (analyse

socio-économique) ou de la VAN (analyse financière) : si le taux d'intérêt r est exogène, il suffit de pondérer les variables par leur probabilité respective de façon à pouvoir calculer pour chaque projet en concurrence l'espérance mathématique de la VAN,

$$\mathbb{E}[\text{VAN}(S, r)] = \sum_{t=0}^n \frac{\mathbb{E}[S_t]}{(1+r)^t} + \frac{\mathbb{E}[R_n]}{(1+r)^n}, \quad (2)$$

ou sa variance

$$\begin{aligned} \text{var}[\text{VAN}(S, r)] &= \sum_{t=0}^n \frac{\text{var}[S_t]}{(1+r)^{2t}} + \frac{\text{var}[R_n]}{(1+r)^{2n}} \\ &+ 2 \sum_{t \neq t'} \frac{\text{cov}(S_t, S_{t'})}{(1+r)^{t+t'}} + 2 \sum_{t=0}^n \frac{\text{cov}(S_t, R_n)}{(1+r)^{t+n}}, \end{aligned} \quad (3)$$

avec $\text{cov}(S_t, S_{t'})$ la covariance de S_t et $S_{t'}$. L'approche moyenne-variance décrite ici constitue une première approximation permettant de tenir compte d'éléments stochastiques dans la VAN. Nous reprendrons dans la section 2.3.3 la discussion de la VAN stochastique.

Lorsque les risques affectant les différentes variables d'un projet ne sont pas indépendants, il est utile de mener en plus une analyse plus approfondie de la distribution de la VAN ; cette analyse permettra par exemple de déterminer avec quelle probabilité la VAN d'un projet sera supérieure à une valeur donnée *a priori*.

Lorsque les risques ne sont pas probabilisables (voir [10]), il faut élaborer tous les scénarios envisageables, puis calculer pour chacun de ces scénarios les indicateurs souhaités (VAN, ...). L'emploi d'un critère issu de la théorie des jeux, tel le critère du Min-Max de regret de Savage, peut également venir compléter cette analyse de projets.

Dans le secteur des transports, il existe plusieurs sources d'aléas possibles :

- la demande,
- les coûts de production : facteurs de production, coûts des matières et de l'énergie, salaires, internalisation des nuisances locales ou globales (effet de serre)...
- le contexte concurrentiel et la régulation (Eurostar),
- le temps d'exécution,
- les variables économiques (macro-économiques, régionales,...),
- les variables financières,

- les risques spécifiques à la concession (durée, obligations, rachat, déchéance,...),
- activation de clauses juridiques,
- les ressources humaines pour la gestion du projet,
- les risques géologiques (cf. la liaison Lyon-Turin).

On distingue également les risques diversifiables et les risques non diversifiables. En particulier, les chocs macro-économiques, le climat et l'environnement géo-politique qui affectent la demande sont la conséquence de chocs non diversifiables. Il est important de différencier les investissements majeurs en terme de degré de diversifiabilité. En effet, les primes de risques à prendre en compte dans l'ACB sont plus élevées pour les risques non diversifiables (voir [9]). La situation est plus complexe dans le cas des PPP (public-private partnership) étant données les difficultés à définir ex ante le partage des risques subis ex post.

Dans le secteur du transport routier par exemple, les facteurs critiques influençant la réussite d'un investissement incluent les prévisions en matière de circulation, le manque d'élasticité de l'investissement (une capacité excessive est souvent requise au début de l'exercice) ou encore l'influence décisive des activités secondaires (l'efficacité d'une autoroute est fonction d'un bon réseau d'embranchement par exemple).

2.3.2. *L'analyse de la sensibilité et des risques*

L'analyse des risques consiste à étudier la probabilité qu'un projet donnera des résultats satisfaisants (en termes de VAN par exemple), ainsi que la variabilité du résultat par rapport à la meilleure estimation préalablement effectuée. La procédure actuellement recommandée pour l'évaluation des risques repose sur les deux éléments suivants : l'analyse de sensibilité et l'analyse de probabilité des risques. Une présentation plus approfondie de ces deux analyses peut être trouvée dans [29] ou [30]. Rappelons que, avant que cette procédure ne soit recommandée, le risque était pris en compte via une augmentation du taux d'actualisation. Dans la procédure actuelle, le taux d'actualisation est choisi indépendamment des risques, et une analyse explicite des risques est demandée (voir [39]).

L'objectif de l'analyse de sensibilité (voir section 7.4 du document [39]) est de sélectionner les variables et paramètres « critiques » du modèle, *i.e.* ceux dont les variations par rapport à la valeur utilisée comme meilleure estimation dans le cas de référence entraînent les modifications les plus importantes sur les indices économiques et financiers calculés (bénéfice actualisé, VAN,...). Parallèlement, les

simulations financières concernent l'étude des capacités de réaction du système après un choc important. Bien évidemment, les hypothèses comportementales utilisées jouent un rôle crucial, généralement sous-estimé. Dans le secteur des transports, l'analyse de sensibilité peut se concentrer par exemple sur les coûts d'investissement et de fonctionnement ou sur la demande attendue, en particulier le trafic généré.

L'analyse de probabilité des risques consiste en l'étude des probabilités d'un projet d'obtenir un taux de rendement satisfaisant et de la variabilité par rapport à la meilleure estimation du taux de rendement.

2.3.3. *Prise en compte du risque dans le secteur des transports*

On se limitera ci-après à l'analyse du risque en termes financiers, à titre d'exemple, la transposition à l'analyse du risque en termes économiques ne soulevant pas de difficulté particulière. Cette section est rédigée à partir de [17]⁴.

VAN stochastique. Le principe du taux d'actualisation de 8 % (qui est censé prendre en compte les risques) permet d'écarter des projets à rentabilité moyenne faible, mais ne s'adresse pas réellement aux sources de risques. Ce critère a aujourd'hui été abandonné. Une autre manière de modéliser le risque (lorsqu'il est modélisé) consiste à prendre le modèle déterministe et à surimposer une variable aléatoire (sur le coût, la demande, etc.). En raisonnant sur base de l'équation de la VAN, on peut ainsi en dériver une version stochastique dans laquelle on s'intéresse à la variable aléatoire VAN_{sto} :

$$VAN_{sto}(I, R, C, i) = \sum_{t=0}^n \frac{-\lambda I(t) + R(p, K_t, \epsilon, t) - C(p, K_t, \eta, t)}{(1+i)^t}$$

où R désigne le revenu du projet, C son coût, I_t est l'investissement en t , K_t représente le stock de capital en t , p le niveau des péages, ϵ et η représentent des termes aléatoires et enfin i est le taux d'actualisation. Notons que ces deux termes ϵ et η peuvent être corrélés (ces corrélations ne sont pas toujours prises en compte, même dans les simulations de Monte Carlo⁵). Dans ce cadre probabiliste, l'évaluation d'un projet

4. Le lecteur pressé pourra passer à la Section 3, sans que la compréhension de ce document en soit altérée. Nous référons aussi le lecteur au chapitre de M. Ivaldi, dans le présent ouvrage.

5. De façon générale une méthode de Monte Carlo estime l'espérance d'une variable aléatoire X par la moyenne arithmétique sur un échantillon de n simulations indépendantes de celle-ci : $\mathbb{E}[X] \approx \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$. La loi forte des grands nombres assure la convergence presque sûre de \bar{X}_n vers $\mathbb{E}[X]$ et le théorème limite centrale fournit l'erreur pour n grand.

se mesure par une densité de probabilité (associée à la variable aléatoire). On s'intéressera, par exemple, à la probabilité que la VAN_{sto} soit plus petite qu'une valeur cible, ou à la probabilité que VAN_{sto} soit inférieure à une valeur donnée, ou encore à la probabilité que la valeur de la VAN_{sto} soit négative (ou l'on utilisera d'autres indicateurs de performance). Cette formulation nous amènera à réfléchir à d'autres stratégies de prise en compte et de gestion des risques (voir la section 3.2.1).

Signalons enfin que les risques majeurs (les catastrophes naturelles, les attentats, les grèves de durée excessive, etc.) ne semblent être que très rarement pris en compte par les opérateurs majeurs.

Fonction de survie. L'un des risques importants pour certains types de projets (par exemple, les chemins de fer) est l'abandon du projet ou de la variante de ce projet parce que son exploitation est devenue trop coûteuse pour l'exploitant. Ceci peut survenir quand il y a un changement de la politique locale de transport ou lorsqu'une entreprise arrête ses activités. Une approche autre que l'utilisation d'une période plus courte pour le test financier serait d'utiliser une fonction de survie. L'excédent brut d'exploitation pour chaque année future serait alors pondéré par un coefficient de survie (par exemple 95 % de survie chaque année). Les coefficients de survie peuvent être calculés pour les projets individuels (en fonction, par exemple, de la durée des mandats politiques régionaux) ou peuvent être des fonctions basées sur l'expérience. La fonction de survie fournit des résultats différents par rapport à la méthode consistant à considérer des périodes d'évaluation plus courtes. Certaines fonctions de survie (exponentielles) conduisent à des résultats similaires à l'utilisation d'un coût de capital accru.

Incitation et asymétrie d'information. Les évaluations des coûts relatifs aux travaux de construction ou de maintenance sont potentiellement biaisées parce que ces coûts ne sont pas totalement internalisés par le principal (constructeur) et parce qu'il n'existe pas un marché totalement compétitif. Les travaux pris en charge par l'agent (l'exploitant) ou par tout autre organisme ne sont pas suffisamment incitatifs étant donné que les bénéfices et les coûts ne sont pas partagés dans le cadre de contrats qui spécifient les prises en charge en cas de dépassement des coûts projetés ou en cas de rallongement de la durée des travaux.

La demande projetée est entachée de risque. Ces risques seront évalués de manière différente par les différents acteurs en présence : États, régions, opérateurs publics et privés.

Les différents agents ont différents niveaux d'information, différentes incitations et différents niveaux de contrôle de la demande. La participation au risque doit être fonction de ces différences. Les asymétries d'information jouent aussi un rôle important dans les négociations en cours (modèle principal/agent). Pour plus de précisions on pourra consulter [33] ou [43] (ouvrage plus synthétique).

Partage des risques entre le principal et l'agent. La participation d'autres partenaires dans l'investissement semble importante : l'investissement dans une ligne de chemin de fer à vocation régionale constitue ainsi un exemple type d'un problème de « hold-up » en économie industrielle. En effet, une fois qu'un responsable d'infrastructure A (le principal) a consenti aux investissements (sunk cost) en infrastructure et en matériel et les a réalisés, il est tentant pour l'exploitant régional B (l'agent) de renégocier les péages vers le bas étant donnée la position captive (et asymétrique) de A (ceci sera moins le cas, une fois que les marchés seront effectivement ouverts à la concurrence internationale).

Le niveau de risque supporté par A est une fonction décroissante de sa participation dans le projet. Prenons l'exemple d'un projet qui nécessite un investissement total de 100. La rentabilité est telle que A investit 20 et les autres partenaires 80. Les bénéfices actualisés pourraient varier entre 80 et 120. Le risque pour A est donc, en termes actualisés, une rentrée nette de +20 à -20 pour un investissement de 20. Imaginons que le même projet est légèrement moins rentable et que ceci amène A à investir uniquement 10. Les risques n'ayant pas changé, A risque maintenant de +20 à -20 pour un investissement de 10. La diminution de la participation a donc donné lieu à une augmentation de ses risques. Le problème de base est que A porte tout le risque financier parce que les autres parties remettent un subside fixe au moment de l'investissement. Dans une entreprise privée, un risque accru pour un investissement nécessite une augmentation de la partie financée par des fonds propres. Cette logique semble inversée dans le cas de A.

Enfin, les différents acteurs en présence possèdent des degrés d'aversion au risque différents. De fait, une règle stricte de non dépassement d'un budget correspond à une aversion au risque localement infinie, ce qui n'est pas raisonnable. Il est vrai que les individus (par exemple, les investisseurs) se focalisent sur certains chiffres (les pertes possibles par exemple) mais il n'y a pas de raison qu'une organisation telle que A possède une aversion pour le risque localement infinie.

De plus, en toute logique, les partages du risque doivent se faire en fonction de l'aversion au risque de chacun des agents. La règle théorique de partage des risques est simple, puisqu'elle préconise que l'agent

doté de l'aversion au risque la plus forte supporte la plus petite quantité de risque. Il reste à déterminer comment les différents acteurs en présence (principalement État, régions, les entreprises A et B), peuvent être classés en fonction de leur degré d'aversion au risque. Le partage des risques peut diminuer son impact pour l'ensemble des agents. Par exemple, l'entreprise B et les régions peuvent participer aux surplus et aux déficits comme le ferait le capital à risque dans le cas d'une entreprise privée.

Flexibilité des prix. Les prix peuvent être déterminés par un monopole en fonction des coûts et de la demande estimée. Ceux-ci peuvent aussi être le résultat de la concurrence entre des opérateurs privés et des opérateurs publics. La prise en compte de la concurrence ne peut être faite sans l'aide d'un modèle économique sous-jacent.

Nous avons développé, dans le cadre des projets REVENUE et FUNDING (projets européens de DG TRENEN) un outil pour l'ACB : MOLINO I (pour REVENUE I) et MOLINO II (pour FUNDING, voir [18] et [19]). Cet outil permet de calculer l'évolution des recettes (transports individuels ou en commun et transports de fret) pour des infrastructures (route, rail, air, etc.). MOLINO, qui est basé sur un modèle CES pour les usagers, ainsi que pour les entreprises, résout deux types de problèmes. Le premier est celui du choix des agents (choix de routes ou de temps de départ). Dans ce cas, l'équilibre est un équilibre entre l'offre et la demande. Lorsque les biens sont des substituts parfaits, l'équilibre est celui de Wardrop. Sinon, il s'agit d'un équilibre qui égalise les temps de trajet généralisé attendus. Le second équilibre est un équilibre économique, en fonction de la structure du marché : monopole d'État, concurrence entre opérateurs privés, entre opérateurs privés et services publics, etc. L'impact de la flexibilité des prix, pris en compte dans MOLINO, est illustré en appendice.

Dans la section suivante, nous rappelons quelques concepts fondamentaux en économie et en finance permettant de comparer différents actifs financiers, et suggérons d'intégrer certains d'entre eux dans l'ACB.

3. MESURES DE RISQUE ET ACB

3.1. *Brefs rappels sur l'évaluation des actifs financiers*

3.1.1. *Le MEDAF*

Le Modèle d'Évaluation des Actifs Financiers (MEDAF), ou Capital Assets Pricing Model (CAPM) développé au début des années 70

par Sharpe est utilisé pour évaluer des actifs financiers dans un marché en équilibre. Nous présentons ici les fondements de ce modèle dans son contexte originel, pour une présentation dans le contexte de la VAN, voir le chapitre de C. Gollier dans cet ouvrage. Ce modèle est basé sur le fait que seul le risque de marché r_M , ou risque diversifiable, est rémunéré par les investisseurs dans un tel marché. La rentabilité R exigée par un investisseur est alors égale au taux de l'argent sans risque r_F majoré d'une prime de risque uniquement liée au risque de marché de l'actif :

$$R = r_F + \beta(r_M - r_F),$$

où β est le coefficient de volatilité de l'actif indiquant la relation existant entre les fluctuations de la valeur de l'actif et les fluctuations du marché.

Ce modèle est devenu la référence théorique de toute l'industrie de la finance. L'hypothèse forte de ce modèle est que, pour choisir entre différents portefeuilles possibles, les individus choisissent sur la base du rendement espéré et de la variance du rendement. On peut montrer que cela revient à supposer que les distributions sont gaussiennes, ou bien que les individus ont des fonctions d'utilité quadratiques.

Bien que très populaires, le critère « moyenne/variance » et le critère « moyenne/écart-type », ne suffisent pourtant pas toujours à comparer deux variables aléatoires, et peuvent même conduire à des aberrations, comme le montrent P. Bernhard et G. Cohen dans [12]. En effet, supposons que l'on souhaite maximiser un rendement J pour les décisions A et B, et pour les « états de la nature » ξ_1 et ξ_2 équiprobables :

Table 3

VAN associées au projet

	ξ_1	ξ_2
A	0	20
B	21	21

On a $\mathbb{E}[J(A), \xi] = 10$ et $\mathbb{E}[J(B), \xi] = 21$. On prend comme mesure de l'incertitude la variance : $\sigma^2(J(A), \xi) = 100$ et $\sigma^2(J(B), \xi) = 0$. Alors, si on choisit de maximiser $\mathbb{E}[J] + \alpha \sigma^2(J)$, avec $\alpha > 0$, la stratégie A devient préférable à la stratégie B dès que $\alpha > 0.11$. Cependant, ce résultat est absurde car la stratégie B est optimale, puisqu'elle domine (rendement supérieur) la stratégie A dans tous les états de la nature.

Une approche systématique pour comparer des variables aléatoires entre elles consiste à utiliser la notion d'ordre stochastique. Un ordre stochastique peut être vu comme un ordre partiel sur un ensemble de fonctions de répartition.

3.1.2 Ordres stochastiques et fonctions d'utilité

Supposons qu'un investisseur compare des rendements aléatoires dont il connaît les lois. Si ses préférences satisfont certaines propriétés, Von Neumann et Morgenstern (1944, voir [40]) ont montré qu'elles peuvent se représenter par l'espérance d'une fonction. Plus précisément :

Définition 3.1 (le critère d'espérance d'utilité). Les préférences d'un individu satisfont au critère d'espérance d'utilité s'il existe une fonction croissante u appelée fonction d'utilité telle que l'individu préfère le rendement aléatoire X au rendement aléatoire Y si et seulement si l'espérance d'utilité de X est supérieure à celle de Y :

$$X \text{ préféré à } Y \Leftrightarrow \mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]. \quad (4)$$

La croissance de la fonction d'utilité exprime simplement le fait que l'investisseur aime la richesse. On remarque que le critère d'espérance de gain est le cas particulier obtenu pour la fonction $u(x) = x$.

La notion d'ordre stochastique permet de déterminer des conditions sur les fonctions de répartition de X et Y qui assurent que l'égalité (4) est vérifiée pour une classe de fonctions d'utilité donnée.

La dominance stochastique à l'ordre 1. Soient X et Y deux rendements aléatoires scalaires.

Définition 3.2 (Dominance stochastique à l'ordre 1). Le rendement X domine le rendement Y au sens de la dominance stochastique à l'ordre 1 ($X \geq_{\text{DS1}} Y$) si et seulement si

$$\psi_X(\eta) \leq \psi_Y(\eta), \quad \forall \eta \in \mathbb{R},$$

où ψ_Z désigne la fonction de répartition de la variable aléatoire Z , pour $Z = X, Y$.

La dominance stochastique à l'ordre 1 peut également être caractérisée à l'aide des fonctions d'utilité. En effet, X domine Y au sens de la dominance stochastique à l'ordre 1 si et seulement si X est préférée à Y par tout individu dont la fonction d'utilité est croissante (individu ayant une utilité marginale du revenu positive) :

Proposition 3.3. *On a l'équivalence suivante :*

$$X \succeq_{\text{DS1}} Y \Leftrightarrow \mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)], \quad \forall u : \mathbb{R} \rightarrow \mathbb{R} \text{ croissante.}$$

Un projet dont le rendement domine le rendement d'un autre projet au sens de la dominance stochastique à l'ordre 1 sera toujours sélectionné. La dominance stochastique à l'ordre 1 est un critère très fort, puisqu'il nécessite que les fonctions de répartition ne se coupent pas. Pour affaiblir cette condition, on peut considérer la dominance stochastique à un ordre supérieur, ou bien considérer d'autres critères tels que ceux présentés plus bas (Conditional Value-at-Risk, ...).

La dominance stochastique à l'ordre 2. X et Y désignent toujours des rendements aléatoires scalaires. Introduisons la fonction ψ_X^2 qui est la primitive de la fonction de répartition :

$$\begin{aligned} \psi_X^2 : \mathbb{R} &\rightarrow \mathbb{R}^+ \\ \eta &\mapsto \int_{-\infty}^{\eta} \psi_X(\xi) d\xi, \end{aligned}$$

continue, positive, croissante et convexe.

Définition 3.4 (Dominance stochastique à l'ordre 2). La dominance stochastique à l'ordre 2 (DS2) est définie par :

$$X \succeq_{\text{DS2}} Y \Leftrightarrow \psi_X^2(\eta) \leq \psi_Y^2(\eta), \quad \forall \eta \in \mathbb{R}.$$

On peut montrer que la dominance stochastique à l'ordre 1 implique la dominance stochastique à l'ordre 2.

Comme pour la dominance stochastique à l'ordre 1, les fonctions d'utilité permettent également de caractériser la dominance stochastique à l'ordre 2. On peut dire que X domine Y au sens de la dominance stochastique à l'ordre 2 si et seulement si X est préférée à Y par tout individu dont la fonction d'utilité est croissante et concave (individu averse au risque et ayant une utilité marginale du revenu positive) :

Proposition 3.5. *(Rothschild et Stiglitz, [42]) On a l'équivalence suivante :*

$$X \succeq_{\text{DS2}} Y \Leftrightarrow \mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)], \quad \forall u : \mathbb{R} \rightarrow \mathbb{R} \text{ continue, croissante et concave.}$$

Moments d'ordre 3 et fonctions d'utilité. Comme nous l'avons rappelé au début de la section, le critère moyenne/variance ne suffit pas

toujours à comparer deux variables aléatoires : si deux variables aléatoires ont même moyenne et même variance, ce critère ne donnera ainsi aucune indication sur la dispersion de la distribution de ces variables par rapport à un rendement de référence par exemple. Afin d'intégrer cette notion de dispersion de la distribution par rapport à un rendement de référence, il est utile de revenir à la théorie de l'utilité espérée et de considérer les moments d'ordre 3 de la fonction d'utilité. En effet dans [38] les auteurs donnent le résultat suivant :

Theorem 3.6. $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ pour toute fonction u telle que $u'''(x) > 0 \forall x \in \mathbb{R}$ si et seulement si les lois de X et Y différent et vérifient :

i) $\mathbb{E}(X) = \mathbb{E}(Y)$,

ii) $\int_0^1 \int_0^z [\psi_Y(y) - \psi_X(y)] dy dz = 0$,

iii) $\int_0^x \int_0^z [\psi_Y(y) - \psi_X(y)] dy dz \geq 0 \forall x \in [0, 1]$ et $> 0 \forall x \in]0, 1[$.

Lorsque les moyennes et les variances de deux variables aléatoires sont identiques, il semble donc utile d'examiner les moments d'ordre 3 des fonctions d'utilité associées afin de mieux capter certains risques (dissymétrie des distributions).

Remarque. En économie, il semble également pertinent de s'intéresser à la dominance stochastique à l'ordre 4, car l'ordre 4 permet de faire le lien entre l'épargne de précaution et l'aversion au risque.

Signalons enfin que dans [4], la notion d'ordre stochastique a été définie pour le cas de variables aléatoires n -dimensionnelles. Ceci peut être utile lorsque l'on doit comparer des variables aléatoires en fonction de plusieurs critères : dans une ACB, on peut par exemple vouloir prendre en compte un critère financier *et* un critère environnemental.

Dans la dernière section de ce document, nous introduisons d'autres mesures de risque issues du monde financier, qui vont permettre de renforcer le contrôle des risques financiers d'un projet d'investissement.

3.2 Mesures de risque et ACB

3.2.1. La Value-at-Risk (VaR)

Définition. La VaR d'un projet A correspond au niveau de VAN minimum, si l'on exclut un ensemble d'événements défavorables ayant une faible probabilité de se produire. Autrement dit, la VaR du projet associée à la VAN $J_A(\xi)$, notée $\text{VaR}_{J_A(\xi)}(p)$, est donc la plus grande

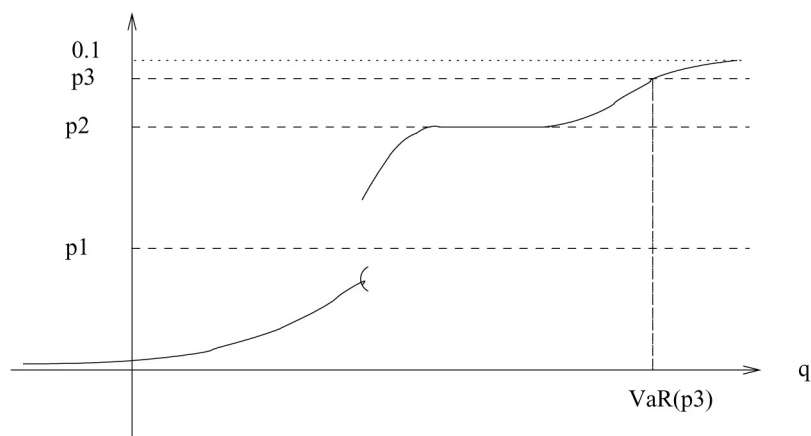
quantité q telle que, avec une probabilité inférieure ou égale à p , la VAN soit inférieure à q :

$$\text{VaR}_{J_A(\xi)}(p) = \max \{q \mid \mathbb{P} \{J_A(\xi) \leq q\} \leq p\}.$$

Notons $\psi_{J_A(\xi)}(q) = \mathbb{P} \{J_A(\xi) \leq q\}$. Lorsque $\psi_{J_A(\xi)}(\cdot)$ est continue et strictement croissante, $\text{VaR}_{J_A(\xi)}(p)$ est l'unique solution de l'équation $\psi_{J_A(\xi)}(q) = p$, ce qui correspond au niveau de probabilité $p3$ sur la Figure 2. Autrement, cette équation peut n'avoir aucune solution (lorsqu'il existe une masse atomique de probabilité, comme pour $p1$), ou au contraire avoir plusieurs solutions (lorsque la densité de probabilité est nulle pour une plage de valeurs, comme pour $p2$).

Figure 2

Value-at-Risk



En pratique, on choisit généralement $p = 0.05$ ou $p = 0.01$. Si la VaR à 5 % d'un projet vaut 100, cela signifie que, « en l'absence d'événements exceptionnels », il y a 95 % de chances que la VAN associée à ce projet soit supérieure à 100.

La VaR d'un projet permet donc d'exprimer en un seul chiffre le montant à risque de ce projet ; elle dépend de :

- la distribution de VAN du projet en fin de période,
- le niveau de confiance p ,
- la période de temps sur laquelle on désire mesurer la VaR.

Méthodes de calculs. Il existe plusieurs méthodes permettant de calculer la VaR : l'analyse historique, la méthode variance-covariance et la simulation de Monte Carlo.

L'analyse historique repose sur l'hypothèse que la distribution des rendements observée à partir des données historiques se reproduira à

l'avenir. Les avantages de cette approche sont d'une part que la mise en application est relativement aisée, et d'autre part, qu'il n'y a aucune hypothèse sur les lois de distribution. Mais cette approche nécessite beaucoup de données historiques, et suppose implicitement que le passé va se reproduire de façon statistiquement identique, ce qui est très discutable.

La *méthode variance-covariance* repose sur l'hypothèse que les rendements du projet et des facteurs de risque ont des distributions normales. Cette hypothèse n'est pas empiriquement justifiée, mais elle permet de simplifier énormément les calculs. Si l'application est aisée, le principal inconvénient de cette méthode est qu'elle sous-estime les événements rares, puisque l'hypothèse de normalité n'est pas empiriquement justifiée.

Enfin, la *simulation de Monte Carlo* est la plus rigoureuse des trois méthodes, mais elle requiert une puissance et un temps de calculs importants.

Les limites de la VaR. Si la VaR fournit un indicateur simple et global de l'exposition d'un projet au risque, elle admet cependant certaines limites.

Dans la mesure où il y a plusieurs méthodes de calcul de la VaR, il est important de se demander si la méthode de calcul choisie est adaptée au projet considéré. C'est ce qu'on appelle le *backtesting* : il s'agit de confronter la VaR calculée avec les rendements effectivement réalisés. Mais cet exercice semble difficile à réaliser lorsque l'on évalue un projet d'investissement en raison de l'horizon d'étude considéré, plusieurs années en général (notons qu'en finance la VaR est calculée à dix jours, ce qui rend l'exercice de backtesting plus pertinent).

Une autre limite de la VaR est directement liée au concept lui-même. En effet, celle-ci ne fournit aucune indication sur l'ampleur des pertes si un événement défavorable devait se produire. Pour remédier à cette limite, on peut s'intéresser à une mesure de risque proche de la VaR, mais qui capte mieux le problème des variations extrêmes : la *Conditional Value-at-Risk*.

3.2.2 La Conditional Value-at-Risk (CVaR)

Définition et propriétés. Par définition, la CVaR d'un projet A correspond à la VAN de ce projet calculée en moyenne sur les cas les plus défavorables. Elle s'écrit de la façon suivante :

$$CVaR_{JA(\xi)}(p) = \mathbb{E} [J_A(\xi) \mid J_A(\xi) \leq VaR_{JA(\xi)}(p)], \quad (5)$$

où $J_A(\xi)$ est une variable aléatoire (la VAN du projet par exemple) et $\text{VaR}_{J_A(\xi)}(p)$ la VaR du projet pour un niveau de confiance p (on choisit généralement $p = 0.05$ ou $p = 0.01$).

Remarque. Nous donnons ici la définition usuelle de la CVaR, valable pour des distributions de VAN continues. Pour des distributions discrètes, la définition de la CVaR est plus complexe car il s'agit d'une moyenne pondérée entre la VaR et l'espérance de la VAN strictement inférieure à la VaR.

La CVaR est caractérisée par la queue de la distribution de la variable aléatoire, et permet ainsi de contrôler la moyenne de cette variable aléatoire pour les pires scénarios. Elle possède également des propriétés intéressantes en pratique, puisqu'il s'agit d'une mesure de risque cohérente (ce qui n'est pas le cas de la VaR).

Remarque. Soit $(\Omega, \mathcal{F}, \mathcal{P})$ un espace de probabilité et \mathcal{X} un ensemble convexe de variables aléatoires de Ω dans \mathbb{R} . On suppose ici que l'on quantifie le risque d'une variable aléatoire représentant des gains. On rappelle qu'une mesure de risque ρ est dite cohérente si elle est :

1. monotone : si un projet rapporte toujours plus qu'un autre, alors il est moins risqué, *i.e.* si

$$X, Y \in \mathcal{X} \text{ et } Y(\eta) \geq X(\eta) \forall \eta \in \mathbb{R}, \text{ alors } \rho(X) \geq \rho(Y) ;$$

2. équivariante par translation : disposer d'une réserve de trésorerie pour le financement d'un projet diminue d'autant le risque de ce projet, autrement dit

$$\text{si } a \in \mathbb{R}^+ \text{ et } X \in \mathcal{X}, \text{ alors } \rho(X + a) = \rho(X) - a ;$$

3. positivement homogène : le risque est proportionnel à la taille du projet, *i.e.* si $b \in \mathbb{R}^+$ et $X \in \mathcal{X}$, alors $\rho(bX) = b\rho(X)$;

4. sous-additive : cette propriété garantit que la diversité diminue le risque, et concerne donc davantage la gestion de portefeuilles financiers :

$$\forall X, Y \in \mathcal{X}, \quad \rho(X + Y) \leq \rho(X) + \rho(Y) .$$

Si $\rho(X)$ mesure les risques associés aux rendements X d'un projet, on ne retiendra que les projets dont les mesures de risque sont élevées.

3.2.3 Downside risk measures

La VaR et la CVaR sont deux des mesures de risques les plus utilisées dans le domaine de la gestion des risques financiers, aussi bien par les banques que par les investisseurs institutionnels. Mais il existe également d'autres classes de mesures de risque, dont l'une d'elles (*downside risk measures*) nous semble particulièrement bien adaptée pour la gestion de projets.

Dans cette section, $J_A(\xi)$ désigne la variable aléatoire représentant la VAN d'un projet associée à la décision A et à la variable aléatoire ξ . Pour chaque décision A , $J_A(\xi)$ est une variable aléatoire ayant une distribution induite par celle de ξ . On suppose que J_A est strictement croissante. La distribution de probabilité sous-jacente de ξ sera supposée avoir une densité, que l'on notera $\phi_A(\xi)$. Les *downside risk measures* sont des mesures définies sur la notion de moments tronqués (introduite en 1977 par Bawa et Lindenberg) ; pour plus de détails sur cette notion, on pourra consulter [6].

Définition et propriétés. La théorie de *Downside Risk Aversion* ([7], [38]) permet de définir des mesures de risque qui sont basées sur les moments tronqués de la distribution de la VAN du projet d'investissement :

$$R_\gamma(\theta) = \mathbb{E} \left[(\max\{\theta - J_A(\xi), 0\})^\gamma \right] = \int_{-\infty}^{J_A^{-1}(\theta)} (\theta - J_A(\xi))^\gamma \phi_A(\xi) d(\xi), \quad (6)$$

avec $\gamma \geq 0$, θ un niveau de VAN de référence (la VAN moyenne du projet par exemple).

Remarque. Le cas où $\gamma \rightarrow \infty$ correspond au risque associé au pire niveau de VAN possible (voir [26] par exemple). Lorsque $\gamma = 0$, cette mesure de risque correspond à la probabilité d'être en dessous d'un certain seuil et on retrouve alors la VaR en choisissant θ (p) tel que $R_0(\theta(p)) = p$. Lorsque $\gamma = 1$ on retrouve l'*expected shortfall*. On montre facilement la relation suivante :

$$R_1(\theta) = p(\text{VaR}_{J_A(\xi)}(p) - \text{CVaR}_{J_A(\xi)}(p)).$$

Un investisseur qui cherche à maximiser l'espérance mathématique de son niveau d'utilité sous ce type de contrainte de risque va donc résoudre le problème d'optimisation suivant :

$$\max_A \left(\mathbb{E}[J_A(\xi)] - M \mathbb{E} \left[(\max\{\theta - J_A(\xi), 0\})^\gamma \right] \right), \quad (7)$$

où M est l'aversion au risque de l'investisseur. On peut alors déduire la classe de fonctions d'utilité associée aux *downside risk measures* :

$$u(x) = \begin{cases} x - M(\theta - x)^\gamma & \text{si } x \leq \theta, \\ x & \text{sinon,} \end{cases}$$

avec $M \geq 0$ (pour assurer que $u(\cdot)$ est croissante) et $\gamma \geq 0$. On se ramène ainsi à la théorie de l'utilité espérée en maximisant $\mathbb{E}[u(x)]$. En choisissant γ petit (resp. grand), l'investisseur s'intéresse aux petits (resp. grands) écarts de VAN en dessous du niveau de référence θ . Lorsque $\gamma < 1$, l'investisseur veut prendre des risques, alors que lorsque $\gamma > 1$, il est averse au risque. La fonction d'aversion relative au risque en dessous du seuil θ est :

$$ARR(x) = -\frac{x u''(x)}{u'(x)} = \frac{M x \gamma (\gamma - 1) (\theta - x)^{\gamma-2}}{1 + M \gamma (\theta - x)^{\gamma-1}}. \quad (8)$$

On observe que pour $1 < \gamma < 2$, l'aversion relative au risque est croissante, alors que pour $\gamma > 2$, cette aversion peut être croissante ou décroissante.

Application à la gestion de projet. Cette classe de mesures de risque nous semble particulièrement bien adaptée dans l'évaluation de projets d'investissements. En effet, considérons le cas où $\gamma = 1$; cette mesure de risque est alors appelée *expected shortfall*. Si l'on choisit comme niveau de référence la VAN de l'infrastructure existante (ce qui correspond à la décision de ne rien faire), on va pouvoir comparer différents projets en terme d'écarts moyens de leur VAN en dessous de la VAN de référence. Cette façon de mesurer le risque s'intègre ainsi facilement dans le cadre d'une ACB.

3.2.4. Ratio d'efficacité

Notons $J_A(\xi)$ une variable aléatoire désignant la VAN d'un projet A, et θ un niveau de VAN de référence (comme la VAN du projet existant). Nous proposons d'introduire le ratio d'efficacité suivant :

$$R_{eff} = \frac{\mathbb{E}[\max\{\theta - J_A(\xi), 0\}]}{\mathbb{E}[\max\{J_A(\xi) - \theta, 0\}]}.$$

Idéalement, un projet devrait avoir un ratio d'efficacité R_{eff} compris entre 0 et 1, autrement dit en moyenne et par rapport à une VAN de

référence donnée, les pertes seraient inférieures aux gains. Le cas où R_{eff} vaut 0 correspond à un projet dont la VAN moyenne est supérieure ou égale à la VAN de référence θ . Enfin, ce ratio permet également de comparer des projets risqués entre eux : plus le ratio R_{eff} est petit, plus le projet est performant.

3.3. Illustration numérique

3.3.1. Le cas de distributions discrètes

Considérons un projet d'investissement A à deux périodes, donnant pour chacun des quatre scénarios ζ_i envisagés les excédents bruts d'exploitation en millions d'euros suivants, avec $EBE0$ l'investissement initial, $EBE1$ l'excédent brut d'exploitation (EBE) à la fin de la première période et $EBE2$ l'EBE à la fin de la seconde période :

Table 4

EBE associés au projet

I	ζ_1	ζ_2	ζ_3	ζ_4
$EBE0$	-10	-10	-10	-10
$EBE1$	5	5	10	10
$EBE2$	7	9	11	15

Les probabilités associées à ces scénarios sont données dans le Tableau 5.

Table 5

Probabilités des scénarios

	ζ_1	ζ_2	ζ_3	ζ_4
probabilité p_{ζ_i}	0.05	0.20	0.5	0.25

On notera VAN_i la VAN du projet pour le scénario ζ_i . On prend comme taux d'actualisation $r = 3.5\%$. Les VAN du projet sont données dans le Tableau 6.

Table 6

VAN associées au projet

VAN_1	VAN_2	VAN_3	VAN_4
1.37	3.23	9.93	13.66

Supposons que la VAN de l'investissement existant (seuil de référence), notée θ , soit égale à 7.

L'*expected shortfall* de ce projet vaut :

$$R_1(\theta) = \sum_{i=1}^4 p_{\xi_i} \max\{\theta - VAN_i, 0\} = 1.04.$$

Ceci signifie donc que l'écart moyen entre la VAN du projet *I* et la VAN du projet actuel, lorsque cet écart est défavorable au projet *I*, sera de 1.04 million d'euros.

Le ratio d'efficacité pour ce projet vaut

$$R_{eff} = \frac{\sum_{i=1}^4 p_{\xi_i} \max\{\theta - VAN_i, 0\}}{\sum_{i=1}^4 p_{\xi_i} \max\{VAN_i - \theta, 0\}} = 0.33,$$

autrement dit, par rapport à l'investissement existant, ce projet générera en moyenne des rendements supérieurs aux pertes.

La distribution de ce projet étant une distribution discrète, on constate que la VaR à 5 % et la CVaR à 5 % sont toutes deux égales à 1.37 million d'euros.

Considérons maintenant un projet B ayant les caractéristiques suivantes :

Table 7

Probabilités des scénarios

	ξ_1	ξ_2	ξ_3	ξ_4
probabilité p_{ξ_i}	0.05	0.2	0.35	0.4

Table 8

EBE associés au projet

<i>I</i>	ξ_1	ξ_2	ξ_3	ξ_4
<i>EBE0</i>	-10	-10	-10	-10
<i>EBE1</i>	7	7	12	12
<i>EBE2</i>	4.54	10	5	13

Table 9

VAN associées au projet

VAN_1	VAN_2	VAN_3	VAN_4
1	6.10	6.26	13.73

L'*expected shortfall* de ce projet vaut 0.74 million d'euros, le ratio d'efficacité vaut 0.27.

Comparons ces deux projets en termes de VaR et de CVaR pour différents niveaux de probabilité :

Table 10
Comparaison des deux projets

projet	VaR à 5 %	CVaR à 5 %	VaR à 6 %	CVaR à 6 %	VaR à 10 %	CVaR à 10 %
A	1.37	1.37	3.23	1.46	3.23	1.83
B	1	1	6.10	1.26	6.10	2.27

On constate que la VaR et la CVaR à 5 % conduisent à retenir le projet A, la VaR à 6 % conduit à retenir le projet B alors que la CVaR à 6 % conduit à retenir le projet A, la VaR et la CVaR à 10 % conduisent à sélectionner le projet B. Ceci montre qu'en fonction du niveau de probabilité retenu, les critères de VaR et de CVaR ne conduiront pas à choisir les mêmes projets. On constate également que pour un niveau de probabilité de 5 %, la VaR et le ratio d'efficacité ne conduisent pas à sélectionner le même projet : la VaR à 5 % conduit à sélectionner le projet A alors que le ratio d'efficacité conduit à sélectionner le projet B. Le choix de la mesure de risque doit donc constituer une étape fondamentale de l'ACB.

3.3.2. Le cas de distributions continues

Considérons quatre nouveaux projets ayant des distributions de VAN continues :

- la VAN du projet A (VAN_A) suit une loi double exponentielle⁶ de paramètres $\eta = 2$ et $\mu = 1$; l'espérance de la VAN vaut alors 2.577 et sa variance 1.645 ;
- la VAN du projet B suit une loi double exponentielle symétrique ($VAN_B = 2 \mathbb{E}[VAN_A] - VAN_A$) ;
- la VAN du projet C suit une loi normale de même moyenne et de même variance ;
- la VAN du projet D suit une loi normale de même moyenne, mais de variance égale à 1.

La Figure 3 représente les densités de ces quatre projets ainsi que leur VaR respective. Sur le graphique représentant les VaR des diffé-

6. Pour mémoire, la fonction de répartition de la loi double exponentielle est $F(x) = e^{-e^{-\frac{(x-\eta)}{\mu}}}$.

rents projets, on constate que si l'on se fixe un niveau de probabilité inférieur à 4 %, le critère de la VaR conduit à retenir le projet A (loi double exponentielle), alors que si l'on adopte un niveau de probabilité supérieur à 4 %, ce même critère conduit à sélectionner le projet D (loi normale de variance 1).

La figure 4 représente les CVaR et expected shortfall de ces quatre projets. Sur le graphique représentant les CVaR, on constate que le critère de la CVaR conduit à retenir le projet A (ceci est vrai tant que le niveau de probabilité reste inférieur à 28 %). Pour le critère d'expected shortfall, le projet retenu dépend du niveau de VAN de référence (projet existant) : si le niveau de référence est inférieur à 2.5, on retient le projet B (loi double exponentielle symétrique), si le niveau de référence est compris entre 2.5 et 3.6 on retient le projet A (loi double exponentielle), enfin si le niveau de référence est supérieur à 3.6, on retient le projet C (loi normale de même variance).

En termes d'analyse de projet, l'utilisation de l'expected shortfall et de la CVaR permet donc d'une part de mesurer l'écart moyen entre la VAN du projet et un seuil de référence (la VAN de l'investissement actuel) lorsque cet écart est défavorable au projet ; et d'autre part de quantifier la VAN moyenne pour les pires scénarios. Le ratio d'efficacité permet quant à lui d'évaluer la performance du projet par rapport à un projet de référence donné.

Figure3
Densité et VaR

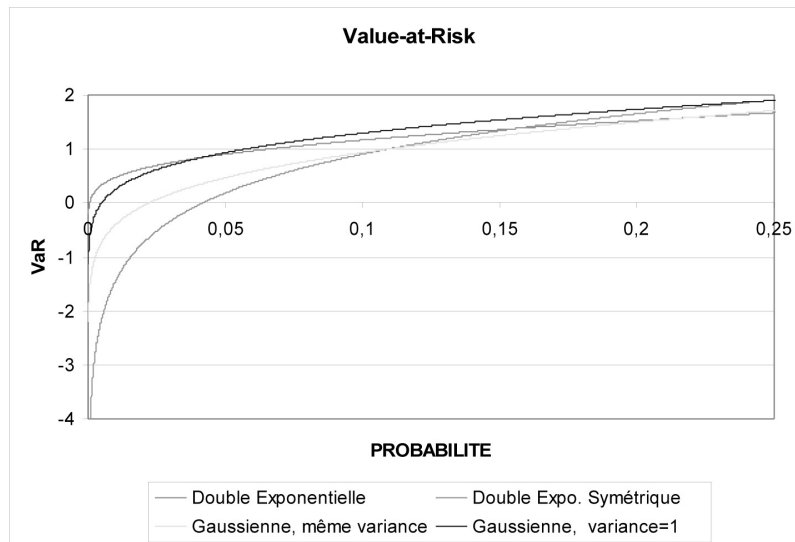
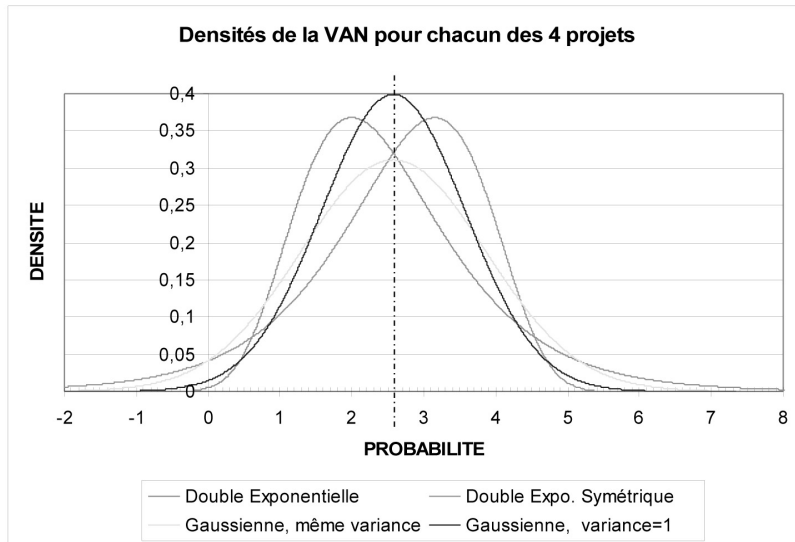
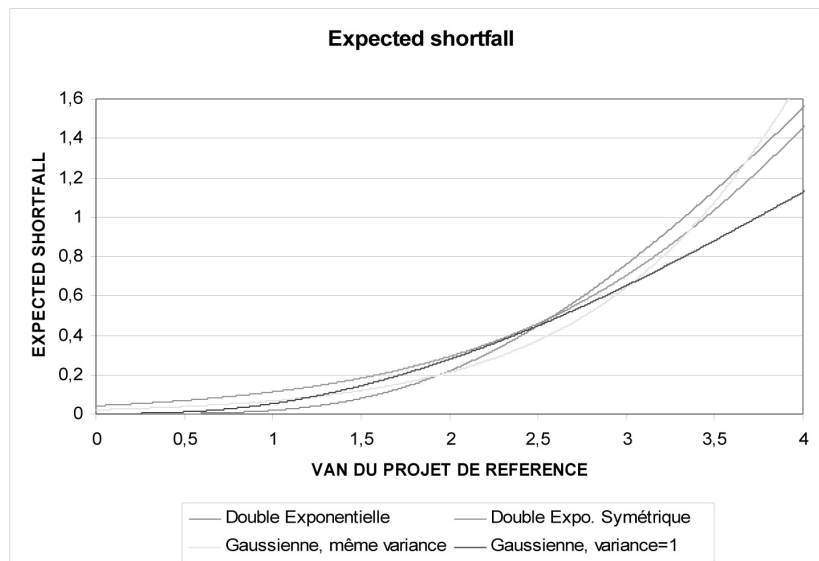
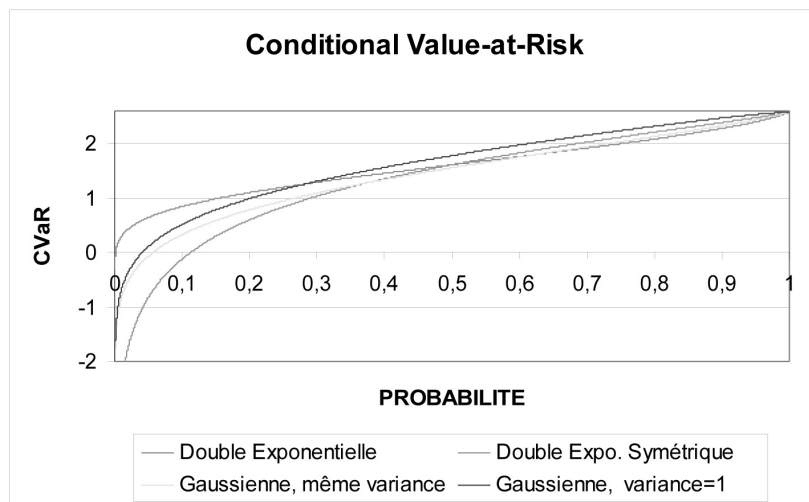


Figure 4
CVaR et Expected Shortfall



4. CONCLUSIONS, RECOMMANDATIONS ET PERSPECTIVES

Dans ce document, notre objectif était de réfléchir à la façon d'intégrer les notions de risque et d'incertitude dans l'analyse des coûts et des bénéfices. Après avoir rappelé les principes de cette analyse, nous avons présenté les techniques actuelles permettant d'intégrer ces deux notions dans l'analyse des projets. L'originalité de notre travail a ensuite été de proposer d'améliorer la gestion du risque liée à un projet d'investissement en intégrant aux critères usuels d'appréciation de ce projet trois mesures de risque supplémentaires (la VaR, la CVaR et l'expected shortfall) et un ratio d'efficacité. Ces propositions permettent de prendre en compte le risque d'une manière compatible avec la théorie économique néo-classique.

« En l'absence d'évènements exceptionnels », la VaR indique le niveau de VAN minimum d'un projet et l'expected shortfall mesure l'écart moyen de la VAN d'un projet par rapport à une VAN de référence lorsque cet écart est défavorable au projet. L'expected shortfall s'intègre bien dans une ACB puisqu'elle permet de comparer un projet donné à un projet alternatif. La CVaR permet de déterminer la moyenne de la VAN sur les plus petites valeurs de VAN d'un projet, et permet donc de mesurer les risques extrêmes. Enfin, le ratio d'efficacité permet d'évaluer la performance du projet par rapport à un projet de référence donné.

D'un point de vue opérationnel, ces indicateurs de risque se calculent grâce à une méthode de Monte Carlo :

- on commence par calculer la VAN (ou toute autre variable aléatoire que l'on souhaite contrôler) pour un scénario en générant un ou plusieurs aléas modélisant les différents risques affectant la VAN (risque géologique, risque de construction, de demande...),
- en répétant ce processus sur un grand nombre de scénarios on construit la distribution de la VAN recherchée ; le nombre de scénarios doit être suffisamment important pour que la distribution soit représentative (pour estimer le nombre de scénarios, on pourra se reporter à la convergence des algorithmes de type Monte Carlo),
- on calcule enfin à partir de cette distribution les indicateurs de risque retenus.

Certains critères de risque que nous proposons (VaR et CVaR) reposent sur des niveaux de confiance, ces niveaux de confiance correspondant à la préférence du décideur (fonction de l'aversion au risque de celui-ci). Nous avons ainsi mis en évidence que pour un critère de risque donné, la classification de projets dépend des niveaux de confiance sélectionnés. Il s'agit d'une conclusion fondamentale du

point de vue d'une ACB, car il peut exister des situations où les différents acteurs concernés (gestionnaires, actionnaires, institutions financières, Etat) possèdent des préférences différentes vis-à-vis du risque se traduisant par des niveaux de confiance différents. Il nous semblerait alors intéressant d'appliquer des méthodes économétriques pour déterminer ces niveaux de confiance.

Au terme de ce travail, nous recommandons dans le cadre d'une analyse de projets de faire successivement :

1. l'analyse des coûts et des bénéfices pour déterminer les taux de rendements des projets (VAN et TRI) et établir des comparaisons avec la situation antérieure, et éventuellement avec d'autres projets alternatifs,
2. l'analyse de sensibilité afin de bien identifier les variables aléatoires qui ont un réel impact sur les résultats attendus,
3. le calcul de mesures de risques et d'indicateurs d'efficacité pertinents pour bien évaluer les risques financiers.

Dans notre étude, nous n'avons pas proposé de traitement spécifique en fonction du type de risque considéré. Il pourrait ainsi être intéressant par la suite de proposer de traiter différemment les risques diversifiables de grandes amplitudes et les risques non diversifiables de petites amplitudes.

Enfin, une autre voie d'amélioration possible pour ce travail pourrait être de développer les propriétés économétriques de la demande ou des coûts. On pourrait ainsi essayer de faire le lien entre l'analyse économique et l'analyse économétrique des données utilisées dans les analyses coûts bénéfices.

Annexe : Flexibilité des politiques de prix

Considérons le cas d'une entreprise en condition de monopole avec demande linéaire. Nous avons :

$$\tilde{B} = -I + \sum_{t=1}^n \frac{1}{(1+i)^t} (\tilde{a} - b p_t)(p_t - \tilde{c}),$$

où I est l'investissement initial, i le taux d'actualisation, a et b se rapportent aux paramètres de la demande et c au coût marginal de production (supposé constant), avec :

$$\tilde{a}_t = \bar{a} + \varepsilon_a(t); \tilde{c}_t = \bar{c} + \varepsilon_c(t),$$

où ε_a et ε_c sont des variables aléatoires de moyenne nulle.

Considérons d'abord le cas où l'entreprise raisonne ex-ante et ignore la possibilité d'adaptation des prix. Cette entreprise a des prix rigides et fait face au problème suivant :

$$\mathbb{E}[\tilde{B}] = -I + (\bar{a} p_t - b p_t^2 - \langle a, c \rangle + b p_t \bar{c}) \sum_{t=1}^n \frac{1}{(1+i)^t}.$$

Conditions du premier ordre : $p^* = \frac{\bar{a} + b \bar{c}}{2b}$, et

$$\mathbb{E}[\tilde{B}] = -I + \mathbb{E}[(\tilde{a}_t - b p^*)(p^* - \tilde{c}_t)] \sum_{t=1}^n \frac{1}{(1+i)^t},$$

avec

$$\mathbb{E}[(\tilde{a}_t - b p^*)(p^* - \tilde{c}_t)] = \Omega = \frac{(\bar{a} - b \bar{c})^2}{4b} - \langle \varepsilon_a, \varepsilon_c \rangle.$$

Dans ce cas, l'entreprise perd lorsque les paramètres de la demande fluctuent au cours du temps. Par contre, l'entreprise flexible, qui optimise au début de chaque période (en supposant que la réalisation de la variable aléatoire soit connue) est : que la valeur moyenne du profit est : $p_t = \frac{a_t + b c_t}{2b}$ de sorte que la valeur moyenne du profit est :

$$\mathbb{E}[(a_t - b p_t)(p_t - c_t)] = \frac{(a_t - b c_t)^2}{4b} = \Phi = \frac{(\bar{a} - b \bar{c})^2}{4b} + \frac{\mathbb{E}[\varepsilon_a - b \varepsilon_c]^2}{4b},$$

où a_t et c_t sont les réalisations des variables aléatoires de \tilde{a}_t et \tilde{c}_t . Dans ce cas, l'entreprise flexible peut augmenter son niveau de profit. En effet, dans le cas certain de l'entreprise qui se trouve dans un environnement certain, nous avons :

$$\mathbb{E}[(a_t - b p_t)(p_t - c_t)] = \frac{(\bar{a} - b \bar{c})^2}{4b}.$$

Par conséquent, la perte de profit due à l'incertitude est égale à : $\langle \varepsilon_a, \varepsilon_c \rangle$ qui peut être positif si a et c sont corrélés positivement, ou négatif dans le cas contraire. Si les paramètres a et c ne sont pas corrélés, alors le profit est le même avec et sans incertitude, même lorsque l'entreprise ne s'adapte pas à l'incertitude.

Notons que l'entreprise qui est flexible possède un niveau de profit plus grand comme :

$$\Phi - \Omega = \frac{\mathbb{E}[\varepsilon_a + b \varepsilon_c]^2}{4b} > 0.$$

Le gain de profit apporté par la flexibilité est positif et égal à (si $\beta = 1/(1+i)$) :

$$\Delta \pi = \frac{\mathbb{E}[\varepsilon_a + b\varepsilon_c]^2}{4b} \sum_{t=1}^n \frac{1}{(1+i)^t} = \frac{(1-\beta^{n+1})}{1-\beta} \frac{\mathbb{E}[\varepsilon_a + b\varepsilon_c]^2}{4b}.$$

On peut ainsi donner un ordre de grandeur au gain de flexibilité. Supposons que $a = 2$, $b = 1$, $c = 1$ et que a et c sont des distributions uniformes sur $[1.5, 2.5]$ et $[2.5, 3.5]$, de variance 8.33 %. Supposons que les variables sont non corrélées (pour simplifier l'exposé), que la période est de 20 ans et que le taux $i = 8\%$, nous obtenons $\beta = 0.926$ et

$$\Delta \pi = \frac{(1 - 0.926^{21})}{1 - 0.926} \frac{(8.33\%)}{2}.$$

Notons que le profit de l'entreprise (le monopole) en situation de certitude est de $4/9$. La fraction relative de profit qui peut être recouvrée si l'entreprise s'ajuste aux fluctuations de la demande et des coûts représente un peu moins que 10 % de son profit.

Références

- [1] Adler H.A., *Economic Appraisal of Transport Projects : a Manual with Case Studies*, John Hopkins University Press, Baltimore, 1987.
- [2] Allais M., *Le calcul économique publique : bilan de quinze années de recherche et de perspectives*, revue d'Economie Politique, 2, 1989.
- [3] Arrow K.J. et Lind R.C., *Uncertainty and the Evaluation of Public Investment Decisions*, American Economic Review, n° 60, pp 364-378, 1970.
- [4] Baccelli F. et Brémaud P. : *Elements of Queuing Theory*, Springer, Second Edition, 2003.
- [5] Bassani A., et Pouyet J., *Strategic Choice of Financing Systems in Regulated and Interconnected Industries*, présentation à l'IDEI, Toulouse, 2003.
- [6] Bawa V. et Lindenberg E., *Capital Market Equilibrium in a Mean-Lower Partial Moment Framework*, Journal of Financial Economics 5, 189-200, 1977.
- [7] Berkelaar A. et Kouwenberg R., *Dynamic Asset Allocation and Downside-Risk Aversion*, Econometric Institute Report EI 2000-12/A, 2000.

-
- [8] Boiteux M., A propos de la critique de la théorie de l'actualisation, *Revue d'économie politique* 5, 1976.
- [9] Brean D. J.S. et Burgess D. F. , Risk-adjusted Discount Rates for Public Sector Investments with illustrations for transportation, mimeo, Université de Toronto.
- [10] Chateauneuf A., Cohen M. et Meilijson I., Four Notions of Mean Preserving Increase in Risk, Risk Attitudes and Applications to the Rank-Dependent Expected Utility Model, *Journal of Mathematical Economics*, 2004.
- [11] Chervel M., *Calcul économique public et planification : les méthodes d'évaluation de projets*, Publisud, Paris, 1987.
- [12] Cohen G. et Bernhard P. : On the Rationality of some Decision Rules in a Stochastic Environment, *IEEE Transactions on Automatical Control*, vol AC-24, no. 5, 1979.
- [13] Commission of European Communities, *Cost-Benefits and Multi-criteria Analysis for New Road Construction*, Euret Program, Brussels, 1992.
- [14] Cooper D., Grey S., Raymond G. et Walker P., *Risk Management Guidelines for Projects and Complex Procurements*, Wiley, 2005.
- [15] Dean J., *Capital Budgeting*, Columbia University Press, 1951.
- [16] Dasgupta P., *Environment Management under Uncertainty*, 1982.
- [17] de Palma A. et Proost S., *Evaluation financière des projets d'infrastructures*, Note d'étude pour la RFF, 2004.
- [18] de Palma A., Proost S., Van der Loo S. et Dunkerley F., *MOLINO II a model for assessing pricing and investment strategies for transport infrastructure with a case study*, livrable du projet FUNDING, DG TRENEN, KUL, Belgique 2006.
- [19] de Palma A., Lindsey R., Proost S. et Van der Loo S., *A Cost-Benefit Analysis of Tunnel Investment and Tolling Alternatives in Antwerp*, *European Transport*, 2005.
- [20] Dinwiddy C. et Teal F., *Principles of Cost-Benefits Analysis for Developing Countries*, Cambridge University Press, 1996.
- [21] Dorval C., *Choix des investissements*, *Techniques de l'Ingénieur*, traité L'entreprise industrielle, 1981.
- [22] Drèze J. et Stern N.H., *The Theory of Cost-Benefit Analysis*, *Handbook of Public Economics*, ed. Auerbach and Feldstein, North Holland, 1987.
- [23] Dupuit J., *De la mesure de l'utilité des travaux publics*, *Annales des Ponts et Chaussées*, (8), 1844.
- [24] ECMT, *Reforming Transport Taxes*, 2003.
- [25] European Conference of Ministries of Transport, *Evaluating Investment in Transport Infrastructure*, Paris, 1992.

- [26] Fishburn P., Mean-Risk Analysis with Risk Associated with Below-Target Returns, *American Economic Review*, 1977.
- [27] Flyvbjerg B., Bruzelius N. et Rothengatter W., *Megaprojects and Risk, An Anatomy of Ambition*, Cambridge University Press, 2003.
- [28] Glaister S., *Fundamentals of Transport Economics*, Blackwell, Oxford, 1981.
- [29] Commission européenne, *Guide de l'Analyse Coûts-Avantages des grands projets*, 1997.
- [30] Commission européenne, *Guide de l'Analyse Coûts-Avantages des projets d'investissement*, 2003.
- [31] Héline R. et Poupard-Lafarge O., *Principes et techniques des investissements, sélection et financement*, Delmas, 1975.
- [32] Kirpatrick C. et Weiss J. (eds), *Cost-Benefit Analysis and Project Appraisal in Developing Countries*, Elgar, Cheltenham, 1996.
- [33] Laffont J.-J. and Tirole J., *A Theory of Incentives in Procurements and Regulations*, MIT Press, 1993.
- [34] Layard R. et Glaiser S. (eds), *Cost-Benefit Analysis*, 2nd edition, Cambridge University Press, 1994.
- [35] Lesourne J., *Calcul économique et réflexion sur la prise en compte des risques*, 1972.
- [36] Lind R.C., Regressing the government's discount rate policy in the light of new theory and data in a world economy with a high degree of capital mobility, *Journal of Environmental Economics and Management*, 18, pp 8-28, 1988.
- [37] Massé P., *Le choix des investissements*, Dunod, 1968.
- [38] Menezes C., Geiss C. et Tressler J., Increasing Downside Risk, *American Economic Review*, Vol. 70, No. 5, 921-932, 1980.
- [39] Ministère de l'Équipement, des Transports et du Logement, *Instruction cadre relative aux méthodes d'évaluations économiques des grands projets d'infrastructures de transport* », 27 mars 2004 mise à jour le 27 mai 2005.
- [40] von Neuman J. et Morgenstern O., *Theory of Games and Economic Behavior*, 1953 edition, Princeton University Press.
- [41] Rockafellar R.T. et Uryasev, Conditional Value-at-Risk for General Loss Distributions, *Journal of Banking & Finance*, vol. 26, pp. 1443-1471, 2002.
- [42] Rothschild M. et Stiglitz J., Increasing Risk : I. A Definition, *Journal of Economic Theory* 2 (3), 225-243, 1970.
- [43] Salanié B., *The Economics of Contracts : A Primer*, MIT Press, 1997.

- [44] Shofield J.A., *Cost-Benefit Analysis in Urban and Regional Planning*, Allen & Unwin, London, 1989.
- [45] Wells G.R., *Highway Planning Techniques. The Balance of Cost and Benefit*, Griffin, London, 1971.

CHAPITRE 5

LA MESURE DE L'UTILITÉ SOCIALE DES INVESTISSEMENTS : L'ENJEU DU PROCESSUS DE PRODUCTION DES VALEURS TUTÉLAIRES

De l'expérience à la prospective

Luc Baumstark¹

Si une collectivité décide d'investir les ressources dont elle dispose dans un ouvrage ou un service, c'est parce que ceux-là produisent une certaine utilité sociale : un théâtre, un hôpital, une route, un commissariat de police, une université ont une utilité pour la collectivité. La difficulté dans la sphère publique, c'est que cette utilité n'a pas forcément de traduction marchande immédiate à opposer aux financements de ces ouvrages ou services. L'utilité économique est en général difficile à chiffrer et le processus de décision se contente bien souvent d'une évaluation « intuitive et politique » : le calcul économique se limite alors aux seuls calculs financiers qui pèsent considérablement dans le processus décisionnel. Cela est insuffisant pour apprécier l'efficacité de la décision : éviter de perdre en réparations ce qu'on croit avoir gagné par un investissement rentable mais polluant, réserver des moyens rares aux projets qui augmentent le plus le bien-être collectif, etc.

Ces questions ne sont pas nouvelles puisque Jules Dupuit écrivait déjà au milieu du XIX^e siècle :

« Le législateur a prescrit les formalités nécessaires pour que certains travaux puissent être déclarés d'utilité publique ; l'économie politique n'a pas encore défini d'une manière précise les conditions que ces travaux doivent remplir pour être réellement utiles ; du moins les idées qui

1. Laboratoire d'Économie des Transports, CNRS, Université Lyon 2, ENTPE.

ont été émises à ce sujet nous paraissent vagues, incomplètes et souvent inexactes. Cependant cette dernière question est plus importante que la première ; des enquêtes plus ou moins multipliées, des lois, des ordonnances ne feront pas qu'une route, un chemin de fer, un canal soient utiles, s'ils ne le sont pas réellement. La loi ne devrait, pour ainsi dire, que consacrer les faits démontrés par l'économie politique. Comment doit se faire cette démonstration ? Sur quelles données, sur quelle formule repose-t-elle ? Comment, en un mot, doit se mesurer l'utilité publique ? Tel sera l'objet de nos recherches ². »

Le calcul économique, technique d'évaluation conçue dès l'origine pour éclairer les décisions publiques, a toujours été contesté, et ce d'autant plus qu'il concerne des secteurs (culturel, santé, sécurité, éducation) qu'on souhaite écarter des logiques marchandes. On lui reproche notamment, et la critique est radicale, de véhiculer une problématique pseudo-scientifique reposant sur les fondements d'une théorie de l'utilité contestable et de renforcer un pouvoir technocratique en décalage avec les processus de décision.

Or, ces contestations récurrentes n'ont pas empêché les économistes de poursuivre leurs investigations. Sous l'impulsion des demandes des administrations et de nombreuses institutions leurs travaux s'étendent désormais à des domaines nouveaux, ce qui n'est pas sans exacerber les réticences. Par ailleurs l'administration française a engagé ces quinze dernières années des travaux importants pour préciser les méthodes utilisées dans les évaluations socio-économiques des investissements de transport. Un des enjeux principaux de ces travaux, qui s'inscrivent dans une longue tradition dans la sphère publique (administration et entreprises publiques) était de préciser les valeurs (prix fictifs) qui devaient permettre d'incorporer aux calculs économiques de rentabilité des investissements publics du secteur des transports, chaque fois que c'est possible, les coûts et les avantages non marchands associés au projet : l'épuisement des ressources non renouvelables, la congestion et les nuisances sonores liées au trafic, les effets de la pollution atmosphérique sur la santé, mais aussi les gains de temps et les vies épargnées etc. L'élaboration de ces valeurs a nécessité une large concertation et des compromis de la part des principaux acteurs représentés dans le groupe de travail. Le consensus, lent à obtenir sur certains dossiers, les études mobilisées pouvant conduire à des appréciations nuancées voire contradictoires, constitue néanmoins

2. J. Dupuit, « De la mesure de l'utilité des travaux publics », *Annales des Ponts et Chaussées*, 2^e semestre, n° 116, 1844, p. 332-375. Jules Dupuit, Inspecteur Général des Ponts et Chaussées, est considéré comme le fondateur du calcul économique public.

aujourd'hui, malgré les grandes incertitudes qui demeurent sur bien des points, une base solide sur laquelle il est possible d'appuyer le calcul économique public.

Il ne s'agit pas ici de refaire la théorie ou l'histoire du calcul économique ni même de revenir sur les outils théoriques et méthodologiques qu'il faut mobiliser pour donner un prix aux externalités produites par les investissements publics³. De très nombreux ouvrages et articles existent sur ce sujet.

Il s'agit de s'interroger sur le processus de production de ces valeurs de référence. L'analyse des modalités de production de ces valeurs, qu'on nomme parfois valeurs tutélaires, est ici entreprise au regard de la question de l'acceptabilité politique des instruments d'aide à la décision. Elle permet de s'interroger plus largement sur l'utilité sociale des institutions, comme le Plan (Centre d'Analyse Stratégique⁴), qui offrent un espace de concertation ouvert, et reconnu, dans lequel il est possible de produire des normes de référence, normes qui par la suite peuvent être plus facilement utilisées par les agents comme un langage commun partagé. La conviction qui sous-tend cette présentation est que l'appropriation finale de l'outil du calcul économique dans le processus de décision, y compris au stade du débat public, dépend de la qualité du processus de production des normes qui permettent d'appréhender l'utilité sociale des investissements publics.

Dans un premier temps, on rappelle les principaux arguments qui sont avancés pour expliquer la difficulté de mise en œuvre de cet outil et qui conduisent chercheurs et praticiens comme les institutions à le rejeter pour d'autres approches jugées plus pertinentes et plus opératoires. On présente ensuite dans un deuxième temps l'analyse des modalités de production des normes de calculs mises en œuvre au Commissariat Général du Plan. L'analyse de ce processus conduit dans un troisième temps à établir une série de recommandations qui, répondant sur certains points aux blocages constatés, sont susceptibles de favoriser l'utilisation de cet outil dans les processus de décision.

3. On trouvera de nombreuses explications dans les rapports du Plan (CGP, 2001 ; 2005) qui décrivent précisément les procédures par lesquelles ces valeurs ont été fixées, en insistant sur les éléments qui faisaient difficulté au sein du groupe de travail.

4. Le 27 octobre 2005, le Premier ministre annonçait la transformation du Commissariat Général du Plan. Le Centre d'analyse stratégique lui a succédé le 6 mars 2006. Le décret instituant sa création (décret n° 2006-260, *Journal Officiel* du 7 mars 2006) précise que ce nouveau Conseil exerce, sous l'autorité du Premier Ministre, des missions de veille, d'expertise et d'aide à la décision pour la conduite des politiques publiques. Afin d'éclairer le Gouvernement dans la définition et la mise en œuvre de ses orientations stratégiques en matière économique, sociale, environnementale ou culturelle, il élabore, soit à la demande du Premier ministre, soit de sa propre initiative dans le cadre d'un programme de travail annuel des rapports, recommandations et avis. <http://www.strategie.gouv.fr>

1. LA MESURE DE L'UTILITÉ SOCIALE DES INVESTISSEMENTS PUBLICS EN QUESTION

Le calcul économique est présenté dans la théorie économique comme un outil d'aide à la décision qui permet de rationaliser les dépenses publiques et de garantir une allocation optimale des ressources : dit autrement, et de manière un peu plus provocante, un tel exercice est censé dire l'intérêt général et défendre le « bien public » face au poids des intérêts particuliers et des groupes de pression. Au cours des années 60, dans la mouvance d'une « planification régénérée », le calcul économique considéré comme le symbole de l'effort de rationalisation de la dépense publique, était même devenu une référence incontournable. Or, en pratique, il s'avère qu'on observe des écarts importants entre les résultats de l'évaluation des projets, lorsqu'elle est entreprise, et les attentes des décideurs publics et, plus généralement encore, les attentes de la société. Les spécialistes et promoteurs de cet outil font lucidement ce constat.

« On aurait pu s'attendre à ce que le calcul économique, symbole de l'effort de rationalisation des choix, soit progressivement généralisé dans l'administration et l'entreprise et que son influence sur les décisions aille croissant. Trente ans après, force est de reconnaître qu'il n'en est rien : à une période faste a succédé une ère de déclin. De plus en plus de décisions échappent à son évaluation, et lorsqu'il est mis en œuvre, ses recommandations sont peu suivies d'effet ⁵. »

En 2005, le dernier rapport de l'administration sur ce sujet, le rapport de la commission présidée par Daniel Lebègue⁶ qui réévalue le taux d'actualisation, fait à son tour ce même constat et l'absence des

5. B. Walliser, E. Quinet, « Splendeur et misère du calcul économique », *Le Monde*, 18 décembre 1990

6. Le rapport Lebègue resitue la question du taux d'actualisation dans le cadre plus général de l'utilisation du calcul économique qu'il souhaite voir se développer. Les principales conclusions sur ce point sont :

- mieux maîtriser et mieux comprendre le calcul économique (formation, diffusion de la culture du calcul économique) ;
- étendre la pratique du calcul économique à tous les secteurs d'activités (transport, énergie, mais aussi santé, défense, recherche, etc.) ;
- recommander ce type d'évaluation aux collectivités territoriales (Région) ;
- se donner les moyens de l'enrichissement du calcul économique (recherche) ;
- renforcer ou créer les dispositifs de contre-expertise (évaluation des études, renforcement des évaluations *ex post*) ;
- pratiquer un calcul économique concerté (règles claires à établir secteur par secteur : rôle du Plan ; concertation européenne).

suites attendues pour y remédier (Etchegoyen, 2006) montre que le signal n'a pas été entendu⁷.

Le calcul économique suscite peu d'intérêt, ou lorsqu'il est évoqué, il se trouve au contraire très fortement contesté⁸ sur le fond, mais aussi, et souvent, dans les modalités pratiques de mise en œuvre. C'est ce qui a conduit dans les évaluations, dans le meilleur des cas, à s'orienter vers d'autres outils jugés plus efficaces.

Sans être exhaustif on peut rassembler les principales critiques autour de quatre points assez différents :

– Un calcul technocratique qui prétend dire l'intérêt général

Le calcul économique est très souvent considéré par ces opposants, comme technocratique et pseudo-scientifique : les résultats ne sont guère compréhensibles que par des techniciens chevronnés ; les calculs agrégés et peu lisibles et finalement peu transparents reposent sur des conventions qui sont souvent arbitraires et discutables. On lui reproche dès lors de pouvoir être manipulé et finalement de confisquer le débat. Pour beaucoup enfin, ces valeurs sont des valeurs de convenue qui présentent une part d'arbitraire souvent importante. Il y aurait même une imposture à considérer qu'elles sont des approximations de valeurs de référence qu'on chercherait à approcher⁹.

Ces calculs sont critiqués en raison même de l'objectif que poursuivent ceux qui les font et qui entendent infléchir les politiques suivies dans le sens de ce qu'ils considèrent comme rationnel. La mesure de la rentabilité, « le taux de rentabilité interne du projet », qui résume l'analyse et les calculs, est souvent ressentie comme un couperet incontournable et déterministe prétendant se substituer à la décision politique¹⁰.

7. On trouvera dans son ouvrage (Etchegoyen, 2005) *Votre devoir est de vous taire*, un chapitre sur le taux d'actualisation et le calcul économique dans lequel il relate et commente ce travail (p. 228-233).

8. Il existe une littérature critique abondante sur le calcul économique et la place toute relative qu'il occupe dans l'évaluation des projets d'investissements publics. On trouvera à titre d'illustration un bon résumé de ces critiques dans Bernard Roy, Sébastien Damart (Roy et Damart, 2001). Cet article n'est pas sans intérêt puisque l'auteur réagit particulièrement aux propositions faites dans le rapport du Commissariat Général du Plan de 2001 qui sera présenté plus bas. On se reportera aussi à la réflexion critique constructive que propose Olivier Godard (Godard, 2004) sur le rôle et les limites de l'analyse économique dans les évaluations économiques. On s'aperçoit que les critiques faites au calcul économique, comme outil d'évaluation, sont bien souvent les mêmes que celles qu'on fait plus généralement à l'analyse économique. Un des enjeux est donc bien la place de l'économiste dans le débat public.

9. « L'analyse coût avantage... ne peut fournir un point de départ objectif scientifiquement mis en évidence à partir duquel il faudrait justifier le fait de s'en écarter... n'est-ce pas utopique d'espérer mettre en évidence une solution voisine d'un optimum afin d'arrêter la décision en justifiant la façon dont elle s'écarte de ce prétendu optimum ? », *op. cit.*, Bernard Roy, p.14

10. La méthodologie traditionnelle conduit en effet à ramener un projet d'investissement à un taux de rentabilité sensé intégrer l'ensemble des avantages et inconvénients du projet. Selon cette

– Un calcul qui réduit l'utilité sociale à la seule sphère marchande

Par ailleurs, cet outil, notamment parce qu'il est basé sur une monétarisation systématique des avantages et des inconvénients pour lesquels il n'existe pas de contrepartie monétaire immédiate (externalités négatives ou positives), parce qu'il ne considère que l'aspect économique des choses et ignore l'ensemble de la politique dans lequel tel investissement est considéré, paraît également trop partiel, incapable de saisir les véritables enjeux des politiques publiques mises en œuvre, et notamment mal armé pour intégrer des considérations redistributives et d'équité ou les considérations d'aménagement du territoire. L'outil, bien adapté à la sphère marchande, semble incapable d'articuler des systèmes de références et des rationalités très éclatés et incommensurables, aussi bien politiques, éthiques et écologiques qu'économiques. Et même dans le cas où certaines de ces externalités peuvent être monétarisées plus facilement, leurs réels effets restent souvent mal appréciés y compris du seul fait de l'insuffisance des connaissances scientifiques (impact de la pollution sur la santé humaine, relations entre le niveau de CO₂ et le changement climatique, etc).

– Un calcul univoque qui fige le débat dans un processus de décision de plus en plus éclaté

Ce réductionnisme, pourtant incontournable du double point de vue scientifique et opératoire, semble ensuite contradictoire avec le souci actuel d'établir de véritables dialogues entre les décideurs, de plus en plus nombreux, et les citoyens¹¹. L'outil apparaît peu adapté à la réalité des processus de décision en vigueur dans lesquels la place de l'État central est contestée, contestation que la décentralisation des responsabilités et des financements d'une grande part des investissements publics accentue très fortement. Supposer en effet, comme c'est le cas dans ce cadre théorique, l'existence d'un décideur, – qualifié dans la théorie des choix publics de dictateur bienveillant et omniscient – placé au-dessus de la mêlée, agissant souverainement au mieux des intérêts bien compris de la collectivité, apparaît insoutenable dans la société actuelle dans laquelle les rapports d'autorité se dissolvent et

règle un investissement qui présente un taux de rentabilité interne supérieur au taux d'actualisation de la collectivité est un investissement profitable, susceptible de créer de la richesse, et qui, sous réserve de disposer des fonds pour le faire, devrait donc être réalisé. On trouvera une présentation complète de ce type d'approche dans un manuel classique comme celui de : Babusiaux D., *Décisions d'investissement et calcul économique dans l'entreprise*, Economica, 1990.

11. Les exigences de concertation se font de plus en plus forte en raison des fréquentes oppositions du citoyen aux décisions publiques. La concertation est inscrite dans le domaine législatif (LOTI, 1982 ; Circulaire Bianco, 1992 ; Loi Barnier, 1995, etc.) même si elle reste dans la pratique encore assez modeste.

les niveaux de décision se multiplient et où la concertation devient le maître mot d'une action réussie. L'approche mono critère basée sur une monétarisation systématique des avantages et des inconvénients ne semble pas adaptée pour permettre un véritable dialogue entre les décideurs au moment des choix.

– Un calcul qui n'a pas les moyens de ses ambitions

L'outil apparaît enfin trop décalé par rapport à l'expertise disponible et se trouve donc finalement impraticable. Les calculs, souvent coûteux et complexes à mettre en œuvre, posent un véritable problème de capacité d'expertise pour celui qui établit le calcul (données disponibles, formation nécessaire, coûts des études...). On conçoit que cet outil puisse être utilisé par les administrations centrales disposant de personnels nombreux et formés, les choses apparaissent plus délicates pour les collectivités territoriales dont le rôle dans les investissements publics s'accroît de plus en plus. Cette situation critique se renforce d'autant plus que l'on enrichit, que l'on complexifie, par ailleurs le calcul économique pour être davantage en phase avec les préoccupations de la société.

Malgré ces critiques très vives, le Commissariat Général du Plan, sollicité à de nombreuses reprises sur ce sujet, a toujours maintenu, sans aucune ambiguïté, la référence au calcul économique traditionnel alors que beaucoup abandonnait cette référence pour s'orienter vers d'autres outils d'évaluation notamment multicritères. Il rappelle que ce calcul n'est pas exclusif d'autres approches, notamment pour les éléments qu'il prend mal en compte, mais qu'il doit être entrepris.

« *L'intérêt de faire du bilan socio-économique, non le critère mais le noyau de l'estimation de la valeur d'un projet, c'est de permettre une analyse des raisons pour lesquelles on est conduit à s'écarter de la solution à laquelle ce seul bilan aurait conduit, et de pouvoir ainsi chiffrer le surcoût de la décision*¹². »

Dans ce cadre les valeurs susceptibles de prendre en compte l'utilité sociale dans les calculs apparaît essentielle. Ne pas chiffrer, c'est risquer de prendre des décisions absurdes ou inéquitables : cela revient à accepter que ces effets comptent pour zéro dans les calculs. C'est aussi empêcher les hommes politiques de connaître le coût de choix qu'ils

12. Rapport Boiteux (2001), *Transports : choix des investissements et coûts des nuisances*, p.16. Marcel Boiteux consacre un paragraphe de son avant propos sur ce sujet. Le rapport Lebègue (*op. cit.*) y revient également à de nombreuses reprises : « *À ces divers titres, le calcul économique apparaît donc, au regard des préoccupations d'efficacité et de bon usage des fonds publics, comme un instrument essentiel de cohérence à utiliser par les administrations tant pour l'ordonnement de leurs activités internes que dans leurs relations avec les autres administrations ou encore avec les collectivités territoriales* », p. 20.

opèrent au nom des convictions qui les ont fait élire. Tout chiffrer, c'est s'exposer tout autant à prendre des mesures perçues comme injustes ou immorales. C'est dans cette tension qu'il faut apprécier le processus de production de ces valeurs de référence.

2. LE SENS POLITIQUE D'UNE ÉLABORATION CONCERTÉE DES VALEURS TUTÉLAIRES : LE RÔLE PARTICULIER JOUÉ PAR LE COMMISSARIAT GÉNÉRAL DU PLAN

Il n'est pas inutile de préciser à ce stade que l'outil du calcul économique a profondément évolué avec les préoccupations de la collectivité, ce qui démontre qu'il n'est pas une technique figée et qu'il peut être en phase avec la prise de conscience collective de certains enjeux. Dans les années 70, les effets externes pris de fait en compte dans les calculs sont essentiellement ceux relatifs au développement massif de la mobilité (gains de temps et de sécurité). Dans les années 80, les préoccupations environnementales apparaissent plus nettement, et les méthodes d'évaluation intègrent alors certains de ces éléments (bruit, pollution atmosphérique). Les débats des années 90 se focalisent plus particulièrement sur l'effet de serre et les impacts sur le climat, on voit alors apparaître les premières tentatives de valorisation de la tonne de carbone. Il n'en demeure pas moins que d'autres préoccupations, notamment celles liées à l'aménagement du cadre de vie urbain restent mal appréhendées dans les méthodes d'évaluation traditionnelles.

Dans cette partie, il convient de rendre compte du processus qui a permis de passer de ces préoccupations plus ou moins diffuses à la production de valeurs normées qui sont sensées intégrer les préoccupations de la collectivité dans la mesure de l'utilité sociale des investissements.

– Une réflexion dans un contexte juridique et réglementaire de plus en plus exigeant

La réflexion du Plan a été menée dans un contexte juridique et réglementaire de plus en plus exigeant en matière d'évaluation. Le calcul économique occupe sinon dans les faits, du moins dans les textes et les discours, une place qui semble devoir être encore renforcée si l'on en juge par les nombreuses recherches soutenues par le ministère de l'Équipement pour améliorer ce type d'expertise.

La LOTI a institué plusieurs notions originales (droit au transport, efficacité économique et sociale, plan de déplacements urbains...) et a imposé dans son article 14 la réalisation d'évaluations socio-économiques et de bilans des grands projets d'infrastructure.

Le décret d'application de l'article 14 de la LOTI¹³ précise le contenu de l'évaluation et du bilan et impose que l'évaluation *a priori* comme le bilan *a posteriori* soit soumis à des obligations de publicité.

En 1996, la loi sur l'air et l'utilisation rationnelle de l'énergie¹⁴ est venue préciser notamment que les études d'impact des infrastructures de transport, qui font partie de l'évaluation *a priori*, comprennent une analyse des coûts collectifs des pollutions et nuisances et des avantages induits pour la collectivité.

La circulaire de la Direction des routes n° 98-99 du 20 octobre 1998 (concernant les méthodes d'évaluation économique des investissements routiers en rase campagne 1998) fixe de nouvelles règles pour l'évaluation socio-économique des investissements routiers, en cohérence avec le rapport Boiteux de 1994 et l'instruction cadre IDRAC de 1995¹⁵, actualisé en mars 2004, puis par l'instruction du 27 mai 2005 pour intégrer les conclusions du rapport Boiteux [2001] sur les valeurs environnementales ainsi que le changement du taux d'actualisation du Plan décidé en janvier 2005.

Toutefois, malgré cet arsenal réglementaire, les pratiques d'évaluation socio-économique restent encore très hétérogènes. C'est le constat que fait par exemple le CERTU¹⁶ après avoir analysé les dossiers de demande de subvention réalisés par les autorités organisatrices de transport pour les projets de transports collectifs en site propre (TCSP), dossiers qui sont instruits par les services de la direction des transports terrestres¹⁷. En milieu urbain où les interactions entre des projets de transport et le développement urbain sont très complexes, les évaluations socio-économiques des projets de transports collectifs en site propre n'ont pas de méthodologie éprouvée et n'ont pas fait l'objet à ce jour de guides de recommandations.

– Une réflexion sur le calcul économique qui prend naissance dans une dynamique de prospective

Les travaux qui vont être évoqués par la suite trouvent leur origine dans une vaste prospective dans le secteur des transports menée au Commissariat Général du Plan dans le début des années 1990 à la

13. Loi n° 82-1153 du 30 décembre 1982.

14. Loi n° 96-1236 du 30 décembre 1996.

15. Instruction cadre du 3 Octobre 1995 du Secrétaire d'État aux transports relative aux méthodes d'évaluation économique des grands projets d'infrastructure de transport.

16. Recommandations pour l'évaluation socio-économique des TCSP, CERTU, Centre d'études sur les réseaux, les transports l'urbanisme et les constructions publiques, ministère de l'Équipement, des Transport et du Logement, avril 2002.

17. Suivant les agglomérations concernées, il s'agit des demandes de prise en considération (DPC), des dossiers d'enquête préalable à la déclaration d'utilité publique (DUP).

demande de plusieurs ministres. Le groupe de travail « Transport 2010¹⁸ », s'était alors heurté à plusieurs questions vives portant notamment sur les méthodes d'évaluation des projets d'infrastructures de transport : il existait sur ce point au sein de l'administration française et chez les opérateurs de transport public, non seulement des pratiques différentes en matière d'évaluation, mais aussi des points de vue de doctrine assez contrastés. Ce rapport dressait le constat de la nécessaire obligation d'un travail interministériel approfondi sur ce point alors que la composante intermodale des projets de transports était de plus en plus prégnante dans les politiques, que les contraintes financières posaient le problème difficile des critères à retenir pour faire les arbitrages les plus pertinents pour la collectivité et surtout parce que le calcul économique (méthode coûts avantages) était fortement contesté au profit de méthodes dites multicritères.

Il convenait donc de faire le tri dans les éléments avancés par les uns ou les autres pour défendre ou contester l'intérêt de certains projets et justifier ou non les dépenses publiques : effet des infrastructures sur l'aménagement du territoire, impacts sur l'environnement et sur d'autres facteurs externes (sécurité, gain de temps pour les usagers, développement et compétitivité économique).

Affirmer l'importance d'une rationalisation des dépenses publiques en matière d'investissements supposait déjà de trouver un consensus sur les méthodes d'évaluation et de calcul de rentabilité des grands projets avec tous les intéressés (ministère de l'Équipement et des Transports, ministère de l'Économie et des Finances, ministère de l'Environnement, entreprises, etc.). C'est à la suite de ces réflexions qu'il a donc été demandé au Plan de poursuivre ses travaux dans ce sens.

– Un groupe de travail pour produire un consensus

Pour traiter ces questions très controversées il paraissait incontournable de trouver une personnalité maîtrisant à la fois les difficultés théoriques et pratiques du calcul économique et n'étant pas impliquée dans le secteur des transports. Aucun consensus ne semblait en effet

18. *Transports 2010*, Rapport du groupe présidé par le Commissaire au Plan, La Documentation Française, 1992. Voir aussi : « Transports : pour une cohérence stratégique », Atelier sur les orientations stratégiques de la politique des transports et leurs implications à moyen terme, présidé par Alain Bonnafous, Commissariat Général du Plan, septembre 1993. Les réflexions ont principalement porté sur les choix d'infrastructures de transport, avec l'objectif de renforcer au niveau national la cohérence de l'ensemble des investissements des opérateurs publics et de ceux prévus aux contrats de Plan État-régions. Plusieurs points ont fait l'objet d'examen critiques et de préconisations : les circuits de financement, la tarification d'usage des infrastructures, et enfin la question des méthodes d'évaluation des projets sur lesquelles les administrations avaient des points de vue divergents.

atteignable sans l'intervention d'une personnalité disposant sur ce sujet d'une autorité scientifique incontestable et qu'on ne pouvait soupçonner de défendre tel ou tel lobby du secteur. Marcel Boiteux¹⁹ disposait non seulement d'une autorité académique reconnue sur ces sujets mais aussi d'une expérience de la pratique du calcul économique hors du champ transport dans le secteur électrique²⁰.

En 1994, le Commissariat Général du Plan publiait les conclusions de ce groupe de travail dans un rapport sur la prise en compte de l'environnement et de la sécurité dans les choix d'investissements de transports, intitulé *Transports : pour un meilleur choix des investissements*²¹ : le « rapport Boiteux », qui devait rapidement devenir une référence dans ce domaine. Ce dernier prend le contre pied d'une tendance critique à l'égard du calcul économique standard qu'on discernait dans bien des pays et des institutions ainsi que dans les milieux universitaires et revient à une orthodoxie plus stricte du calcul économique. On y trouvera pour la première fois, des recommandations précisant la meilleure manière de prendre en compte les effets environnementaux. Des valeurs sont proposées pour les intégrer dans le calcul économique, elles seront très vite reprises dans de nombreuses enceintes administratives et universitaires. Ce rapport eut un retentissement assez considérable puisque, quelques mois après sa publication

19. Normalien, agrégé de mathématique et diplômé de l'Institut d'études politiques, Marcel Boiteux a été attaché au Centre national de la recherche scientifique (CNRS) (1946), avant d'accomplir, à partir de 1949, l'essentiel de sa carrière au sein d'Électricité de France (EDF.), dans laquelle il prendra les fonctions de directeur des études économiques à la direction générale (1958). En 1967, il est nommé directeur général de l'entreprise avant d'en devenir le président du conseil d'administration (1978-1987) puis le président d'honneur. Il mène parallèlement une activité d'enseignement en économie et des travaux de recherche fondamentaux et appliqués sur les questions relatives à la tarification et au calcul économique public. Depuis, il est membre de très nombreux conseils, comités et instituts. Il est élu membre de l'Académie des Sciences morales et politiques en 1992.

20. Il n'est toutefois pas inconnu du secteur des transports puisqu'il préside l'Instance d'Évaluation de la politique des transports de la Région Parisienne (1995-1999)

21. Commissariat Général du Plan, *Transport pour un meilleur choix des investissements*, Groupe de travail présidé par Marcel Boiteux, La Documentation Française, Paris, 1994. On retient de ce rapport le tableau qui propose pour chaque externalités des valeurs monétaires (valeurs du temps, vie humaine épargnée, bruit, effet de serre, pollution de l'air). Mais ce rapport ne se limite pas à ces quelques chiffres. Il rappelle notamment, ce qui n'allait pas de soi à l'époque, que le calcul économique reste, malgré ses insuffisances, ce qu'il y a de mieux pour évaluer les projets d'investissements, que les externalités négatives et positives doivent être incorporées aux calculs lorsque c'est possible et, dans le cas contraire, que les implications non monétarisables doivent être présentées sous la forme d'argumentaires. Déjà, le rapport évoque sans donner de précisions que l'évaluation doit prendre en compte les risques et les incertitudes et que les modèles de trafic, qui sont au cœur de l'évaluation, doivent être audités (cellule d'évaluation indépendante). Des précisions sont apportées sur des aspects plus techniques du calcul comme la difficile question la situation de référence à prendre en compte dans les évaluations, et des recommandations plus politiques sur la transparence des études. Le rapport recommande sur ce point qu'on procède à une normalisation des études.

officielle, le gouvernement émettait une instruction-cadre qui rendait effectives ses principales recommandations.

Quelques années plus tard, de nouvelles interrogations se posent sur ces méthodes notamment en raison des difficultés que rencontrent les évaluateurs pour mener l'analyse des projets de transport en milieu urbain. Le calcul économique se heurte effectivement dans la pratique au fait que les taux de rentabilité socio-économique des investissements des transports collectifs ne sont pas excellents au regard de certains investissements routiers. Cet écart est généralement expliqué par le fait que la valeur du temps écrase toutes les autres valeurs environnementales et que de ce fait pour prendre en compte les avantages environnementaux relatifs des modes de transport alternatif à la voiture particulière il devenait impératif de relever les valeurs environnementales utilisées dans les calculs sous peine de disqualifier l'outil du calcul économique lui-même.

C'est ainsi qu'en 1999, le Plan reçoit une demande conjointe des ministères des Transports et de l'Environnement pour réévaluer les valeurs notamment en milieu urbain. Ces valeurs devaient permettre d'incorporer aux calculs économiques de rentabilité des investissements publics du secteur des transports, notamment en milieu urbain où les méthodes traditionnelles semblaient peu adaptées, les avantages et les coûts non marchands que ces investissements impliquent : l'épuisement des ressources non renouvelables, la congestion et les nuisances sonores liées au trafic, les effets de la pollution atmosphérique sur la santé, mais aussi les gains de temps et les vies épargnées.

Un groupe de travail est constitué à nouveau sous la présidence de Marcel Boiteux qui accepte de reconduire les travaux de 1994²². Ce groupe de concertation a été constitué de manière à préserver certains équilibres entre les modes de transport d'une part et les différentes administrations d'autre part. Il réunissait pour ce faire fonctionnaires en charge des dossiers de transport, universitaires, représentants des grandes entreprises publiques (RATP, SNCF, etc.) et d'autres organismes comme ceux représentant les constructeurs automobiles par exemple.

Ce groupe qui a travaillé pendant près de 2 ans, n'avait pas pour objectif de faire ou même de commander des études particulières. Il s'agissait plus modestement, mais l'exercice était extrêmement difficile, de préciser et de décider les valeurs à retenir au regard des

22. Commissariat Général du Plan, *Transport : choix des investissements et coûts des nuisances*, Groupe présidé par Marcel Boiteux, La Documentation Française, Paris, 2001.

pratiques étrangères et compte tenu des principaux résultats obtenus par les études les plus récentes.

Pour avancer concrètement dans les travaux, le président a donc désigné un expert pour chacune des externalités environnementales en lui demandant de préparer une première réunion au cours de laquelle il devait faire le point sur l'ensemble des études menées depuis 5 ans sur ces sujets, et de proposer une série de recommandations concernant les valeurs et leur utilisation, propositions sur lesquelles le groupe de travail serait ensuite appelé à se prononcer. Ces travaux préparatoires ont été menés dans la plus totale liberté, le Commissariat Général du Plan offrant ses services (logistiques, rapporteurs), pour aider les présidents de séance à dégager les principales recommandations.

Ces réunions n'ont pas pu, pour la plupart, déboucher sur un consensus clair. Et pour les sujets les plus conflictuels, il a même fallu enclencher de nouvelles discussions regroupant les principaux protagonistes. Ces discussions bilatérales devaient permettre au rapporteur général et au président du groupe d'apprécier les points de vue, de mesurer la robustesse des argumentations développées en confrontant, si nécessaire, ces argumentaires à la critique d'autres experts.

Suite à ces différentes réunions, le Plan proposa au groupe pour chacun des chapitres une première proposition de rédaction. Celle-là faisait état des différends entre les membres du groupe en mettant en évidence les principaux enjeux du débat et les éléments d'incertitude. Il proposait ensuite une série de valeurs en prenant bien soin de définir les hypothèses qui avaient été retenues pour les fixer et de préciser les recommandations quant à leur utilisation.

L'ensemble de ces chapitres a fait ensuite l'objet d'une adoption collective, qui sur certains points a pu se terminer par un vote lorsque les différends étaient irréductibles. Pour les valeurs les plus controversées (la tonne de carbone, ou la valeur des impacts de la pollution atmosphérique sur la santé par exemple qui furent les points les plus discutés), le président disposait du pouvoir pour trancher les débats après avoir épuisé toutes les solutions de compromis possibles. La rédaction garde volontairement la trace de ces négociations. Les arbitrages ont été faits, c'était un impératif, ils ont été faits dans la plus totale transparence.

– Des valeurs produites pour les débats futurs

Le rapport transmis au gouvernement a fait l'objet de discussions et d'une expertise au Conseil Général des Ponts et Chaussées qui a proposé de retenir l'ensemble des recommandations. Ces valeurs ont ensuite été traduites dans une circulaire qui s'impose aux différentes

directions en charge de mener les évaluations des grands projets d'infrastructures. Ces valeurs, devenus des valeurs de référence, se trouvent dès aujourd'hui utilisées de fait dans des travaux académiques et administratifs ; elles correspondent à un besoin très manifeste. Le risque toutefois, c'est que les conditions de leur production et les recommandations qui les accompagnent soient oubliées. Le rapport précise leur caractère éphémère. Elles n'ont de sens du point de vue économique que si on se donne dans le même temps les moyens pour les contester et les affiner en permanence.

On notera notamment, que les accords obtenus au sein du groupe l'ont été bien souvent sous réserve d'engager des études complémentaires. Les débats ont montré à de nombreuses reprises que les études étaient généralement insuffisantes pour assurer certains arbitrages. Ces valeurs sont des valeurs tutélaires, mais qui, dans la mesure du possible, restent proches de ce qu'on pense savoir au regard des comportements des agents économiques ou des modèles de simulation lorsqu'elles ne sont pas directement accessibles. Le rapport propose pour ce faire toute une série d'études jugées prioritaires et un grand nombre d'entre elles font l'objet aujourd'hui de travaux dans le cadre d'un vaste programme de recherche qui mobilisent un grand nombre de chercheurs dans le secteur des transports.

Produire ces valeurs, c'est aussi engager un débat dans la société pour révéler le prix que collectivement nous sommes prêts à payer pour intégrer les considérations environnementales dans les investissements qui engagent l'avenir. On trouvera ci-dessous un exemple des valeurs produites par ce groupe de travail.

Tableau 1

**Valeur de la pollution atmosphérique
(euros par unités de trafic €/100.veh.km ; €/100.trains.km²³)**

<i>Valeur 2000 en véh-km</i>	Valeur moyenne	Urbain dense	Urbain diffus	Rase campagne
VP	0,9 €	2,9 €	1,0 €	0,1 €
PL	6,2 €	28,2 €	9,9 €	0,6 €
Bus	-	24,9 €	8,7 €	0,6 €
<i>Valeur 2000 en train-km</i>				
Train diesel (fret)	-	458 €	160 €	11 €
Train diesel (voyageurs)	-	164 €	57 €	4 €

23. Rapport Boiteux 2001, *op. cit.*, p. 138-139.

Plus récemment, en juin 2004, le Commissariat Général du Plan, à la demande du Premier ministre et dans le prolongement des décisions du Comité interministériel d'aménagement et de développement du territoire (CIADT, réunion du 18 décembre 2003), a engagé la révision du taux d'actualisation qui est depuis 1985 la référence utilisée dans l'évaluation socio-économique des projets d'investissement publics. Le taux d'actualisation public – on le qualifie dans la littérature de taux d'impatience – est un élément clef du calcul économique public qui permet de rendre comparable dans les calculs les avantages et les coûts d'un projet qui s'échelonnent dans le temps. Ce taux, le prix accordé au temps, dont la dernière valeur avait été fixée en 1986, faisait en effet, et ce depuis plusieurs années, l'objet de contestations assez fortes notamment parce que son usage conduit mécaniquement à minorer considérablement ce qui se passe dans le futur éloigné. La décision de sa révision a été prise suite au rapport d'audit sur les grandes infrastructures de transport effectué par l'Inspection générale des Finances et le Conseil général des Ponts et Chaussées début 2003, qui conduisait à rejeter bon nombre de projets d'investissements dont la rentabilité était jugée insuffisante. Les élus ont eu alors le sentiment que ces conclusions, dont on expliquait qu'elles étaient liées en partie à l'utilisation d'un taux d'actualisation très élevé, remettaient en cause leur propre légitimité à décider des grandes orientations en matière d'investissement, alors que celles-là pouvaient se justifier au regard d'autres critères. Le taux d'actualisation s'est ainsi trouvé au centre d'une polémique qui imposait dès lors d'engager une révision maintes fois repoussée. Ne pas réviser le taux risquait en effet de fragiliser le calcul économique.

En produisant ces différentes valeurs de référence et en s'engageant dans l'exercice de révision du taux d'actualisation, le Plan engage une démarche bien particulière. La commission composée principalement d'experts et de hauts fonctionnaires avait pour tâche délicate de produire ce qu'on appelle des valeurs tutélaires, des valeurs de référence. Cette production apparaît paradoxale puisqu'elle est fondamentalement motivée par la volonté de rendre le calcul économique plus performant, plus en phase avec les préoccupations de la société, et donc plus acceptable politiquement, alors que d'une certaine manière, ce type de production semble renforcer son caractère arbitraire et donc ce qui motive son rejet. La révision du taux d'actualisation, ou encore la production des valeurs de la vie humaine ou de la tonne de carbone réalisée en 2000 par exemple, a impliqué, à plusieurs reprises, des arbitrages sur des questions qui relèvent bien souvent de choix de société : contrainte budgétaire, prise en compte des risques, des enjeux environnementaux de court et de long terme, voire de très

long terme, niveau de solidarité à maintenir entre les générations, considérations relatives à l'aménagement du territoire et à la cohésion sociale, etc.

3. LE PROCESSUS DE PRODUCTION DES VALEURS TUTÉLAIRES DE RÉFÉRENCES DANS LA CONSTRUCTION D'UNE NOUVELLE GRAMMAIRE DU CALCUL ÉCONOMIQUE

La description du mécanisme qui a présidé à l'élaboration de ces références fait apparaître plusieurs éléments fondamentaux.

– La construction d'un langage commun

La production de ce système de prix relatif permet de traduire en langage commun des préoccupations très diverses, de confronter les discours tenus et les argumentaires aux raretés du monde dans lequel nous vivons et obligent à s'interroger sur les arbitrages que fait la collectivité. C'est dans les processus longs et complexes qui mènent aux grandes décisions que le calcul économique s'insère comme un éducateur du jugement. Le calcul économique est ainsi présenté comme facilitant la discussion en groupe. Le calcul économique oblige à parler le même langage et contraint à présenter les arguments dans un cadre logique et normalisé. Dans ce processus, les valeurs tutélaires qui ont été fixées sont objectives au sens où elles ne sont pas construites consciemment par les groupes de personnes concernés par le projet dont on évalue l'intérêt. Ces valeurs donnent un cadre reconnu par l'ensemble des parties pour les débats et les échanges.

Ces valeurs ne sont pas pour autant reconnues comme telles. Pour l'être, elles doivent passer par un processus de production très particulier qu'il convient d'analyser.

– Un processus de production en concertation²⁴

Les travaux menés par les groupes de travail ont permis de passer sur chacun des points importants de l'évaluation d'une multiplicité de points de vue à un seul. C'est l'objectif fondamental qui est recherché par un travail dont la nature collective est décisive. Ce processus de

24. Dans une situation d'aide à la décision, le mot concertation s'applique à un type de conception du dialogue entre acteurs (et/ou groupes) qui vise à progresser vers une compréhension claire et commune des positions de chacun des acteurs (et/ou groupes) sur la question posée, l'objectif de ce dialogue ou de cette réflexion devant être de conduire une décision et/ou de choisir une façon de formuler et de résoudre un problème spécifique autant que possible d'une manière consensuelle, Roy, in Gal, 1999, p. 1-6, Gal T., Stewart T.J., Hanne T. (eds), 1999, *Multicriteria decision making – Advances in MCDM models, algorithms, theory, and applications*, Kluwer Academic Publishers, Boston.

convergence qui n'est pas simple à obtenir demande du temps : permettre à chaque membre du groupe de s'approprier la connaissance suffisante sur des sujets qu'il ne maîtrise pas forcément au départ et qui sont souvent polémiques pour pouvoir véritablement entrer par la suite dans un jeu de réflexion collective. Le processus de travail doit permettre de rassembler les informations souvent éparses, de focaliser l'attention et le débat sur les principaux enjeux en faisant la part des choses entre ce qui est essentiel et ce qui reste secondaire, entre ce qui fait l'objet d'un consensus et ce qui est moins robuste et plus discutable, puis, ce qui constitue l'objectif essentiel de la démarche, de trancher. Ce processus de discussion²⁵ permet ainsi de passer de préoccupations souvent générales à une règle opératoire, ou à une valeur monétaire, directement utilisable dans les évaluations socio-économiques menées de manière totalement décentralisée. Plusieurs phases de maturation sont à distinguer.

Dans une première phase, les synthèses réalisées par les équipes de préparation ont permis d'avoir une réflexion plus générale sur les méthodologies réalisées en France à l'étranger et sur leurs résultats. Elles ont également fait apparaître les inévitables limites de ces exercices qui tiennent aussi bien à la complexité des phénomènes observés, aux multiples composantes de la nuisance étudiée (économique, physiologique, psychologique, sanitaire), qu'à l'hétérogénéité des unités et des modalités de mesure, aux incertitudes des relations causales, à la variabilité des impacts en fonction des situations géographiques ou topographiques, à la dispersion des appréciations individuelles et collectives des nuisances en fonction des conditions de vie des intéressés, de leurs revenus, de leurs caractéristiques sociales ou individuelles.

Le travail de convergence passe alors à une deuxième phase de concertation et de négociations plus difficiles. Comment passer d'un état de l'art plus ou moins décisif selon les thèmes à des valeurs précises. Les méthodes qui permettent de produire des valeurs monétaires diffèrent d'une externalité à une autre en raison des spécificités de chacune d'entre elles : on conçoit qu'on ne monétarise pas de la même manière la tonne de carbone dans 20 ans, qu'un gain de temps ou l'impact sur la mortalité d'une augmentation d'un certain niveau des particules dans l'atmosphère. Les méthodes diffèrent et chacune d'entre elles ont des inconvénients et des avantages. Certains protoco-

25. Ce travail est considérable. Pour le rapport Boiteux de 2001 c'est plus d'une centaine de personnalités d'horizons divers qui ont été rassemblées, des milliers de mails échangés, deux années de travail, de nombreuses réunions plénières, des sous groupes de travail, des auditions d'experts, sans compter toutes les réunions informelles et les concertations au sein même des institutions partie-prenantes du processus.

les d'enquêtes sont très développés, d'autres beaucoup moins. Bien évidemment les valeurs sont très variées et la synthèse paraît difficile. Là où les enjeux sont importants, les négociations deviennent difficiles, les uns et les autres étant tentés de refuser l'utilisation de telle ou telle étude jugée insuffisante, trop partielle.

La valeur retenue *in fine* est, par définition, un compromis entre des approches qui peuvent être très différentes (études particulières, dire d'experts, etc.). Ces valeurs mêlent tout à la fois des éléments scientifiques mais aussi d'autres considérations. La décision peut parfois être directement tirée d'une étude particulière, mais elle peut tout aussi bien, et c'est souvent le cas, résulter d'un simple calcul de moyenne entre les extrêmes d'une fourchette voire d'un vote entre plusieurs options. L'enjeu de ce processus est donc bien de produire un compromis qui doit nécessairement être obtenu alors que certains points restent et resteront dans l'obscurité.

– Un processus de production dynamique : des valeurs à titre conservatoire

Dans ce processus de production, les membres de ces groupes ont dû à chaque fois se déterminer sur la base de données et d'études, qui restaient insuffisantes. Une fois les arguments présentés, les connaissances disponibles exposées, il fallait trancher. Marcel Boiteux a synthétisé cette situation par une formule très évocatrice : « il faut se jeter à l'eau ». Ce processus de concertation a pu ainsi se terminer par un vote lorsque le consensus tardait à se faire ce qui a fait réagir assez violemment certains scientifiques qui ne comprenaient pas qu'on puisse mettre leurs travaux aux enchères en modulant plus ou moins des résultats et en les simplifiant à l'extrême.

Décider, dans l'incertitude, dans le compromis... Il s'agit là d'un impératif. Si en raison de ces quelques incertitudes on renonce à ces calculs, certaines nuisances, la valeur de certains effets, continueront d'être écartés des bilans, et seront donc comptées pour zéro. Et là où certains groupes d'intérêts y trouveront avantage, rien ne les empêchera de continuer à introduire dans leurs calculs des valeurs très élevées alors qu'on se contentera ailleurs de valeurs beaucoup plus faibles. Cette situation a conduit à proposer des valeurs à titre conservatoire sous contrainte que les pouvoirs publics engagent dans des délais assez brefs un ensemble de travaux pour conforter ou, le cas échéant, modifier les premiers compromis de ce rapport²⁶.

26. On notera par exemple que le rapport Boiteux (2001) prend soin de récapituler dans le dernier chapitre (et non dans une annexe) l'ensemble des études et recherches qui semblait

Ce processus apparaît vertueux. Les membres du groupe ne prétendaient pas construire de nouvelles valeurs scientifiques, ni même saisir une vérité qui reste insaisissable, mais afficher une valeur de référence la plus consensuelle possible au regard des connaissances disponibles. Dans ce processus de négociation qui s'instaurait, chacun était invité à objectiver sa contestation.

La qualité de ce compromis, et donc son acceptabilité sociale, réside aussi dans la possibilité de le contester. Pour que cette contestation puisse être crédible il est nécessaire de prévoir de manière claire l'organisation de révisions régulières. Ceci suppose donc une permanence de l'institution qui les produit. Il est souvent rappelé dans les rapports du Plan que ces chiffrages n'ont pas vocation à s'éterniser. La construction de ces prix non marchands doit donc être appréciée, à l'image de la formation des prix sur les marchés, comme une étape d'un processus d'erreurs et de corrections successives. Chaque étape est franchie en fonction de l'état des connaissances et des prises de conscience de la collectivité. La qualité de ces valeurs tutélaires dépend donc de la dynamique dans laquelle s'inscrit leur production. Un premier prix a été fixé. Il doit s'imposer à tous. Mais la contestation reste ouverte. Il s'agit là d'un processus positif, fécond qui invite au débat.

– Une valorisation au plus près des préférences des individus

On trouve dans les différents rapports du Plan l'expression de « valeur tutélaire ». Mais alors que cette notion s'appliquait dans le passé à une valorisation que l'État, dans sa sagesse, promulguait pour le bien de ses administrés sans avoir nécessairement à épouser leur point de vue, il s'agit ici d'une valeur que les groupes de travail se sont efforcés de fixer en analysant le comportement des gens (valeurs révélées) ou leurs réponses à des enquêtes (valeurs déclarées). On retrouve ici un problème très classique en économie publique sur la fonction d'utilité collective. Les arguments de cette fonction s'appuient-ils sur les seules fonctions de satisfaction ou intègre-elle des considérations qui transcendent les individus eux mêmes. Les derniers travaux ont pris l'option, chaque fois que cela était possible, de fonder le processus de production de ces valeurs sur la base des observations des compor-

nécessaire d'entreprendre pour fonder davantage les décisions qui avaient été prises : « *Le groupe a fréquemment conclu que l'état des données existantes sur les différents effets externes du transport, particulièrement en milieu urbain, ne suffisait pas pour établir solidement un barème de valeurs à retenir dans les évaluations socio-économiques. Chaque fois que cela a été possible, le groupe a proposé des valeurs a priori. Mais, ces valeurs ont été prises à titre conservatoire et il importe que les pouvoirs publics engagent dans des délais assez brefs un ensemble de travaux pour conforter ou, le cas échéant, modifier ces premiers compromis.* »

tements pour faire en sorte que les valeurs de référence ne soient pas en totale déconnexion avec ce que révèlent les comportements des agents économiques.

Ces études, qui ont joué un grand rôle pour la définition de certaines valeurs, conduisent à des ordres de grandeur ou à des fourchettes, mais pas à des valeurs précises. L'État intervient alors pour en normaliser les résultats et pour faire en sorte que tous les intéressés utilisent dans les évaluations, jusqu'à nouvel ordre, la même valeur. Ledit État ne cherche donc pas à se placer au-dessus de ce que révèle l'étude des comportements et des opinions des gens, il normalise les résultats de ces analyses. De ce fait ces valeurs ne tirent pas leur légitimité du caractère scientifique qu'on peut leur donner, même si les résultats de nombreux travaux et recherches ont été nécessaires pour les produire, mais plutôt de la procédure de concertation à laquelle se sont associés les mondes académique et administratif. Ces valeurs font donc autorité mais ne sont pas arbitraires.

Cette approche se heurte à certaines limites qui obligent de composer avec cette règle. On peut en effet considérer que dans bien des cas les agents économiques ont une appréciation erronée, insuffisante, des effets externes qu'on cherche à intégrer dans les calculs et que de ce fait, il est nécessaire de les corriger. C'est par exemple une des justifications qui a incité le groupe de travail à ne pas fonder strictement le nouveau taux d'actualisation sur les taux d'intérêts réels de long terme, même si la décision finale a intégré cette information.

– L'institutionnalisation du consensus

La reconnaissance de ces valeurs et finalement leur utilité dans le processus d'arbitrage dépend en grande partie de la légitimité de l'institution qui a porté ce travail, des personnalités qui se sont impliquées à des titres divers dans la réflexion collective, de la diversité des membres du groupe. Il paraît difficile d'imaginer réaliser un tel processus de concertation et de compromis à une échelle plus vaste.

Malgré tout, l'élaboration des valeurs monétaires se heurte en permanence à la question de la légitimité de celui qui les produit. Certains ont pu ainsi estimer que le Commissariat Général du Plan était un camp retranché de technocrates défendant une orthodoxie économique décalée par rapport aux enjeux de la société. D'autres estiment au contraire qu'il est un creuset dans lequel s'élaborent de manière concertée, à échelle réduite, les valeurs qui seront intégrées par la suite dans les calculs publics permettant d'apprécier l'utilité sociale des investissements. Il est toutefois significatif d'observer la rapidité avec laquelle ces valeurs ont fait référence dans de très nombreux travaux y

compris à l'extérieur du cadre administratif et ce sans aucune obligation particulière. Même si ces valeurs ne sont pas celles que retiennent les promoteurs des études, ils se sentent dans l'obligation de se positionner par rapport à ce cadre de référence. Et cela constitue une avancée très significative dans l'objectivation de l'utilité sociale de telle ou telle décision.

Par ailleurs, on notera que, dans la procédure actuelle ces valeurs qui ont été définies à travers toute une procédure de confrontations d'experts, de débats entre ministères représentant des intérêts très différents, restent des propositions faites au gouvernement. Le passage entre ces valeurs et leur utilisation effective au sein de l'administration passe donc par une validation politique. Certains considèrent que, compte tenu des impacts que ces valeurs ont sur les résultats des évaluations, cette transmission et une validation au sein d'un ministère dans le cadre d'une circulaire restent insuffisantes pour surmonter les critiques rappelées plus haut et estiment que la légitimité de ces travaux se trouverait renforcée s'ils faisaient l'objet d'une discussion au sein d'une commission parlementaire puis d'un vote au Parlement : cela apporterait une reconnaissance claire au compromis établi dans un groupe restreint et permettrait de dépasser la multiplicité des argumentaires.

4. CONCLUSION : AU-DELÀ DES INCANTATIONS ET DES LAMENTATIONS

L'utilité sociale du calcul économique, qui est souvent contestée, dépend en grande partie de l'usage qui peut en être fait et surtout de son appropriation par les acteurs eux-mêmes dans le cadre des discussions et des compromis présidant à la décision. L'appropriation collective de ce référentiel passe par une procédure dans laquelle les pouvoirs publics se trouvent nécessairement impliqués. Les travaux menés par le Commissariat Général du Plan constituent une expérimentation intéressante de ce point de vue. La qualité du processus de production de ce cadre référentiel, véritable bien collectif, représente un enjeu important pour améliorer l'organisation du débat sur l'utilité sociale des dépenses publiques.

Le dernier rapport du Plan concernant le calcul économique établit un diagnostic sévère sur la capacité des pouvoirs publics à procéder effectivement aux mesures de l'utilité sociale attendue des décisions d'investissement. Le groupe de travail, encouragé par les propos souvent très acerbes des experts et hauts fonctionnaires sur la situation actuelle, a réitéré avec force – c'est son message principal – des recom-

mandations anciennes et récurrentes sur l'importance de l'usage du calcul économique dans la sphère publique, constatant malheureusement que bien souvent ces travaux ne sont pas réalisés et ne jouent pas le rôle qui devrait être le leur. Cela peut s'interpréter de plusieurs façons :

– Difficulté de ressources, c'est-à-dire à trouver des personnes pour réaliser ces études qui supposent une formation adéquate. Ces difficultés sont d'autant plus difficiles que l'administration concernée ne dispose pas de moyens d'expertise suffisants pour mener ces travaux.

– Difficulté opératoire pour mettre en œuvre ces recommandations sur la mesure de l'utilité sociale qui restent dès lors purement incantatoires. Ces problèmes opératoires sont multiples : comment produire et utiliser les valeurs susceptibles de mesurer certains effets non marchands qui légitiment généralement la décision publique ? Comment offrir une réponse concrète aux responsables de secteurs qui ne disposent ni d'études normées et validées spécifiques ni de pratiques systématiques ou d'un savoir faire sur lesquels s'appuyer, comme cela peut exister dans le secteur des transports ?

– Difficulté d'opportunité, c'est-à-dire difficulté à percevoir l'utilité réelle de ces études alors que bien souvent le calcul économique pèse peu dans la décision finale, difficultés qui traduisent d'une certaine manière l'absence de volonté politique pour, d'une part, réaliser ces études et pour, d'autre part, leur faire jouer un rôle dans le débat public.

Il convient de bien mesurer les enjeux de ces difficultés et d'en tirer les conséquences.

Soit l'on considère que l'usage du calcul économique public est un outil socialement utile participant à la défense de l'intérêt général : utile pour animer le débat (un outil pour faciliter et approfondir le débat au sein de la sphère administrative, un outil pour enrichir et élever le débat public), utile pour éclairer les décisions en situant une décision particulière par rapport à l'ensemble des dépenses publiques, utile dans un processus de rationalisation des dépenses publiques et de réforme de l'État. Et dans ce cas, il y a une réelle urgence à dresser un plan d'action pour réactiver ces pratiques et leur donner une chance réelle d'exister notamment en fixant le rôle qu'elle peuvent jouer dans le processus de décision.

Soit l'on considère que cet outil n'est pas adapté au processus de décision : par quoi le remplace-t-on ? Dans ce cas, il faut laisser ce champ de réflexion aux historiens, aux universitaires et éviter de fantasmer sur un calcul économique virtuel, idéal, qui au mieux est instru-

mentalisé à l'occasion d'un débat spécifique comme on peut le constater aujourd'hui autour de tel ou tel projet.

Il est clair que le développement effectif du calcul économique ne se fera pas sans un investissement très important en études et en moyens, et que ces derniers ne seront obtenus que si ces études ont un réel enjeu stratégique. En rester aux lamentations et aux incantations revient à perdre son temps, s'entêter revient à prendre le risque de friser le ridicule, la population, les élus ayant le sentiment, souvent à juste titre, que ces calculs restent totalement décalés par rapport aux enjeux. La loi organique relative aux lois de finances du 1^{er} août 2001 (LOLF) qui réforme en profondeur la gestion publique propose de nouvelles pratiques qui imposeront sans doute, dans les faits, la production et l'utilisation d'indicateurs complexes, de guides méthodologiques, permettant de mesurer l'utilité sociale des dépenses publiques²⁷ engagées (efficacité socio-économique des programmes engagés, qualité des services rendus aux usagers, efficacité de la gestion). Ce nouveau cadre apparaît plus favorable pour la diffusion de l'outil du calcul économique standard et la production de valeurs monétaires de référence.

Pour que cet outil puisse être accepté dans le débat et pour que les acteurs se l'approprient, plusieurs éléments doivent être réunis :

- Garantir la neutralité des valeurs de références

La production des valeurs tutélaires, des méthodes doivent pouvoir être déterminée de manière indépendante des préoccupations du moment, éloignée des lobbies qui se forme sur tel ou tel projet. Déjà, ces valeurs de références doivent dépasser le seul cadre d'un secteur particulier alors que jusqu'à maintenant les valeurs évoquées ici ont été conçues et discutées dans le cadre du seul secteur des transports. Ensuite, ces valeurs doivent être produites dans un lieu ouvert et reconnu associant des personnalités diverses garantissant et crédibili-

27. « LOI organique n° 2001-692 du 1^{er} août 2001 relative aux lois de finances », *JO* n° 177 du 2 août 2001.

Il faut rapprocher de ce texte de loi de nombreuses réflexions autour des questions relatives à la régulation publique qui ont été menées dans d'autres instances administratives et dans les milieux académiques. La régulation par l'information suppose en effet de disposer de batteries d'indicateurs reconnus y compris pour des éléments difficiles à saisir. À titre d'exemple on peut citer le travail sur les indicateurs de performances que mène l'Institut de la Gestion Déléguée (<http://www.fondation-igd.org>) dont les membres représentent l'ensemble des acteurs et partenaires concernés en France par la gestion des services publics locaux : associations d'élus locaux et des cadres dirigeants des collectivités locales ; associations de consommateurs, grandes administrations (DGCCRF, DGCL, etc.) ; opérateurs de service public privés ou filiales de groupes publics ; personnalités françaises et internationales. La production de tels indicateurs répond à une demande sociale qu'on trouve par exemple exprimée dans une charte des services publics locaux qu'ont signée le 16 janvier 2002 l'IGD et les trois grandes Associations d'élus, Maires de France, Départements et Régions.

sant les normes proposées. Le processus lui-même de production doit être transparent pour pouvoir être contesté et pour ouvrir ainsi le jeu à de nouvelles négociations.

– Assurer la continuité du processus de co-production du système de valeurs

Il est important également que le processus ne soit pas figé pour qu'il reste en phase avec la réalité sociale. Cela suppose la permanence dans le temps du cadre de production de ces valeurs. Ce cadre doit assurer le va-et-vient entre les utilisateurs de ces valeurs, les attentes et les préoccupations de la population comme des décideurs, ainsi qu'avec la recherche théorique et appliquée. Il est notamment nécessaire d'élargir le spectre des valeurs en s'attachant à ce qui est le plus difficile à appréhender et qui touche par exemple à l'équité territoriale et sociale, ou encore à l'appréhension des effets de long terme.

– Diminuer le coût d'usage des outils

Une des difficultés importantes pour la diffusion de ces systèmes de prix relatif est à trouver dans le coût d'usage des valeurs. Cela passe par le développement de guides des bonnes pratiques en prenant à bras le corps les difficultés de méthode pour appréhender la mesure et la valorisation des effets externes diffus et du risque par exemple. Ce guide, issu des échanges des praticiens et des universitaires, pourrait permettre de dégager des recettes opératoires susceptibles d'aider et de faciliter la diffusion de ces pratiques d'évaluation. L'élaboration de ce guide permettrait aussi de lister les points sur lesquels des études et des recherches mériteraient d'être lancées. Cela passe ensuite par l'animation et la formation sur le long terme des personnes susceptibles de concevoir, de réaliser ou de piloter des travaux de ce type.

– Incrire la production de ces valeurs dans un processus politique d'évaluation

Ces valeurs (un système de prix relatif pour des biens qui n'ont pas de prix) correspondent à un bien collectif que la collectivité doit décider de mettre à disposition des utilisateurs potentiels. On peut considérer qu'il existe des usagers potentiels intéressés à se référer à des valeurs non contestables pour alimenter leur argumentaire, demander des compensations ou des indemnités, défendre ou promouvoir un projet, etc. Mais cette demande ne s'exprime guère. Et l'investissement dans ce processus de production très coûteux ne peut sans doute être entrepris que s'il est porté par une volonté politique qui exige, en contrepartie des dépenses publiques consenties, un réel effort pour mesurer l'utilité sociale qu'elles produisent. Celle-là doit se traduire de manière claire, si l'on souhaite que des forces vives de l'administration et des universitaires consacrent du temps et de l'énergie pour produire de telles valeurs.

Pour terminer, il est assez difficile de convaincre les politiques de ce que la collectivité peut gagner en disposant de tels outils et de telles valeurs. Il est clair que produire ces outils, faciliter leur usage suppose d'y consacrer des moyens importants. Ce document montre précisément ce que représente cette production. Mais ne pas s'engager résolument dans cette voie conduit à réduire dans les faits la mesure de l'utilité sociale aux seuls flux financiers ce qui constitue un appauvrissement considérable du rôle de la puissance publique, puisqu'on s'interdit alors de « mesurer » le bien être que produisent les dépenses publiques. L'intérêt commun impose de lutter contre l'appauvrissement et l'affaiblissement du calcul économique, car il s'agit de la seule alternative possible au calcul politique ou strictement financier pour saisir l'intérêt général dans le long terme.

Références

- Abraham C., « Rationalité et sécurité », *Analyses et Prévisions*, (4), 1970.
- Baumstark L., « La construction de valeurs socio-économiques environnementales : l'économiste dans la posture du passeur », Colloque AFSE, Rennes, 2004.
- Babusiaux D., *Décisions d'investissement et calcul économique dans l'entreprise*, Economica, 1990.
- Bernard A., « Repenser le calcul économique public », Colloque AFSE, Rennes, 2004.
- Boiteux M., « Echanges et controverses, À propos de la critique de la théorie de l'actualisation telle qu'employée en France », *Revue d'Économie Politique*, N° 86, 1976.
- Boltanski L. et Thévenot L., *De la justification. Les économies de la grandeur*, Gallimard, (Coll. NRF – Les Essais), Paris, 1991.
- Commissariat Général au Plan, *Calcul économique et planification*, La Documentation Française, Paris, 1973.
- Commissariat Général au Plan, *Calcul économique et décisions publiques*, La Documentation Française, Paris, 1979.
- Commissariat Général au Plan, *Calcul économique et résorption des déséquilibres*, La Documentation Française, Paris, 1983.
- Commissariat Général du Plan, *Transports : pour un meilleur choix des investissements*, Rapport du groupe de travail présidé par Marcel Boiteux, La Documentation Française, Paris, 1994.
- Commissariat Général du Plan, *Transports : choix des investissements et coûts des nuisances*, Rapport du groupe de travail présidé par

- Marcel Boiteux, Rapporteur général L. Baumstark, La Documentation Française, Paris, 2001.
- Commissariat Général du Plan, *Le prix du temps et la décision publique*, Rapport du groupe présidé par D. Lebègue, Rapporteur général L. Baumstark, Coordinateur P. Hirtzman, La Documentation Française, Paris, 2005.
- Desaigues B. et Point P., *Économie du patrimoine naturel. La valorisation des bénéfices de protection de l'environnement*, Paris, Économica, 1993.
- Dupuit J., (1844« De la mesure de l'utilité des travaux publics », *Annales des Ponts et Chaussées*, 2^e semestre, n° 116, 1844, p. 332-375.
- Etechegoyen A., « *Votre devoir est de vous taire.* » *Carnets de voyage d'un philosophe à gauche et à droite*, l'Archipel, 2006, 263 p.
- Godard O., « Environnement, modes de coordination et systèmes de légitimité : analyse de la catégorie de patrimoine naturel », *Revue économique*, 41(2), 1990, p. 215-42.
- Godard O., « Autour des conflits à dimension environnementale : Évaluation économique et coordination dans un monde complexe », École Polytechnique, CNRS, Cahier n° 2004-012, 2004.
- Guesnerie R., « De l'utilité du calcul économique public », *La lettre de l'AFSE*, n° 59, 2004.
- Guesnerie R., *Dupuit hier et aujourd'hui*, Working paper, n° 2004-22, Delta, 2004.
- Henry C., « La microéconomie comme langage et enjeu de négociations », *Revue économique*, 35(1), janvier, 1984, p. 177-97.
- Inspection générale des finances et Conseil Général des Ponts et Chaussées, *Rapport d'audit sur les grands projets d'infrastructures de transport*, (N° 2002 M-026-01, N° 2002 0190-01), Administration française, Paris, 2003.
- Ministère des Transports, de l'Équipement, du Tourisme et de la Mer, « Instruction-cadre relative aux méthodes d'évaluation économique des grands projets d'infrastructures de transport », 25 mars 2004, mise à jour le 27 mai 2005, 58 p.
- Roy B. et Damart S., « L'analyse Coûts Avantages, outil de concertation et de légitimation ? », Projets et politiques de transport : Expertises en débat, *Métropolis*, n° 108/109, 2001, p.7-16.
- Walliser B. et Quinet E., « Splendeur et misère du calcul économique », *Le Monde*, 18 décembre 1990.
- Willinger M. , « La méthode d'évaluation contingente : de l'observation à la construction des valeurs de préservation », *Natures, Sciences, Sociétés*, 4(1), 1996, p. 6-22.

CHAPITRE 6

L'INFLUENCE RELATIVE DES DIFFÉRENTES VALEURS TUTÉLAIRES : UNE ÉTUDE PAR LA SENSIBILITÉ DES INDICATEURS SOCIO-ÉCONOMIQUES

Guillaume Chevasson¹

L'objectif de ce chapitre est double :

– Il est de discuter des variables utilisées pour mener à bien une évaluation socio-économique. Nous discuterons d'une part de l'influence des valeurs tutélaires dans les résultats de rentabilité socio-économique. La question sera de savoir quel est l'impact de ces valeurs sur le bénéfice, le but n'étant en aucun cas de dénigrer les valeurs tutélaires actuelles des effets non marchands et d'en proposer de nouvelles, mais il est de participer à la réflexion sur les enjeux du calcul économique. Nous discuterons d'autre part de l'influence d'un certain nombre d'autres valeurs qui sont liées à un projet d'infrastructure de transport et qui participent avec plus ou moins d'importance à sa rentabilité.

– Il est de mener une réflexion sur les valeurs retenues pour le calcul de rentabilité socio-économique afin de voir si elles sont cohérentes avec ce que les actions politiques peuvent nous enseigner. Nous focaliserons l'analyse sur deux valeurs tutélaires, les valeurs du temps et la valeur de la tonne de carbone, en nous interrogeant sur la relation vitesse – valeurs du temps et vitesse – valeur de carbone. Nous montrerons ainsi que la réduction des vitesses observée depuis quelques temps peut s'apparenter, d'une part à une considération pour des valeurs du temps plus faibles et d'autre part à une considération pour une tonne de carbone plus élevée.

1. Doctorant au Laboratoire d'Économie des Transports, Lyon

L'évaluation socio-économique, dont le principe est la valorisation dans les projets des avantages non marchands, repose sur une démarche de monétarisation des principaux effets identifiés, à savoir le gain de temps pour les usagers essentiellement, la diminution des pollutions classiques liées aux transports, la sécurité et le bruit (surtout en milieu urbain).

L'estimation monétaire de ces différents avantages présente cependant un caractère nécessairement incertain, de par la nature non 'palpable' de ces effets. Or, dans un bilan socio-économique, leur part est souvent prépondérante dans les résultats finals. Les valeurs prises en compte dans la valorisation de ces avantages relèvent donc un caractère décisif. C'est pourquoi le calcul économique fait souvent l'objet de vives critiques sur la portée réelle des évaluations de projets.

Pourtant, rappelons que si les indicateurs de rentabilité (Bénéfice pour la collectivité, Taux de Rendement Économique, etc.) sont déterminés, la démarche de l'évaluation est loin d'être déterministe. Son rôle est bien d'aider et de conforter le politique dans son choix d'investissements, en rapport avec les attentes réelles de la collectivité, et non de dicter les choix à opérer. Cela est d'autant plus vrai que des pondérations explicites ainsi que des tests de sensibilité existent, mettant en évidence le caractère non définitif du calcul économique.

En 2001, le Commissariat Général du Plan publiait un rapport intitulé *Transports : choix des investissements et coûts des nuisances* qui présentait l'ensemble des avancées réalisées dans l'analyse des nuisances et de leurs conséquences et actualisait les valeurs tutélaires qui étaient alors en vigueur, suite au précédent rapport sorti en 1994².

Pendant la rédaction de ce rapport, des discussions « animées » ont eu lieu entre l'ensemble des représentants institutionnels, scientifiques et experts, pour savoir quelle valeur associer à tel ou tel paramètre. De l'ensemble de ces débats ont émergé les valeurs que nous connaissons aujourd'hui. Ainsi, et au-delà du fait que l'accroissement du niveau de vie et de la richesse du pays majore en volume les différentes valeurs, l'actualisation de ces dernières a représenté pour certaines d'entre elles un véritable bond en avant dans la valorisation des avantages et nuisances liés au transport routier, entérinée par la collectivité.

Suite à ce rapport, le ministère des Transports a adopté une instruction-cadre³ qui rend effectives les principales recommandations

2. Commissariat Général du Plan, *Transports : pour un meilleur choix des investissements*, La Documentation Française, Paris, 1994.

3. La *circulaire* explicite le contexte dans lequel s'inscrivent les évaluations de projet, les objectifs poursuivis et met en perspective la démarche d'évaluation dans un processus continu d'évolution. L'*instruction* définit les grands principes et les différentes étapes de l'évaluation,

de ce rapport. Cette instruction, comme l'étaient les précédentes, doit être utilisée par les décideurs locaux, régionaux ou nationaux pour mesurer la rentabilité de chaque projet. Elle est donc représentative de ce qu'est, en théorie et dans la pratique, le calcul économique aujourd'hui en France.

L'objectif étant de mettre en lumière les enjeux du calcul économique tel qu'il est actuellement pratiqué par la collectivité lorsqu'elle lance des études d'investissement d'infrastructures de transport, nous avons construit, sur la base d'instructions ministérielles, un outil qui va nous permettre de simuler des études de rentabilité de projet de transport. L'outil s'appuie en partie sur les annexes techniques de l'instruction de 1998 relative aux méthodes d'évaluation économique des investissements routiers en rase campagne⁴ ainsi que sur l'instruction-cadre du 25 mars 2004 et 27 mai 2005 relative aux méthodes d'évaluation économique des grands projets d'infrastructures de transport⁵.

Nous présenterons tout d'abord l'outil de simulation de manière synthétique ainsi que les différentes valeurs tutélaires utilisées (1). Nous présenterons ensuite les résultats des études de sensibilité des indicateurs socio-économiques (2), tout d'abord aux valeurs tutélaires (2.1) puis à d'autres variables, celles que nous désignerons comme étant en amont du calcul économique et qui, pour certaines, jouent un rôle déterminant (2.2). Nous mettrons en évidence les relations qui existent entre les valeurs du temps, la vitesse et la valeur de la tonne de carbone en imaginant différents scénarios sur le niveau de ces valeurs (3). Nous concluons sur l'utilité que peut avoir une telle présentation (4).

1. PRÉSENTATION DE L'OUTIL DE SIMULATION

1.1. *Un cadre d'analyse simplifié*

L'idée générale de l'outil est de simuler l'évaluation d'un projet d'infrastructure de transport fictif. En partant d'un scénario de départ, on crée une nouvelle infrastructure et compte tenu des coûts (investis-

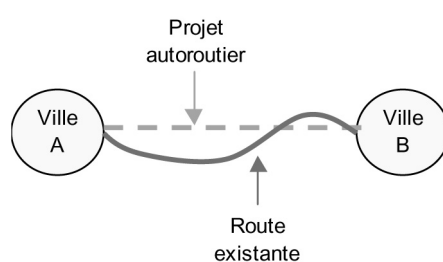
explicitant la méthode dans ces grandes lignes. Les *annexes techniques* déclinent les concepts en pratiques opérationnelles pour les différentes étapes de l'évaluation et définissent les outils à utiliser.

4. Ministère de l'Équipement, des Transports et du Logement, Direction des routes, Méthodes d'évaluation économique des investissements routiers en rase campagne, septembre 1998.

5. Ministère de l'Équipement, des transports, du Logement, du Tourisme et de la Mer, Instruction-cadre relative aux méthodes d'évaluation économique des grands projets d'infrastructures de transport, 25 mars 2004, mise à jour le 27 mai 2005.

sement, exploitation, entretien), des avantages (sécurité, pollution, gain de temps) ainsi que des valeurs tutélaires des effets externes, le projet est « validé » ou non par un certain nombre d'indicateurs socio-économiques (Bénéfice actualisé pour la collectivité, Taux de Rendement Économique).

Figure 1
Représentation du contexte d'analyse



L'hypothèse de départ est l'existence d'une route de type « route nationale », entre une ville A et une ville B. La route arrivant à saturation, la collectivité étudie la possibilité de construire une nouvelle infrastructure de type « autoroute ». Le projet est considéré comme un investissement routier en rase campagne. Il n'y a pas d'autres alternatives pour effectuer ce trajet (pas d'autres routes, pas de voies ferrées). L'année de mise en service est prévue en 2005, l'année d'actualisation est l'année qui précède la mise en service, soit 2004. Le taux est de 4 % et suit la règle d'évolution comme il est précisé dans l'instruction⁶ cadre 2005.

Tableau 1
Caractéristiques du cas théorique

	Longueur Km	Capacité en uvp ⁷ par sens	Vitesse VP/PL Km/h	Trafic initial Uvp	Croissance ⁸ du trafic par an (2025)
Route	110	1 730 uvp/h	90/80	12 000 véh/j 15 % de PL	VP 1,8 % PL 1,5 %
Autoroute	90	3 460 uvp/h	130/90	-	

6. 4 % jusqu'en 2034, puis 3,5 % jusqu'en 2054 et 3 % après.

7. Uvp : unité voiture particulière

8. DAEI-SES, La demande de transport en 2025, octobre 2004

1.2. Construction de l'outil

La mesure de la rentabilité d'un projet de transport nécessite un volume important d'informations. Le calcul repose avant tout sur la construction d'un échancier détaillé des dépenses d'investissement, des recettes de péages, des coûts d'exploitation et des effets non marchands monétarisables.

Nous pouvons décomposer l'outil en différents modules.

Le premier module rassemble l'ensemble des hypothèses nécessaires à la conduite de l'évaluation. Il contient :

- les variables descriptives de l'état initial qui serviront de base pour décrire l'évolution de la situation de référence : la longueur de l'itinéraire, le niveau de trafic, la répartition VP/PL, la répartition VP essence/diesel, la capacité de l'infrastructure, les vitesses réglementaires, les coûts de circulation, les taux accidentogènes et les investissements éludés ;
- les variables descriptives du projet de transport envisagé qui serviront à établir les différents coûts mais aussi à préciser la variation de qualité du service de transport offert : la longueur de l'itinéraire, la capacité de l'infrastructure, les vitesses réglementaires, le coût de construction, les dépenses d'exploitation et d'entretien, les tarifs de péage, les coûts de circulation et les taux accidentogènes ;
- les variables macro-économiques qui serviront à faire évoluer différents paramètres et permettent d'établir les projections sur la durée de vie du projet : les taux de croissance du trafic, le taux de croissance de la consommation finale des ménages, le taux de croissance du produit intérieur brut et le taux d'inflation ;
- les valeurs tutélaires des effets non marchands : valeurs de la pollution atmosphérique, valeurs du temps voyageurs et marchandises, valeurs de la vie humaine et valeur de la tonne de carbone.

Le deuxième module concerne la prévision de la demande et le calcul des trafics entre les itinéraires routier et autoroutier. Il se compose de deux éléments :

- le calcul des temps de parcours, à l'aide de courbes débit-vitesse de type BPR⁹, qui vont servir en partie à déterminer les coûts de circulation sur les différents itinéraires. Les calculs s'effectuant en

9. Formule employée et développée par U.S. Bureau of Public Road. La formulation simplifiée est : $T = T_{\min} [1 + a(\frac{q}{q_{\max}})^b]$ dans laquelle T et q désignent respectivement le temps de parcours et le débit du chaînon, T_{\min} , le temps de parcours minimum, c'est-à-dire à vitesse libre, q_{\max} , la capacité maximale du chaînon, a et b, deux paramètres calibrés à partir de mesures.

partie sur la base des variables descriptives de l'état initial et du scénario d'aménagement (qualité de service) ;

- le calcul de l'affectation du trafic entre les itinéraires à partir des coûts de circulation à l'aide du modèle « logit logarithmique », dit « loi d'Abraham¹⁰ ».

Le calcul des trafics et leur répartition sur les deux infrastructures constituent la base sur laquelle l'évaluation socio-économique est construite.

Le troisième module est celui de l'évaluation proprement dite. Il sert à quantifier les effets nets du projet, c'est-à-dire la variation relative des différents paramètres entre la situation de référence et la situation avec projet. Il comprend :

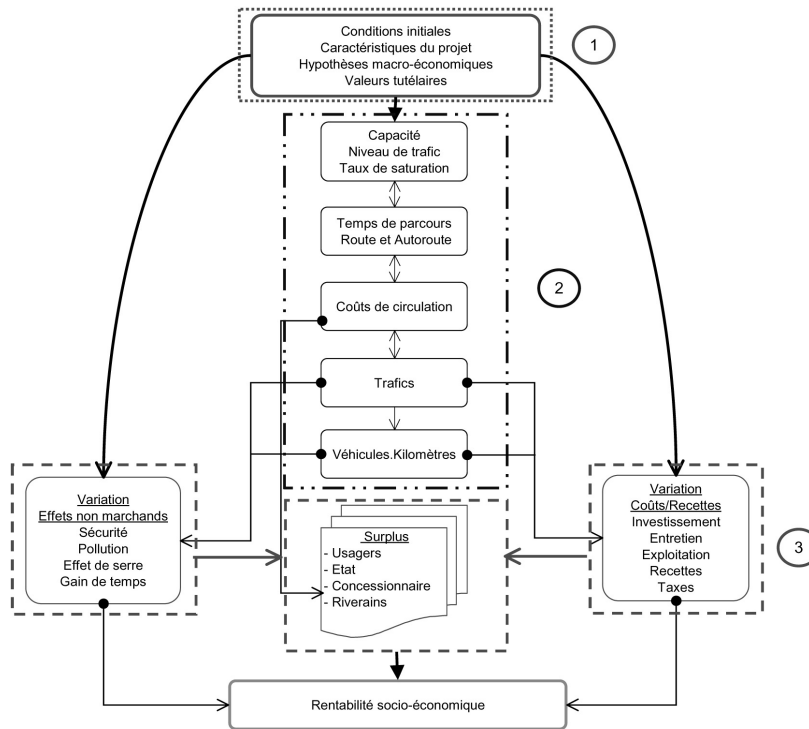
- les effets directs du projet, comme les coûts d'investissement, les investissements érudés, les coûts d'exploitation, le niveau de trafic et les recettes, qui serviront à mesurer la rentabilité économique du projet ;

- la variation du surplus des agents concernés par la transformation d'état du système. Cette variation par rapport à la situation de référence sert à calculer la rentabilité socio-économique du projet.

Le fonctionnement de l'outil est assez simple. À partir des valeurs que nous entrons dans le premier module, nous déterminons le scénario (ou situation) de référence et le scénario d'aménagement. Ces scénarios seront utilisés dans le deuxième module et serviront de base aux calculs des temps de parcours sur chaque itinéraire. Ces temps de parcours permettront de déterminer des coûts de circulation, qui sont à la fois la clef de répartition des trafics entre la route et l'autoroute et à la base du niveau de trafic induit par l'autoroute. À partir de cette répartition des trafics et de leur niveau, nous entrons dans le troisième module qui calcule, outre les coûts et recettes du projet, les coûts de l'insécurité routière, de la pollution atmosphérique et de l'effet de serre. Les surplus des différents agents concernés par le projet s'ajoutent au calcul des coûts marchands pour déterminer le bénéfice actualisé pour la collectivité.

10. $\frac{T_1}{T_2} = \frac{C_2}{C_1}$ avec $T_1 + T_2 = T$ trafic total de la relation origine-destination, T_1 et T_2 sont les trafics à attribuer aux deux itinéraires et C_1 et C_2 les coûts de circulation sur ces itinéraires.

Figure 2
Architecture de l'outil de simulation



1.3. Les valeurs tutélaires : rappel et règle d'évolution

Les valeurs tutélaires présentées ici sont celles que recommande d'utiliser la nouvelle instruction. Elles trouvent leur origine dans le rapport du Commissariat Général du Plan¹¹. L'année de référence est l'année 2000. Les valeurs utilisées dans l'outil suivent les règles d'évolution qui sont indiquées à la suite de chaque tableau.

11. Commissariat Générale du Plan, *Transports : choix des investissements et coûts des nuisances*, La Documentation Française, Paris, 2001.

- Les valeurs du temps
 - Les valeurs du temps voyageur en interurbain

En euro 2000 par voyageur/heure	Distances inférieures à 50 km	Distances comprises entre 50 et 400 km	Distances supérieures à 400 km
Route voyageur	8,94	$0,016 \times d + 8,1$	14,5

Ces valeurs évoluent en fonction de la dépense de consommation des ménages par tête, en monnaie constante, avec une élasticité de 0,7.

- Les valeurs du temps marchandise

Mode	Euro/tonne/heure	Euro/heure
Route marchandise	0,45	6,75

L'évolution sera égale aux 2/3 de l'évolution du PIB en volume. À cette valeur du temps des marchandises, il faut ajouter la variation du coût d'exploitation des transporteurs due aux gains ou pertes de temps. Pour la route, la valeur de 31,4 euros par heure (en euro 2000) a été retenue. Cette valeur n'est pas indexée et sera la même en euros constants dans l'avenir.

- Les valeurs de la vie humaine

En Euro 2000	Transport routier
Tué	1 000 000
Blessé grave	150 000
Blessé léger	22 000

Les valeurs évoluent comme la consommation finale des ménages par tête.

- Les valeurs de la pollution atmosphérique

En euro/100 véh.km	Urbain dense	Urbain diffus	Rase campagne
VP	2,9	1,0	0,1
PL	28,2	9,9	0,6

Ces valeurs évoluent comme la dépense de consommation des ménages par tête diminuée de 5,5 % par an pour les véhicules légers et de 6,5 % par an pour les poids lourds sur la période 2000-2020, afin de

tenir compte des améliorations dans les émissions de polluants des véhicules dues au progrès technique.

- La valeur de la tonne de carbone

2000 – 2010	Après 2010
100 € la tonne de carbone, soit 6,6 centimes d'euro par litre d'essence et 7,3 centimes d'euro par litre de diesel	+ 3 %/an

- La valeur du bruit

Dans les zones non habitées, les dommages causés par le bruit peuvent être négligés. Dans le cadre de l'outil, l'hypothèse a été faite que le long des deux infrastructures routières, il n'y avait pas d'habitations. Le trafic est un trafic d'échange uniquement.

1.4. Le bilan coûts-avantages actualisé pour la collectivité

Tableau 2

Bilan coûts-avantages actualisé pour la collectivité

Bilan coûts-avantages actualisé pour la collectivité		
Agent économique	Nature des avantages pris en compte	
Usagers VP	Avantages temps Avantages confort Avantages frais de fonctionnement Avantages péage	
Somme actualisé des avantages usagers VP		(1)
Usagers PL	Avantages temps Avantages frais de fonctionnement Avantages péage	
Somme actualisé des avantages usagers PL		(2)
Puissance publique	Recettes fiscales sur les usagers VP Recettes fiscales sur les usagers VP et PL Recettes fiscales sur les usagers induits Recettes fiscales sur les péages Sécurité routière	
Somme actualisé des avantages de la puissance publique		(3)
Riverain	Pollution atmosphérique Effet de serre	
Somme actualisé des avantages de la puissance publique		(4)

Bilan coûts-avantages actualisé pour la collectivité	
Agent économique	Nature des avantages pris en compte
Concessionnaire	Recettes de péages
Somme actualisé des avantages du concessionnaire	(5)
Somme actualisé des avantages tout agent économique confondu	(6)=(1+2+3+4+5)
Dépense d'exploitation	(7)
Avantage net global	(8) = (6-7)
Investissement	(9)
Bénéfice actualisé pour la collectivité	(10) = (8-9)

Le bilan coûts-avantages actualisé pour la collectivité comprend :

- les bilans des usagers de la route pour les Voitures Particulières (1) et pour les Poids Lourds (2) qui comprennent les avantages de temps, les avantages de confort pour les VP uniquement, les avantages de frais de fonctionnement (entretien, dépréciation, carburant) et les avantages de péages ;

- le bilan de la puissance publique (3) qui comprend les recettes fiscales des usagers VP sur les frais de fonctionnement, les recettes fiscales sur les consommations de carburant VP et PL, les recettes fiscales liées aux péages, les recettes fiscales liées au trafic induit et les coûts liés à la sécurité routière ;

- le bilan des riverains (4) qui comprend le coût lié à la pollution atmosphérique et le coût lié à l'effet de serre. Concernant le coût lié à l'effet de serre, nous l'avons mis dans cette catégorie d'acteur, bien qu'il s'agisse d'un phénomène mondial ;

- le bilan du concessionnaire (5) qui comprend les recettes liées au péage.

La somme de ces bilans (6) constitue les avantages du projet auxquels nous retranchons les dépenses d'exploitation (7) pour avoir l'avantage net global (8). Le bénéfice actualisé pour la collectivité (10) s'obtient en retranchant le coût d'investissement (9) à l'avantage net global.

2. SENSIBILITÉ DES INDICATEURS DE RENTABILITÉ SOCIO-ÉCONOMIQUE

L'objet de l'analyse de la sensibilité est de faire ressortir les variables et paramètres « critiques » de l'évaluation de rentabilité, c'est-à-dire ceux dont les variations, positives ou négatives, par rapport à la valeur utilisée dans le cas de référence, ont l'effet le plus prononcé sur

le bénéfice ou le TRE en ce sens qu'ils entraînent les modifications les plus importantes de ces indicateurs. L'objectif ici est de montrer, dans un premier temps, comment une ou des variations dans le système de prix relatifs influencent les résultats de rentabilité socio-économique. En d'autres termes, il s'agit de connaître l'influence de ces valeurs dans les résultats. Par influence, nous faisons référence au poids relatif de ces valeurs dans la balance de l'analyse coûts-avantages, dont le résultat peut être représenté par le bénéfice actualisé pour la collectivité

Afin de tester la sensibilité des valeurs tutélaires sur les résultats, nous sommes partis d'une méthode simple dans la pratique mais qui appelle quelques recommandations d'usage. Nous limiterons notre analyse à deux indicateurs : le bénéfice actualisé pour la collectivité et le Taux de Rendement Economique (TRE). Le principe étant de considérer, par rapport à un scénario de base¹² (base 100), l'effet d'une variation (de +/-10%) des valeurs sur le niveau de ces deux indicateurs. De cette gymnastique, nous pouvons faire sortir ce que nous appellerons des élasticités « apparentes ». Ces élasticités ne sont valables que dans le contexte de l'outil et de ses hypothèses de base concernant le scénario de référence et le scénario d'aménagement. Il n'est pas possible de les généraliser à tous les projets d'infrastructures routières, mais ces résultats sont satisfaisants dans la mesure où ils donnent des ordres de grandeurs de l'impact de chaque valeur tutélaire et ils hiérarchisent l'importance de chacune, les unes par rapport aux autres. La sensibilité du bénéfice pour la collectivité s'analyse en termes d'évolution en *pourcentage* par rapport à notre scénario de base, tandis que la sensibilité du TRE s'analyse en termes d'évolution en *points de pourcentage*. Les résultats ici sont à prendre avec précaution. Il faut les considérer plutôt comme des ordres de grandeur que comme des résultats précis, car la sensibilité des indicateurs aux variables n'est pas une fonction linéaire, ce qui explique pourquoi il y a un signe ~ (environ) devant les résultats d'élasticité.

2.1. Sensibilité des indicateurs de rentabilité socio-économiques aux valeurs tutélaires

Cette série de variables concerne les valeurs tutélaires des effets non marchands. Quatre effets différents sont analysés : les valeurs du

12. Le scénario de base prend en compte les valeurs officielles concernant les valeurs tutélaires des effets non marchands, les valeurs unitaires concernant les coûts de circulation, les ratios d'insécurité routière, les vitesses réglementaires, le coût d'investissement, les dépenses d'exploitation et d'entretien, le taux d'actualisation, le taux de croissance du trafic et les hypothèses macro-économiques. Concernant les paramètres liés aux longueurs, aux capacités et au trafic existant à la mise en service, ils sont au choix de l'utilisateur.

temps, les valeurs de la vie humaine, la valeur de la tonne de carbone et les valeurs de la pollution atmosphérique.

Tableau 3
Relations entre valeurs tutélaires et indicateurs de rentabilité

Variation		Bénéfice	TRE
Valeurs du temps voyageurs et/ou marchandises	▲	▲	▲
	▼	▼	▼
Valeurs de la vie humaine	▲	▲	▲
	▼	▼	▼
Valeur de la tonne de carbone	▲	▲	▲
	▼	▼	▼
Valeurs de la pollution atmosphérique ¹³	▲	▲	▲
	▼	▼	▼

▲ : augmentation de la valeur / ▼ : diminution de la valeur

Lecture du tableau : une variation à la hausse des valeurs du temps a un effet à la hausse sur le bénéfice et sur le TRE. A l'inverse, une variation à la baisse des valeurs de la vie humaine a un effet à la baisse sur le bénéfice et sur le TRE.

Tableau 4
Résultats des études de sensibilité aux valeurs tutélaires

Elasticités apparentes	Rentabilité socio-économique	
	Bénéfice actualisé pour la collectivité	Taux de Rendement Economique
	En %	En pts de %
Valeurs du temps voyageurs et marchandises	~ 1,0	~ 0,15
Valeur du temps voyageurs	~ 0,96	~ 0,14
Valeur du temps marchandises	~ 0,04	~ 0,01
Valeurs de la vie humaine	~ 0,24	~ 0,03
Valeur de la tonne de carbone	~ 0,016	~ 0,002
Valeurs de la pollution atmosphérique	~ 0,0003	~ 0,00001

Lecture du tableau : une augmentation de 1 % de la valeur de la tonne de carbone diminue le bénéfice actualisé pour la collectivité de 0,016 % et baisse le TRE de 0,002 point. Et inversement.

13. Une hausse des valeurs de la pollution atmosphérique a un impact positif sur le bénéfice et le TRE, car le coût de la pollution est lié aux véhicules-kilomètres parcourus. Or, dans notre exemple, la longueur de l'autoroute est plus faible que celle de la route et le trafic autoroutier, même augmenté du trafic induit, engendre des véhicules.kilomètres inférieurs à ceux du trafic routier dans la situation de référence. Cela explique la relation de cause à effet que nous observons ici.

Les résultats du tableau suggèrent plusieurs commentaires.

– La valorisation du temps constitue bel et bien le volet prépondérant dans la rentabilité socio-économique des projets (élasticité proche de 1). Cette importance s'explique par le fait que les valeurs du temps influencent fortement le niveau des coûts de circulation qui, eux-mêmes, déterminent la répartition du trafic entre les itinéraires et également le niveau de trafic induit par la mise en place de la nouvelle infrastructure. Or, le niveau de trafic et sa répartition entre les différents itinéraires sont à la base des bilans des usagers de la route où la valorisation du temps occupe une place importante.

– Les gains de sécurité constituent le deuxième volet important. Ils peuvent s'expliquer par deux facteurs : des valeurs de la vie humaine élevées et un différentiel accidentogène entre les deux infrastructures important. Pour un même niveau de trafic le nombre d'accidents est beaucoup plus élevé sur route que sur autoroute et les ratios d'insécurité¹⁴ sont à l'avantage de l'autoroute. Ainsi le coût d'insécurité sur autoroute est 4,7 fois moins important que sur la route.

– Le coût de l'effet de serre ne peut pas influencer à lui tout seul le niveau de rentabilité d'un projet (élasticité très faible : 0,016). Nous reviendrons sur ce point dans la partie 3.

– La pollution atmosphérique ne joue qu'un rôle de figuration dans la détermination de la rentabilité. Il faut faire attention ici au fait que les valeurs utilisées dans l'outil sont celles recommandées en rase campagne et sont donc beaucoup plus faibles qu'en urbain dense (29 fois pour les VP et 47 fois pour les PL). De plus, les valeurs de la pollution diminuent dans le temps du fait qu'elles évoluent à la hausse comme la consommation finale des ménages par tête mais à la baisse du fait de la prise en compte du progrès technique dans le temps. Or la deuxième évolution est plus importante que la première.

2.2. Sensibilité des indicateurs de rentabilité socio-économique aux variables en amont du calcul économique

Les variables sont dites « en amont du calcul économique » car elles diffèrent d'un projet à un autre. La longueur, le coût de construction, la capacité de l'infrastructure, le trafic de base, etc., sont des variables indépendantes du calcul économique, contrairement aux valeurs tutélaires qui sont les mêmes selon les projets.

14. Le nombre de tués pour 100 accidents est de 21,7 sur route type 7m alors qu'il n'est que de 9,3 sur autoroute interurbaine. Le nombre de blessés graves est de 49,6 sur route et de 26,3 sur autoroute. Le nombre de blessés légers est de 114,5 sur route et de 125,3 sur autoroute. Les taux d'insécurité résultent ici d'une moyenne sur les années 1999 à 2001.

Tableau 5
**Relation entre différents paramètres, indicateurs de rentabilité
 et niveau de trafic**

Variation		Bénéfice	TRE
Longueur de tracé du projet	▲	▲	▲
	▼	▼	▼
Vitesses sur route	▲	▲	▲
	▼	▼	▼
Trafic existant à la mise en service	▲	▲	▲
	▼	▼	▼
Vitesses sur autoroute	▲	▲	▲
	▼	▼	▼
Vitesses sur réseaux	▲	▲	▲
	▼	▼	▼
Croissance du trafic	▲	▲	▲
	▼	▼	▼
Capacité de l'autoroute	▲	▲	▲
	▼	▼	▼
Coût de circulation ¹⁵ (hors VDT)	▲	▲	▲
	▼	▼	▼
Coût d'investissement	▲	▲	▲
	▼	▼	▼
Evolution de la CFM/t	▲	▲	▲
	▼	▼	▼
Evolution du PIB	▲	▲	▲
	▼	▼	▼

▲ : augmentation de la valeur / ▼ : diminution de la valeur

Lecture du tableau : Une variation à la hausse de la longueur du tracé du projet a un effet à la baisse sur le bénéfice et le TRE. A l'inverse, une variation à la baisse des coûts d'investissement a un effet à la hausse sur le bénéfice et le TRE.

15. $C = h \cdot T + i \cdot L + m + p$, avec h (valeur du temps), T (Temps de parcours), i (Malus d'inconfort), L (Longueur du parcours), m (Dépenses monétaires de fonctionnement des véhicules), p (Péage éventuel).

2.2.1. Variables des niveaux de service offerts

Cette première série de variables prend en compte les niveaux de service offerts par les différents itinéraires de transport en concurrence. Ces variables jouant sur le niveau de congestion influent sur les temps de parcours et donc sur l'avantage relatif offert par le projet.

Tableau 6
Sensibilité des indicateurs aux niveaux de service offert

Elasticités apparentes	Bénéfice pour la collectivité			Taux de Rendement Economique		
	-10 %	+10 %	e	-10 %	+10 %	e
Longueur du tracé du projet	44,5 %	-44,8 %	~ 4,5	8,7	-7,8	~ 0,8
Vitesses sur route	31,8 %	-23,3 %	~ 2,8	4,6	-3,5	~ 0,4
Vitesses sur réseaux	20,3 %	-22,2 %	~ 2,1	3,0	-3,5	~ 0,3
Trafic comptabilisé à l'année d'étude du projet	-14,5 %	14,7 %	~ 1,5	-2,1	2,1	~ 0,2
Vitesses sur autoroute	-10,5 %	2,0 %	~ 0,6	-1,5	0,1	~ 0,1
Capacité de l'autoroute	-0,02 %	0,06 %	~ 0,004	-0,01	0,01	~ 0,001

Lecture du tableau : une augmentation de 1 % de la longueur du tracé du projet diminue le bénéfice actualisé pour la collectivité de 4,5 % et baisse le TRE de 0,8 point. Et inversement.

– La longueur du projet influence fortement les résultats. En diminuant la longueur du projet, nous touchons directement aux temps de parcours. Par conséquent, les coûts de circulation sont considérablement abaissés et le bilan pour les usagers est augmenté et inversement, c'est ce qui explique en grande partie l'impact de ce paramètre sur les résultats. Ce résultat souligne l'impact potentiel sur les résultats de rentabilité que peut avoir le choix de la bande lors de l'Avant Projet Sommaire dans laquelle s'inscrit l'infrastructure et/ou le choix de variantes de tracé d'un projet.

– « L'élément majeur du processus d'évaluation, mais en même temps celui dont la fragilité a les plus lourdes conséquences, est la prévision des trafics¹⁶. » Cette phrase résume parfaitement l'importance des études de trafic dans la détermination de la rentabilité socio-économique d'un projet d'infrastructure de transport. Le trafic existant avant la mise en service est aussi un élément clef puisqu'un projet de transport, pour être légitimé, s'appuie essentiellement sur le niveau

16. Page 31, Commissariat Général au Plan, *Transports : pour un meilleur choix des investissements*, La Documentation Française, novembre 1994.

de trafic 'potentiel'. L'élasticité du bénéfice au niveau de trafic est forte. Ce qui met l'accent sur l'impact d'une étude de trafic qui aurait été mal réalisée. Une surestimation du trafic existant ne serait pas neutre vis-à-vis de la rentabilité socio-économique d'un projet et inversement. Par ailleurs, plus le niveau de trafic existant est faible et plus la sensibilité des résultats à une variation de cette variable est importante.

– Les vitesses sur les réseaux déterminent en grande partie les coûts de circulation. En touchant aux vitesses, sur la route, sur l'autoroute ou sur les deux, les temps de parcours s'allongent ou s'abaissent. Or quand nous savons que les gains de temps pour les usagers de la route constituent le volet prépondérant dans la détermination de la rentabilité du projet, il est facile d'en déduire qu'une variation sur les vitesses a des effets importants sur les indicateurs de rentabilité socio-économique.

Nous nous trouvons ici avec cette variable devant un phénomène de cause à effet qui a besoin d'explication. En effet, il paraît étonnant qu'une variation de la vitesse sur autoroute ait aussi peu d'impact sur le bénéfice. Les élasticités sont relativement faibles par rapport à ce que nous nous attendions. Nous pouvons expliquer ces résultats. Le résultat des élasticités est une moyenne des deux variations (+/-). Or, une baisse de la vitesse n'a pas le même impact qu'une hausse. La sensibilité du bénéfice est forte à la baisse des vitesses (-11 %) mais très peu à la hausse (+2 %). Cette différence s'explique en partie par le fait que les usagers roulent à la vitesse calculée par le modèle d'affectation qui se base sur la vitesse réglementaire pour calculer la vitesse « libre ou à vide ». Donc à la hausse, le modèle exclut la possibilité pour les usagers de rouler moins vite. Au final les gains de temps pour les usagers engendrés par la hausse des vitesses sont plus que compensés par les frais de fonctionnement des véhicules, notamment par la hausse de la consommation de carburant. Et cette compensation vient atténuer l'effet sur le bénéfice.

– La capacité de l'autoroute influe sur le débit maximal d'écoulement du trafic et par conséquent sur la relation débit-vitesse et donc sur le temps de parcours. Cette relation de cause à effet pointe du doigt la question du bon calibrage des infrastructures. Dans notre cas de figure, la capacité n'a que très peu d'influence sur le bénéfice.

2.2.2. Variables des coûts du projet

Cette série de variables concerne l'investissement. La première variable concerne le coût du projet¹⁷. La deuxième concerne les dépenses d'entretien et d'exploitation ainsi que les grosses réparations et investissement complémentaire sur autoroute en service (ICAS). Ces éléments sont centraux notamment pour le concessionnaire, puisque qu'ils influent directement sur son Excédent Brut d'Exploitation.

Tableau 7

Sensibilité des indicateurs aux coûts d'investissement

Elasticités apparentes	Bénéfice pour la collectivité			Taux de Rendement Economique		
	-10 %	+10 %	e	-10 %	+10 %	e
Coût d'investissement	2,7 %	-2,7 %	~ 0,3	1,9	-1,6	~ 0,2
Entretien, exploitation, réparations	1,4 %	-1,4 %	~ 0,1	0,2	-0,2	~ 0,02

Lecture du tableau : une augmentation de 1% du coût d'investissement baisse le bénéfice actualisé pour la collectivité de 0,3 % et le TRE de 0,2 point. Et inversement.

– En général, étant donné le niveau de connaissance et d'expérience en matière de construction, le coût d'investissement est une donnée qui peut être estimée, sans grande difficulté, de façon précise. Néanmoins, des risques de dérive des coûts existent, notamment lors de retards de chantier, de risque géologique, etc. Une surestimation, comme une sous-estimation des coûts d'infrastructure peuvent tout de même avoir des effets non négligeables sur les indicateurs de rentabilité socio-économique.

– Les dépenses d'entretien, d'exploitation et les grosses réparations sont des coûts qui viennent limiter l'excédent brut d'exploitation du concessionnaire, mais leur poids dans la détermination du bénéfice pour la collectivité et du TRE est limité.

2.2.3. Variables influençant la répartition du trafic

Cette série de variables concerne la détermination du coût de circulation sur chaque itinéraire, élément central dans la répartition des trafics. Parmi elles, nous trouvons les valeurs du temps et les frais de fonctionnement des véhicules.

17. Le coût d'investissement du projet s'étale ici sur 7 années (n-4 à n+2), avec une proportion plus importante des coûts se situant entre n-3 et n.

Tableau 8
Sensibilité des indicateurs aux coûts de circulation

Elasticités apparentes	Bénéfice pour la collectivité			Taux de Rendement Economique		
	-10 %	+10 %	e	-10 %	+10 %	e
Coûts de circulation VP/PL	-10,4 %	10,4 %	~ 1,04	-1,5	1,5	~ 0,15
Valeurs du temps voyageurs et marchandises	-10,0 %	9,9 %	~ 0,99	-1,5	1,4	~ 0,15
Frais de fonctionnement	-0,6 %	0,4 %	~ 0,05	-0,03	0,01	~ 0,02
Coût de circulation (hors VDT) VP	-0,9 %	0,7 %	~ 0,08	-0,08	0,05	~ 0,01
Coût de circulation (hors VDT) PL	0,3 %	-0,3 %	~ 0,03	0,05	-0,05	~ 0,005
Coût de circulation VP	-10,4 %	10,4 %	~ 1,04	-1,5	1,5	~ 0,15
Valeurs du temps voyageurs	-9,6 %	9,5 %	~ 0,96	-1,4	1,4	~ 0,14
Péage	2,0 %	-2,0 %	~ 0,21	0,3	0,4	~ 0,04
Malus d'inconfort	-3,6 %	3,6 %	~ 0,36	-0,5	0,5	~ 0,05
Carburant	1,5 %	-1,5 %	~ 0,15	0,2	-0,2	~ 0,02
Entretien, Dépréciation	-0,8 %	0,8 %	~ 0,08	-0,2	0,2	~ 0,02
Coût de circulation PL	-0,1 %	0,1 %	~ 0,01	-0,01	0,01	~ 0,001
Valeur du temps marchandises	-0,4 %	0,4 %	~ 0,04	-0,1	0,1	~ 0,01
VDT + coût d'exploitation	-1,7 %	1,7 %	~ 0,17	-0,35	0,35	~ 0,035
Coût d'exploitation	-1,4 %	1,3 %	~ 0,13	-0,3	0,3	~ 0,03
Péage	0,7 %	-0,8 %	~ 0,08	0,01	-0,01	~ 0,01
Carburant	-0,2 %	0,2 %	~ 0,02	-0,03	0,03	~ 0,003
Entretien	-0,3 %	0,3 %	~ 0,03	-0,04	0,04	~ 0,004

Lecture du tableau : une augmentation de 1 % du tarif de péage VP baisse le bénéfice actualisé pour la collectivité de 0,21 % et le TRE de 0,04 point. Et inversement.

Les coûts de circulation sont la clef de répartition entre les trafics. Ils déterminent également le niveau de trafic induit par la mise en place de la nouvelle infrastructure. Les résultats des élasticités nous montrent que :

- les coûts de circulation pris dans leur ensemble et sans la valorisation du temps n'ont pas de réel impact sur le bénéfice et le TRE (élasticités de 0,08 % pour les VP et de 0,03 % pour les PL) ;
- pour les VP, le tarif de péage et le malus d'inconfort sont les deux variables qui peuvent influencer les résultats ;
- pour les PL, le coût d'exploitation est un élément pouvant influencer les résultats avec également les tarifs de péages ;
- les coûts de circulation des PL ont une influence limitée par rapport à ceux des VP. Cela est dû à l'importance relative des trafics VP par rapport à celui des PL ;
- la valeur du temps reste le paramètre clef.

2.2.4. Variables macro-économiques

Cette série de variables concerne les hypothèses macro-économiques qui vont permettre de faire évoluer différents paramètres : les valeurs tutélaires (pollution, valeurs de la vie humaine), les valeurs liées aux coûts de circulation (valeur du temps, malus d'inconfort) et le niveau des trafics (voyageurs et marchandises).

Tableau 9

Sensibilité des indicateurs aux hypothèses macro-économiques

Elasticités apparentes	Bénéfice pour la collectivité			Taux de Rendement Economique		
	-10 %	+10 %	e	-10 %	+10 %	e
Croissance du trafic	-3,0 %	3,0 %	~ 0,3	-0,3	0,3	~ 0,03
Croissance voyageurs	-2,6 %	2,7 %	~ 0,27	-0,3	0,3	~ 0,03
Croissance marchandises	-0,3 %	0,3 %	~ 0,03	-0,03	0,03	~ 0,003
Evolution de la CFM/t	-3,1 %	3,2 %	~ 0,3	-0,03	0,03	~ 0,03
Evolution du PIB	-0,1 %	0,1 %	~ 0,01	-0,01	0,01	~ 0,001

Lecture du tableau : une augmentation de 1 % du trafic voyageurs augmente le bénéfice actualisé pour la collectivité de 0,3 % et le TRE de 0,03 point. Et inversement.

– Le taux de croissance du trafic est une variable importante dans l'estimation de la rentabilité d'un projet et, paradoxalement, c'est la variable la plus sujette à des erreurs d'appréciation¹⁸. Car celle-ci dépend d'un certain nombre de facteurs qui sont totalement indépendants du projet de transport (prix du baril de pétrole (\$), taux de change (€à\$), augmentation du PIB, dépense de consommation des ménages, etc.). L'élasticité des résultats aux taux de croissance est d'environ 0,3. Ce résultat peut paraître faible mais il ne faut pas oublier que la variation porte sur le taux de croissance. Et généralement dans les études de sensibilité lors des évaluations, les résultats sont soumis aux hypothèses basses et hautes et celles-ci varient de +30 % à -30 % par rapport aux hypothèses moyennes. Une surestimation ou une sous-estimation du taux de croissance du trafic peut avoir des effets certains sur la rentabilité d'un projet.

18. Pour exemple, nous pouvons citer le cas du viaduc de Millau dont le trafic dépasse d'environ 20 % les prévisions qui avaient été réalisées. Même si, pour ce cas, il y a certainement un effet « site d'attraction » qui perdurera les premières années, du fait de l'ampleur et de la réputation de l'ouvrage, cela montre que la prévision de trafic attendu est un exercice difficile et source d'erreur d'appréciation, tant à la hausse qu'à la baisse.

– L'évolution de la consommation finale des ménages par tête a de nombreux liens avec d'autres variables. Elle détermine leur évolution (valeurs du temps voyageur, valeurs de la pollution atmosphérique, valeurs de la vie humaine). Ces liens expliquent qu'une variation du niveau de la CFM/t a des effets non négligeables sur les résultats de rentabilité. De plus, la même remarque que celle sur le taux de croissance du trafic peut s'appliquer ici, à savoir que nous appliquons une variation sur un taux de croissance. Il faut noter que les hypothèses dans l'instruction qui encadre le taux moyen de croissance sont de plus ou moins 25 %, c'est-à-dire largement plus élevées qu'une variation de +/-10 %. De telles variations auraient des impacts beaucoup plus élevés sur les indicateurs de rentabilité.

– L'évolution du produit intérieur brut influe uniquement sur l'importance de l'évolution de la valeur du temps pour les marchandises. Son impact est limité. La même remarque que précédemment s'applique ici : il s'agit de variations en pourcentage sur des taux de croissance. *Les taux de croissance du trafic sont ici considérés indépendamment de l'évolution du PIB*, ce qui, bien sûr, n'est pas le cas dans la réalité. Dans le cas contraire, une variation du PIB aurait beaucoup plus d'importance. Si nous prenons les hypothèses de croissance du PIB retenues¹⁹, le bénéfice baisse de 6,7 % dans le cas de l'hypothèse faible (PIB : +1,5 %/an) et augmente de 9,2 % dans le cas de l'hypothèse forte (PIB : +2,3 %/an). L'évolution du PIB n'est donc pas neutre vis à vis des résultats de rentabilité.

Au vu de ces résultats (§2.1, §2.2), il apparaît clairement que la rentabilité socio-économique d'un projet dépend très fortement des hypothèses retenues et qu'elle s'appuie d'un côté sur un certain nombre de variables qui sont susceptibles de conditionner fortement ou assez fortement le niveau de bénéfice pour la collectivité et d'un autre côté sur des variables qui jouent un rôle moins important ou infime.

– La valorisation des coûts externes (pollution atmosphérique, effet de serre, sécurité routière) marque la prise en compte des effets non marchands par la collectivité, mais le niveau des valeurs tutélaires est beaucoup trop faible pour pouvoir influencer significativement les résultats de rentabilité socio-économique ; seules les valeurs tutélaires de la vie humaine ont une influence assez significative sur les résultats.

– La valorisation du temps est l'élément clef dans la détermination du bénéfice pour la collectivité. Par conséquent, tout ce qui modifie le

19. Cf. DAEI-SES, La demande de transport en 2025, octobre 2004.

temps de parcours (longueur, vitesse, niveau de trafic existant) a un impact déterminant sur le niveau de rentabilité.

3. L'ART DE FAIRE DU CALCUL ÉCONOMIQUE À L'ENVERS OU COMMENT METTRE EN ÉVIDENCE LA PRÉFÉRENCE RÉVÉLÉE DE LA PUISSANCE PUBLIQUE POUR UNE VALEUR DE LA TONNE DE CARBONE PLUS ÉLEVÉE

Diminuer les vitesses de circulation n'est pas une idée nouvelle. Depuis longtemps l'Agence de l'Environnement et de la Maîtrise de l'Énergie (ADEME) le préconise dans le but de faire baisser la consommation de carburant et les émissions de CO₂. Récemment, suite à la hausse du prix du pétrole, le ministre des Transports avait 'lancé un pavé dans la marre' en émettant l'idée de limiter à 115 km/h la vitesse sur autoroute. Derrière cette idée, il y avait une justification de réduire la consommation de carburant (et donc les émissions de CO₂) et de soulager le porte-monnaie des automobilistes. Devant une levée de boucliers, cette idée a été abandonnée. Mais le fait même que le gouvernement envisage cette possibilité, et que, par ailleurs, il poursuive sa politique de réduction des vitesses, peut nous laisser supposer que la puissance publique accorde à la valeur de la tonne carbone une valeur supérieure à ce qu'elle est actuellement (100 €). En effet, lorsque nous savons qu'une baisse de la vitesse sur autoroute équivaut à une baisse des valeurs du temps et donc à une baisse du bénéfice pour la collectivité, il n'y a pas derrière cette idée une justification économique au sens d'une recherche d'efficacité. En revanche, il y a peut être derrière cette idée une préférence implicite pour une valeur de la tonne de carbone plus élevée.

L'objectif de cette partie est de montrer que ceux qui prônent une baisse des vitesses accordent indirectement à la tonne de carbone une valeur plus importante qu'elle ne l'est aujourd'hui.

Nous partirons du constat que la valorisation du temps est le volet prépondérant dans la détermination de la rentabilité des projets et que, face à elle, la valorisation de la tonne de carbone n'a qu'un poids très limité (3.1) Cette différence mettra en lumière la préférence de la collectivité pour le temps. Pourtant en montrant qu'il existe une relation entre les valeurs du temps et la vitesse (3.2), nous pourrons nous demander si, derrière les politiques en faveur d'une réduction des vitesses engagées depuis quelques années, qui ne vont donc pas dans le sens de l'efficacité économique, ne se cache pas une préférence des décideurs en faveur de l'environnement, puisque baisser les vitesses équivaut à prendre une valeur de la tonne de

carbone plus élevée (3.3). Nous limiterons notre analyse à la vitesse autorisée sur autoroute.

3.1. *Valorisation du temps versus Valorisation de l'effet de serre*

Les résultats présentés ici font suite aux résultats précédents concernant les élasticités. Nous avons poussé les curseurs, à l'extrême quelquefois, afin de faire apparaître le fossé entre le temps et l'effet de serre.

Valeur de la tonne de carbone

Comme nous l'avons vu précédemment, les indicateurs de rentabilité socio-économique ont une sensibilité très faible aux variations de la valeur de la tonne de carbone. Les résultats du tableau 10 le confirment : pour que le poids de l'effet de serre puisse réellement influencer le bénéfice, il faudrait multiplier la valeur de la tonne de carbone actuelle par environ 7,2 (si nous retenons comme impact significatif sur le bénéfice une variation de -10 %).

Tableau 10

Augmentation de la valeur de la tonne de carbone

	Variation	Equivalence en €	Bénéfice
Valeur de la tonne de carbone	-100 %	0	+1,6 %
	-50 %	50	+0,8 %
	-40 %	60	+0,6 %
	-30 %	70	+0,5 %
	-20 %	80	+0,3 %
	-10 %	90	+0,2 %
	+10 %	110	-0,2 %
	+50 %	150	-0,8 %
	+100 %	200	-1,6 %
	+400 %	500	-6,4 %
	+900 %	1 000	-14,5 %
	+1400 %	1 500	-22,5 %

Lecture du tableau : une augmentation de 100 % de la valeur de la tonne de carbone diminue de 1,6 % le bénéfice pour la collectivité. Une valeur du carbone nulle (-100 %) entraîne une hausse du bénéfice de 1,6 %.

Une telle augmentation de la valeur de la tonne de carbone serait-elle justifiée aujourd'hui ? Si nous regardons ce qui se passe sur le marché des permis d'émissions négociables, une hausse semble être à l'heure actuelle incohérent et inapproprié. En effet la valeur de la tonne de CO₂ tourne autour de 22,6 € la tonne²⁰. Ce prix équivaut à 83,7 €²¹ la tonne de carbone. Bien que ce prix ait déjà dépassé quelquefois la barre des 27 € la tonne de CO₂ (équivalent à 100 € la tonne de carbone), il reste en dessous de la valeur fixée par le Commissariat Général du Plan.

Il faut souligner quand même que cette valeur est proche de celle recommandée par le Commissariat Général du Plan et qu'elle reste sujette à des variations potentielles (à la hausse), surtout si les quotas d'émission venaient à être modifiés (à la baisse). Mais il faut souligner également que ces variations sont limitées et qu'une tonne de carbone multipliée par 7,2, soit 720 € la tonne, est une opération irréalisable, dans les conditions actuelles du marché.

Valeurs du temps

Les valeurs du temps ont un très fort impact sur les résultats. Les résultats du tableau 11 le confirment. La valorisation du temps apparaît bien comme l'élément prépondérant dans le calcul de rentabilité socio-économique du projet. Mais, de même qu'il paraît irréaliste de multiplier la valeur du carbone par 7,2, il paraît tout aussi irréaliste de vouloir continuer à augmenter les valeurs du temps indéfiniment. Surtout que la tendance aujourd'hui est à l'augmentation de la congestion à l'entrée des agglomérations (ce qui fait perdre une partie des gains de temps accumulés auparavant) et à la réduction des vitesses sur la route (ce qui équivaut à une baisse des valeurs du temps).

20. Moyenne des prix à la clôture sur les 123 derniers jours de cotation (du 24/06/05 au 21/12/2005). (<http://www.powernext.fr>)

21. En prenant 0,27g de carbone comme équivalence pour 1 gramme de CO₂.

Tableau 11

Variation des valeurs du temps

	Variation	Impact sur le bénéfice
Valeurs du temps	-50 %	-52,9 %
	-40 %	-41,6 %
	-30 %	-30,7 %
	-20 %	-20,2 %
	-10 %	-10,0 %
	+10 %	+9,9 %
	+20 %	+19,6 %
	+30 %	+29,2 %
	+40 %	+38,7 %
	+50 %	+48,2 %

Lecture du tableau : une augmentation des valeurs du temps de 10 % augmente le bénéfice pour la collectivité de 9,9 % ; une diminution de 10 % entraîne une baisse d'environ 10,0 %.

Au vu des résultats des tableaux, il est évident que le fossé qui sépare la valorisation de l'effet de serre et celle du temps est immense. Ce fossé peut s'interpréter comme la différence de degré de considération de la collectivité vis-à-vis de ces deux effets. Pourtant la réduction tendancielle des vitesses, amorcée depuis quelques années semble venir contredire cette conclusion hâtive. En effet, il semble qu'implicitement certains acteurs de la sphère institutionnelle, de par les choix qu'ils effectuent, donnent à la tonne de carbone une valeur supérieure à celle qui est actuellement utilisée par le calcul économique.

3.2. La relation Vitesse – Valeurs du temps

Depuis un certain temps, nous assistons à une tendance à la réduction de la vitesse tant en urbain qu'en interurbain. Les politiques de sécurité routière n'y sont bien sûr pas étrangères : mise en place des radars, vitesse limitée à 110 km/h sur certains tronçons autoroutiers, renforcements des contrôles, permis probatoire, aggravation des sanctions pour les récidivistes, automatisation de la chaîne contrôle-sanction, campagnes d'information et d'incitation par le biais des médias, etc. La puissance publique a ainsi multiplié les mesures en faveur d'une réduction du nombre de tués sur les routes françaises. La

poursuite de cet objectif s'est traduit par une réduction des vitesses moyennes observées.

Dans la suite, nous nous intéressons *uniquement à la baisse des vitesses sur autoroute*. Si nous prenons le cas des autoroutes de liaison, la vitesse moyenne (VP+PL) est passée de 107,5 km/h en 2000 à 105,8 km/h en 2004, soit une baisse de 1,6 %.

Tableau 12

Vitesses moyennes constatées sur autoroutes de liaison entre 2000 et 2004

	Mode	2000	2001	2002	2003	2004
Moyenne	Ensemble	107,5	107,0	107,9	105,8	105,8
Autoroutes de liaison	VP	124,0	124,0	123,5	119,0	118,5
	PL	91,0	90,0	92,3	92,7	93,0

Source : Observatoire National Interministériel de Sécurité Routière
VP : moyenne jour/nuit ; PL : moyenne 2,3,4 essieux

Maintenant, regardons la relation qui existe entre une baisse des vitesses et les valeurs du temps. Pour arriver à ces résultats, nous prenons comme référence le bénéfice issu du scénario de base et nous recherchons les valeurs du temps qui entraînent la même variation à la baisse de ce bénéfice que celle engendrée par la baisse des vitesses.

Tableau 13

Relation entre Vitesses sur autoroute et Valeurs du temps

Vitesses sur autoroute	Une baisse des vitesses de :	Vitesses en km/h	est équivalente à une baisse des VDT de :
Vitesses limites : VP : 130 km/h PL : 90 km/h	-5 %	VP : 123,5 – PL : 85,5	-4,9 %
	-10 %	VP : 117 – PL : 81	-10,5 %
	-15 %	VP : 110,5 – PL : 76,5	-18,3 %
	-20 %	VP : 104 – PL : 72	-27,3 %

Lecture du tableau : une baisse des vitesses sur autoroute de 10 % est équivalente à une baisse des valeurs du temps de 10,5 %.

Ce que nous pouvons dire au vu des résultats précédents est que la baisse des vitesses observées depuis l'année 2000 peut s'apparenter à une baisse des valeurs du temps. Nous avons ici une relation qui, du point de vue du calcul économique, peut sembler paradoxale, puisque d'un côté la collectivité a augmenté les valeurs du temps et d'un autre

côté elle agit d'une manière qui sous-entend qu'elle les considère moins élevées. En d'autres termes, nous avons d'un côté une valorisation du temps qui constitue l'essentiel des avantages résultant d'un projet et qui détermine en grande partie sa rentabilité et d'un autre côté nous assistons à des politiques de réduction de la vitesse qui vont à l'encontre de cette règle.

Qu'est-ce que cela peut signifier ? Quelle explication pouvons-nous trouver à ce paradoxe ? En nous intéressant maintenant à la relation qui existe entre vitesses/valeurs du temps et valeur de la tonne de carbone, nous pouvons peut-être trouver un début d'explication.

3.3. *L'hypothèse implicite des décideurs en faveur d'une valeur de la tonne de carbone plus élevée*

Nous avons donc une relation entre vitesses et valeurs du temps. Maintenant regardons l'effet d'une variation des valeurs du temps ou des vitesses sur le niveau implicite de la tonne de carbone.

Valeurs du temps / Valeur de la tonne de carbone

Pour arriver à ces résultats, nous prenons comme référence le bénéfice issu du scénario de base et nous recherchons la valeur de la tonne de carbone qui a le même impact sur le bénéfice qu'une baisse des valeurs du temps.

Tableau 14

Relation entre valeurs du temps et valeur de la tonne de carbone

	Une baisse des VDT de :	est équivalente à une hausse de la valeur du carbone de :	Valeur de la tonne de carbone en €	Impact sur le bénéfice
Valeurs du temps	-0,16 %	+10 %	110	-0,2 %
	-0,81 %	+50 %	150	-0,8 %
	-1,62 %	+100 %	200	-1,6 %
	-6,46 %	+400 %	500	-6,4 %
	-14,42 %	+900 %	1 000	-14,5 %
	-22,23 %	+1400 %	1 500	-22,5 %

Lecture du tableau : une baisse de 1,6 % des valeurs du temps est équivalente à une augmentation de 100 % de la valeur de la tonne de carbone.

Vitesses / Valeur de la tonne de carbone

De la même façon, regardons l'effet d'une variation des vitesses sur le niveau implicite de la tonne de carbone. Pour arriver à ces résultats, nous prenons comme référence le bénéfice issu du scénario de base et nous recherchons la valeur de la tonne de carbone qui a le même impact sur le bénéfice qu'une baisse des vitesses.

Tableau 15

Relation entre vitesses et valeur de la tonne de carbone

	Une baisse de la vitesse sur autoroute de :	Soit : vitesses en km/h VP / PL	est équivalente à une hausse de la valeur t/Cb de :	Valeur de la tonne de carbone en €	Impact sur le bénéfice
Vitesses limites sur autoroute VP :130 km/h PL :90 km/h	-1,0 %	128,7 / 89,1	+65 %	165	-1,0 %
	-1,6 %	127,9 / 88,6	+84 %	184	-1,4 %
	-2,0 %	127,4 / 88,2	+132 %	232	-2,1 %
	-5,0 %	123,5 / 85,5	+301 %	401	-4,8 %
	-10,0 %	117,0 / 81,0	+654 %	754	-10,5 %
	-11,5 %	115 / 79,6	+794 %	894	-12,8 %
	-15,0 %	110,5 / 76,5	+1150 %	1 250	-18,5 %
	-20,0 %	104,0 / 72,0	+1733 %	1 833	-27,9 %

Lecture du tableau : une baisse de 5,0 % des vitesses sur autoroute (VP : 123,5 km/h et PL : 85,5 km/h) a le même impact sur le bénéfice qu'une hausse de 301 % de la valeur de la tonne de carbone.

3.4. Résultats

Considérons ici deux cas de figure : le premier concerne la baisse des vitesses observées sur les autoroutes de liaison, le deuxième concerne l'idée de faire passer la vitesse sur autoroute à 115 km/h.

1) Nous considérons uniquement la baisse des vitesses sur autoroutes observées entre 2000 et 2004.

Nous supposons que le coût monétarisé de toute tonne de carbone rejetée dans l'atmosphère est pris en compte dans la détermination des indicateurs de rentabilité socio-économique et nous supposons également que les vitesses moyennes sont proportionnelles aux vitesses réglementaires.

En considérant que la baisse des vitesses observées entre 2000 et 2004 (-1,6 %) soit prise en compte par le calcul économique lors de l'évaluation du projet, il apparaît que celle-ci peut s'apparenter à une prise en compte de la valeur d'une tonne de carbone de l'ordre de 184 € valeur 2005²². Par ailleurs, cette valeur correspondrait à un coût de 12,1 centimes d'euro par litre d'essence et de 13,4 centimes d'euro par litre de diesel. La valeur de la tonne de carbone trouvée serait donc de 184 € jusqu'en 2010, puis augmenterait de 3 %/an, conformément aux recommandations de la nouvelle instruction.

2) Nous considérons la proposition de baisser la vitesse réglementaire sur l'autoroute à 115 km/h au lieu de 130 km/h, ce qui correspondrait à une baisse de 11,5 %.

Nous supposons toujours que le coût monétarisé de toute tonne de carbone rejetée dans l'atmosphère est pris en compte dans la détermination des indicateurs de rentabilité socio-économique. La baisse des vitesses s'effectue ici sur les vitesses réglementaires.

– Si la baisse concernait uniquement les voitures particulières. Elle correspondrait à une prise en compte de la valeur de la tonne de carbone d'environ 575 € valeur 2005. Par ailleurs, cette valeur correspondrait à un coût de 37,6 centimes d'euro par litre d'essence et de 41,8 centimes d'euro par litre de diesel.

– Si la baisse concernait les voitures particulières et les poids lourds (c'est-à-dire que nous appliquons la même baisse aux vitesses réglementaires VP et PL). Elle correspondrait à une prise en compte de la valeur de la tonne de carbone d'environ 894 € valeur 2005. Par ailleurs, cette valeur correspondrait à un coût de 58,5 centimes d'euro par litre d'essence et de 65,0 centimes d'euro par litre de diesel.

Du point de vue du calcul économique et des méthodes et pratiques des évaluations de rentabilité socio-économique, ces résultats signifient qu'il est identique d'augmenter la valeur de la tonne de carbone et de la fixer à 575 € que d'abaisser la vitesse réglementaire des voitures particulières à 115 km/h sur autoroute. Le même remarque s'applique aux autres cas de figure.

L'avantage qu'apportent les gains de temps dans la rentabilité socio-économique d'un projet n'est plus à démontrer. Pourtant les valeurs du temps en interurbain sont peut-être surévaluées par rapport à ce qu'estime la collectivité. Les résultats nous ont permis de montrer que les valeurs retenues pour le calcul de la rentabilité socio-

22. Année de la mise en service de l'autoroute. Année à partir de laquelle nous calculons les avantages.

économique (valeurs du temps et valeur de carbone) ne sont pas toujours cohérentes avec ce que les actions politiques peuvent nous enseigner.

Bien que les résultats doivent être considérés avec précaution, ils ont l'intérêt de montrer que derrière l'idée d'abaisser les vitesses, se cache peut-être, hormis une préférence déclarée pour la sécurité routière, une préférence pour une valeur de la tonne de carbone plus grande. Le glissement dans la hiérarchie des priorités entre la sécurité routière et les émissions de CO₂ est assez net. Le nombre de morts sur les autoroutes à péage a fortement baissé en France. Une réduction des vitesses maximales aurait de moins en moins d'effets sur le résultat final. Ce serait donc bien pour des questions environnementales et énergétiques que pourrait se programmer une baisse de la vitesse maximale.

Abaisser la vitesse réglementaire sur autoroute pour les voitures particulières serait une manière pour la collectivité d'afficher sa volonté de réduire les émissions de CO₂. Cette solution révélerait qu'implicitement la collectivité accorde à la valeur de la tonne de carbone un montant plus élevé que celui qui est officiellement fixé et à l'inverse un montant aux valeurs du temps plus faible.

4. CONCLUSION

Quelle peut être l'utilité d'une telle démarche ? Son objectif n'étant pas de donner des valeurs tutélaires qui fassent que le calcul économique soit à l'abri de critiques de tout bord, nous pourrions être interrogés alors sur l'apport d'un tel travail.

Les résultats que nous avons présentés ici n'ont qu'un seul but, celui de mettre en lumière les implications des choix effectués par la collectivité dans la valorisation des effets non marchands et également dans la construction des outils du calcul économique.

Comme nous l'avons vu, les indicateurs de rentabilité socio-économique, notamment le bénéfice pour la collectivité, sont très sensibles aux variables qui ont une influence sur les gains de temps (longueur du tracé, vitesses, valeurs du temps, trafic existant avant la mise en service). Si certaines variables sont indépendantes d'un projet à un autre, certaines, comme les valeurs du temps, sont du ressort de la puissance publique. Par conséquent, le choix de leur niveau n'est pas neutre dans la détermination de la rentabilité des projets.

Dans un autre registre, les valeurs environnementales, notamment celle de la tonne de carbone, n'ont qu'une place de représentation au

sein de la détermination de la rentabilité socio-économique. Mais les augmenter afin que leurs impacts soient plus importants ne serait pas forcément la solution, puisque, d'une part il faudrait les augmenter beaucoup, ce qui ne serait pas sans conséquence sur les autres secteurs de l'économie et d'autre part il y a d'autres façons (détournées) de les prendre en considération, si nous en croyons les résultats précédents.

Références

Commissariat Général au Plan, *Transport : choix des investissements et coûts des nuisances*, La Documentation Française, Paris, 2001.

Ministère de l'Équipement, des Transports et du Logement, Direction des Routes, « Instruction relative aux méthodes d'évaluation économique des investissements routiers en rase campagne », octobre 1998.

Ministère de l'Équipement, des Transports, de l'Aménagement du territoire, du Tourisme et de la Mer, « Instruction-cadre relative aux méthodes d'évaluation économique des grands projets d'infrastructures de transport », 25 mars 2004, Mise à jour le 27 mai 2005.

DEUXIÈME PARTIE

DES PROJETS À LA PROGRAMMATION DES INVESTISSEMENTS

Selon une intuition répandue, l'abaissement du taux d'actualisation de 8 % à 4 % à la suite du rapport¹ du Commissariat Général du Plan de février 2005 et de l'instruction cadre du 27 mai 2005 du ministre de l'Équipement, des Transports et du Logement, devrait entraîner une augmentation du nombre des projets d'investissement socio-économiquement rentables. Par voie de conséquence, une augmentation de l'enveloppe des financements est nécessaire. Or, cet accroissement demeure problématique. Les finances publiques sont en effet contraintes dans le cadre du Pacte européen de stabilité et de croissance. Le déficit budgétaire et la dette publique doivent respecter des seuils, respectivement 3 % et 60 % du PIB, et elles peinent à le faire. À ces exigences, s'ajoute un débat récurrent sur le niveau des prélèvements obligatoires. De ce télescopage entre une demande potentielle croissante de fonds publics et un resserrement des disponibilités naît une interrogation sur la programmation des projets socio-économiquement rentables. Si l'on ne peut pas tous les financer, dans quel ordre et selon quel calendrier faut-il les réaliser ? Lesquels sont à écarter ?

L'intuition d'un accroissement des projets socio-économiquement rentables est-elle fondée ? Certes, il est clair, selon la théorie de la croissance optimale élémentaire que si le taux d'actualisation diminue, le capital par tête de croissance équilibrée augmente, donc le volume de l'investissement aussi. Mais l'abaissement du taux d'actualisation atténuant l'écrasement des avantages futurs, il peut rendre non rentables des projets dont les avantages futurs sont en fait des inconvé-

1. Commissariat Général du Plan (2005), *Le prix du temps et la décision publique*, président Daniel Lebègue, coordonnateur Philippe Hirtzman, rapporteur général Luc Baumstark, La Documentation Française.

nients, comme par exemple les pollutions, planétaires ou locales² ; la résultante de ces deux effets est donc ambiguë. En outre, comme le souligne Christian Gollier (cf. Première partie), le taux d'actualisation antérieur de 8 % intégrait une prime de risque « forfaitaire », alors que l'adoption du taux d'actualisation de 4 % doit s'accompagner d'un traitement explicite des risques, projet par projet, de sorte que, au total sur l'ensemble des projets, les effets se compensent au moins partiellement et peut-être totalement. Ce point restant à élucider, on considérera ci-après, à titre de conjecture, la perspective d'une augmentation de la demande de fonds publics.

Il convient ici de rappeler que la question de la tension entre volume des projets socio-économiquement éligibles et enveloppe de fonds publics disponible n'est pas nouvelle. Ainsi, dans la première circulaire interministérielle que l'on peut considérer comme relativement achevée quant à sa cohérence méthodologique, celle du 20 janvier 1970³, il était recommandé d'appliquer un coefficient de rareté des fonds publics à l'analyse coût avantage. Cette préconisation partait du constat que tous les projets réputés rentables, au sens où ils présentaient un bénéfice net actualisé (BNA) positif ne pouvaient être financés.

Le coefficient de rareté appliqué à toute dépense publique abaisse, bien entendu, les soldes des bilans actualisés des projets. Cela revient à mettre en cohérence la capacité de financement des projets et le taux d'actualisation officiel, ainsi que cela a pu être recommandé par la littérature spécialisée⁴ puis exprimé d'une manière plus générale⁵. Les débats se sont poursuivis au sein du Commissariat Général du Plan. On en trouve la trace notamment dans les rapports *Calcul économique et planification*⁶ (1973), *Les choix d'investissements publics décentrali-*

2. Exemple : soit un projet dont l'investissement est 1 et qui comporte dans un an un bénéfice de 20 et dans deux an un dommage de 20.

Calcul du bénéfice actualisé :

Avec 8 % : $B = 20 * \left[\frac{1}{1.08} - \left(\frac{1}{1.08} \right)^2 \right] - 1 = 0.37$: le projet est à retenir.

Avec 4 % : $B = 20 * \left[\frac{1}{1.04} - \left(\frac{1}{1.04} \right)^2 \right] - 1 = -0.26$: le projet est à rejeter.

3. Ministère de l'Équipement (1970), *Calcul de rentabilité des investissements routiers*, Circulaire du 20 janvier 1970, division des études et programmes.

4. Abraham C. et Thédié J. (1962), *Rentabilité des travaux routiers*, Rapport définitif du cycle d'étude 1961-1962, Ministère de l'Équipement, Direction des routes et de la circulation routière.

5. Abraham C. et Thomas A. (1970), *Microéconomie, décisions optimales dans l'entreprise et dans la nation*, Dunod.

6. Commissariat Général du Plan (1973), *Calcul économique et planification*, Président Ullmo Y., rapporteur Bernard A., La Documentation Française.

*sés en période de croissance ralentie*⁷ (1981) et *Calcul économique et résorption des déséquilibres*⁸ (1983). La conclusion à laquelle ces rapports parviennent est que l'introduction d'une contrainte de rareté des fonds publics éloigne de la situation d'équilibre optimal de premier rang, qui sous-tend la théorie de l'actualisation, et conduit à la recherche d'un optimum de second rang⁹. Et que, « dans une optique de décentralisation des choix macro-économiques du Plan à l'intérieur de chaque enveloppe sectorielle d'équipements, un procédé approximatif mais commode est d'appliquer un coefficient majorateur affectant les dépenses – d'investissement et d'exploitation – ainsi que la part des avantages couverts par des recettes monétaires, de façon à réduire la rentabilité des projets d'investissement en compétition dans la mesure nécessaire pour que les financements des projets rentables s'adaptent aux ressources disponibles¹⁰ ».

Jacques Thédié¹¹ (1983) a développé cette approche en précisant que le coefficient majorateur est en fait annuel et peut varier « de manière très importante » d'une année à l'autre, en fonction tant de la contrainte budgétaire annuelle que de la liste des projets candidats. Il préconise pour chaque projet de calculer à l'aide de ces coefficients majorateurs annuels le « bénéfice actualisé généralisé (BAG) » dans les différentes hypothèses d'année de mise en service et d'inscrire l'opération l'année qui permet de maximiser ce BAG.

La question a également été examinée par André Laure et Claude Abraham dans leur article « Étude des programmes d'investissements routiers »¹² (1984). Les auteurs mentionnent d'abord la règle selon laquelle, s'il n'existait pas de restriction budgétaire, il conviendrait de réaliser chaque projet à une date optimale, qui est celle où le taux de rentabilité socio-économique immédiate est égal au taux d'actualisation ; règle qui est applicable si l'on ne s'éloigne pas trop de certaines hypothèses simplificatrices. Les auteurs abordent ensuite « la recherche de l'emploi optimum de crédits insuffisants ». Ils proposent une

Voir notamment pages 62 et sq. « L'incompatibilité du taux d'actualisation et de l'enveloppe d'équipements publics : vrai ou faux problème ? ».

7. Commissariat Général du Plan (1981), *Les choix d'investissements publics décentralisés en période de croissance ralentie*, président Malinvaud E., rapporteur Balladur J.-P., mimeo.

8. Commissariat Général du Plan (1983), *Calcul économique et résorption des déséquilibres*, président Malinvaud E., président du groupe technique Guesnerie R., rapporteurs, Walliser B. et Goudard D., La Documentation Française.

9. Voir notamment « Calcul économique et planification », *jam. cit.*, page 21.

10. *Ibid.*, page 88.

11. Thédié J. (1983), « Du choix des investissements sous contrainte financière », *Annales des Ponts et Chaussées*, 1^{er} trimestre.

12. Abraham C. et Laure A. (1984), « Étude des programmes d'investissements routiers », *Annales des Ponts et Chaussées*, 1^{er} trimestre 1983.

méthode imaginant que l'on puisse réaliser seulement une fraction de chaque projet, cette fraction s'appliquant tant à l'investissement qu'aux avantages correspondants. Ils en concluent que « l'on obtient le programme optimum, c'est-à-dire le revenu actualisé maximum, en choisissant chaque année les opérations par ordre décroissant du ratio « avantages de l'année sur investissement » jusqu'à épuisement des crédits disponibles et en recommençant l'année suivante avec les opérations restantes ». Ces résultats reposent toutefois sur des hypothèses simplificatrices, notamment sur celle-ci : si la croissance des avantages d'un projet est une année donnée supérieure à la croissance des avantages d'un autre projet cette année-là, il en va de même pour toutes les autres années envisagées.

Par rapport à cet état des lieux préexistant, les dispositions accompagnant la récente diminution du taux d'actualisation de 8 % à 4 % apportent des éclairages complémentaires. On en retiendra ici deux qui ont trait aux finances publiques. L'un concerne le coût d'opportunité des fonds publics, censé refléter la perte d'efficacité socio-économique (distorsions) due aux impôts. L'autre porte sur l'existence supposée d'un plafonnement des fonds publics disponibles, susceptible de se traduire par un coût fictif de rareté (variable duale associée à la contrainte quantitative).

Ainsi, le rapport du Plan précité (2005) précise (page 87) sur ces deux points : « Le rapport recommande sur la question très controversée du coût d'opportunité des fonds publics de réviser le coefficient de 1,5 proposé par le Plan en 1985 ainsi que les modalités de son usage. Il recommande par ailleurs de renforcer le mécanisme de sélection des projets de manière à utiliser les ressources financières disponibles au mieux des intérêts des collectivités en s'appuyant sur le ratio bénéfice actualisé par euro public dépensé. ».

De son côté, l'instruction cadre du 27 mai 2005 fournit en son annexe III les précisions suivantes :

- en matière de coût d'opportunité des fonds publics, « le calcul de rentabilité socio-économique prendra en compte les flux budgétaires nets sur la durée de vie du projet (en provenance de l'ensemble des collectivités publiques) que l'on majorera par un coefficient de 1,3 d'opportunité des fonds publics » ;
- en matière de contrainte budgétaire, « afin de tirer le meilleur parti d'un financement public limité, la règle de classement des projets doit être non pas le bénéfice actualisé induit par le projet, mais le bénéfice actualisé par euro public dépensé, prenant en compte l'ensemble des dépenses publiques et des recettes publiques éventuelles au cours de la durée de la vie du projet (actualisés au taux de 4 %) ;

en conséquence, on déterminera pour chaque projet, en sus des indicateurs de rentabilité socio-économique, le bénéfice actualisé pour la collectivité divisé par la valeur actualisée nette des dépenses budgétaires de toutes les collectivités publiques (nettes des recettes fiscales éventuellement générées), pendant la durée du projet. Cet indicateur pourra faire l'objet d'une analyse de sensibilité ».

Les réflexions engagées à la suite de la récente diminution du taux d'actualisation permettent-elles d'aller plus loin ? Les trois contributions qui suivent s'attachent à préciser plusieurs aspects relatifs à la programmation des investissements en situation de rationnement des fonds publics.

Dans le chapitre 7, Alain Bonnafous et William Roy examinent d'abord le taux de subvention nécessaire pour assurer la rentabilité financière d'un projet. Puis, considérant un ensemble de 17 projets d'autoroute à péage, ils recherchent sous contrainte budgétaire l'ordre de réalisation optimal de ces projets, au moyen d'une simulation en nombres entiers. Ils comparent ensuite ce résultat au classement obtenu en utilisant divers indicateurs, comme le taux de rentabilité socio-économique, le taux de rentabilité interne et le ratio bénéfice actualisé par euro investi, pour conclure que ce dernier indicateur est celui qui fournit la meilleure approximation. Ils étayent enfin cette conclusion en utilisant le procédé de fractionnement des projets et le théorème de Kuhn et Tucker relatif à l'optimisation sous contraintes.

Dans le chapitre 8, Emile Quinet et Alain Sauvart proposent une méthode générale fondée sur le procédé de fractionnement des projets et sur l'algorithme de programmation linéaire pour résoudre la question de la programmation dynamique optimale en temps discret des projets, que ceux-ci soient indépendants ou liés, que la contrainte budgétaire soit connue avec certitude ou probabilisée. Les auteurs insistent sur la nécessité de rechercher, pour chaque projet, sa date de réalisation optimale, qui dépend beaucoup – outre la contrainte budgétaire – de l'évolution de l'investissement en fonction de la date de mise en service et de la convexité de la courbe des avantages par comparaison avec les autres projets candidats. Des multiples simulations qu'ils ont faites, il ressort que, parmi les différents indicateurs envisageables, le mieux corrélé avec la programmation optimale semble être le taux de rentabilité immédiate et non pas le ratio bénéfice actualisé par euro public investi. Ils préconisent de décentraliser le choix entre plusieurs variantes d'un même projet, mais d'établir ensuite au niveau central l'ordre chronologique de réalisation tenant compte des contraintes budgétaires.

Joël Maurice, dans le chapitre 9, reprend l'examen de la programmation dynamique en temps discret, sous contrainte annuelle de fonds publics, en utilisant le procédé de fractionnement des projets et en interprétant les résultats de la programmation linéaire à la lumière du théorème de Kuhn et Tucker. L'auteur montre que la détermination d'un « prix fictif annuel de rareté des fonds publics » permettrait de décentraliser complètement la recherche de l'optimum par maximisation du bénéfice fictif actualisé : choix de la variante optimale, y compris du mode de financement, et de la date optimale de mise en service. Il examine le critère auxiliaire du ratio bénéfice actualisé par euro public dépensé et souligne qu'il convient de considérer autant de valeurs de ce ratio que d'années envisageables pour la mise en service. Il en conclut que ce ratio ne fournit qu'un substitut imparfait à la communication de la série des prix fictifs des fonds publics.

L'auteur propose en outre une piste de réflexion sur l'optimisation de l'appel aux financements privés, dont le remboursement sur fonds publics serait gagé sur la chronique, supposée connue, des enveloppes budgétaires disponibles. Cette démarche conduirait à une valeur unique du « prix fictif de rareté des fonds publics » au lieu d'une série de prix fictifs annuels ; cette valeur pourrait être déterminée à dire d'experts (comme le fut la révision du taux d'actualisation). Le calcul du bénéfice fictif actualisé relatif à chaque projet au niveau décentralisé en serait facilité, renforçant ainsi les raisons de considérer ce critère comme le plus approprié et d'en recommander l'usage et la diffusion. Le recours au critère ratio bénéfice par euro public dépensé serait lui aussi simplifié, le montant des euros publics dépensé devant toutefois être actualisé en utilisant non pas le taux d'actualisation, mais le taux d'intérêt réel sans risque ; ce critère auxiliaire resterait néanmoins imparfait et susceptible de conduire à des ambiguïtés.

CHAPITRE 7

ÉVALUATION, FINANCEMENT ET PROGRAMMATION DES INVESTISSEMENTS

Alain Bonnafous et William Roy¹

L'articulation entre évaluation et modalités de financement pose beaucoup de questions vives qui peuvent être aussi bien théoriques que pratiques. Au niveau théorique, par exemple, peut se poser la question du taux d'actualisation cohérent avec une croissance économique optimisée : ce taux étant supposé proprement calculé, doit-il en résulter nécessairement que tout projet présentant une valeur actualisée nette positive soit effectivement réalisé ? et que soit ainsi déterminé le niveau optimal du budget public qu'il convient de mobiliser pour les investissements de transport ? Plus généralement, comment doivent s'articuler, d'une part, le calcul économique et la théorie du rendement social, d'autre part, l'analyse macro-économique de la détermination des enveloppes budgétaires ?

À un niveau pratique, le problème se présente de manière plus brutale dans la mesure où cette enveloppe budgétaire consacrée aux investissements de transport est tout bonnement exogène. La capacité de financement public se présente alors comme une contrainte et le choix des investissements comme une optimisation de second rang dont la fonction objectif reste la somme des valeurs actualisées nettes (socio-économiques) dégagées par l'ensemble des projets réalisés.

Dans cette mise en ordre, il convient en outre de s'interroger sur les conséquences d'une tendance lourde qui voit se généraliser les cofinancements entre l'utilisateur et le contribuable. C'est, en somme, le rapport « Lebègue » (Commissariat Général du Plan, 2005) et la circulaire ministérielle qui en a résulté qui appellent cette réflexion renouvelée. Consacré à la « Révision du taux d'actualisation des investisse-

1. Laboratoire d'Économie des Transports, CNRS, Université Lyon 2, ENTPE

ments publics », ce rapport, en préconisant une baisse sensible du taux d'actualisation² du Plan, implique une multiplication des projets réputés rentables pour la collectivité et accroît donc l'écart entre le budget nécessaire à leur réalisation et les capacités de financement public. C'est pourquoi il suggère aussi de prendre en compte cette contrainte en retenant l'ordre de classement des projets selon les valeurs décroissantes du ratio : **VAN / euros de fonds publics investis**.

Il s'agit, en somme, d'un ratio qui a vocation, pour désigner les projets prioritaires, à se substituer à l'utilisation du taux d'actualisation du Plan comme seuil de rentabilité des projets réputés profitables pour la collectivité (les projets étant d'autant plus attrayants que leur TRE est supérieur à ce taux du Plan).

Si le numérateur de ce ratio est à première vue (à première vue seulement) indépendant des modalités de financement, il n'en va pas de même du dénominateur. On peut, en effet, établir une relation fonctionnelle entre, d'une part, le taux de subvention que nécessitera chaque projet, et d'autre part, le taux de rentabilité interne (TRI) du projet. Le TRI est un indicateur purement financier, un certain niveau de subvention permettra d'atteindre le TRI exigé par l'opérateur. Ce dernier peut varier selon la nature des opérateurs (EPIC, SEM, société privée,...), leurs compétences, mais aussi selon le montage financier et les modalités de couverture des risques qu'il comporte.

Le problème se complique du fait que le TRI propre au projet et le taux de subvention qui en découle ne sont pas indépendants de variables qui peuvent faire partie des éléments de proposition contenus dans les réponses aux appels d'offre. Par exemple, un système de péages peut être proposé dans un dossier candidat qui serait relativement éloigné de la tarification d'usage de l'infrastructure pratiquée jusque là et le surplus global du projet peut s'en trouver modifié car la demande finale dépend de la tarification. Le numérateur du ratio n'est donc pas, en fait, indépendant du système de financement et la mise en œuvre de l'évaluation suppose, dans ces conditions, que soient précisées les modalités de calcul (avant appel d'offre) de la VAN socio-économique³ et du taux de subvention. Ce calcul doit-il se faire sur la base d'une tarification standard ? d'une tarification qui optimise le TRI ? d'une tarification qui optimise le ratio $VAN_e/\text{subvention}$?... Si

2. Cf. le chapitre 4.

3. La valeur empirique de la VAN socio-économique est, dans les circulaires ministérielles, « le bénéfice socio-économique actualisé » ou BNA. Ce chapitre ne propose pas de revenir sur les règles de calcul pratique du bénéfice actualisé, et ne considérera que sa valeur théorique : la VAN socio-économique.

cette dernière hypothèse est cohérente avec la fonction objectif de l'État, la tarification correspondante reste à établir.

À supposer que cela soit fait, la conséquence logique serait que cette tarification soit imposée par le cahier des charges de l'appel d'offre. Cela pose évidemment la question de la relation entre méthode d'évaluation et dispositif de sélection des opérateurs et, plus généralement, celle du rôle dévolu aux calculs de rentabilité financière.

La réponse à la question consisterait à préciser comment (ou selon quelles variantes) pourrait être établi le calcul du taux de subvention en comparaison avec les calculs actuellement pratiqués. À défaut de ce genre de résultat, nous supposons le problème résolu ou, plus précisément, nous lui proposerons, *dans une première section*, une solution qui consiste à établir, sur la base d'un projet standard, la relation entre le taux de subvention qui assure la rentabilité financière d'un projet et les paramètres économiques de ce projet standard en nous appuyant sur une formalisation déjà publiée (Bonnaïfous, 2002).

Nous montrerons *dans une deuxième section* que, sur cette base, une programmation « optimale » peut être recherchée par simulation numérique et comparée aux programmations qui résultent des différents ratios (ordre des TRE, ordre des TRI, ratio VAN_e /subvention). Limité au cas particulier de 17 projets autoroutiers, cet exercice apporte, dans ce cas précis, une forme de preuve (tout du moins au sens des physiciens) de l'efficacité du ratio pour la détermination du programme optimal. Cette section s'appuie également sur des résultats déjà publiés (Bonnaïfous et Jensen, 2005).

Une première démonstration formelle de cette propriété sera proposée *dans la troisième section*. Celle-ci a le mérite d'explicitier les hypothèses sous lesquelles l'usage du ratio conduit bien à un programme qui optimise le rendement social. Ces hypothèses ne sont bien évidemment pas toujours réalistes. C'est une première approche qui sera généralisée et détaillée dans le chapitre 9 de Joël Maurice et discuté par les simulations numériques d'Emile Quinet et Alain Sauvant au chapitre 8, dans le prolongement des travaux d'Abraham et Laure (1959).

Une *quatrième section* mettra l'accent sur les conséquences de la prise en compte du coût d'opportunité des fonds publics, sur la tarification des péages d'infrastructure. C'est un exemple d'application, qui met en évidence les conséquences multiples du coût de rareté des fonds publics

Nous concluons ce chapitre par les quelques préconisations concrètes, qui peuvent être raisonnablement proposées à l'issue de

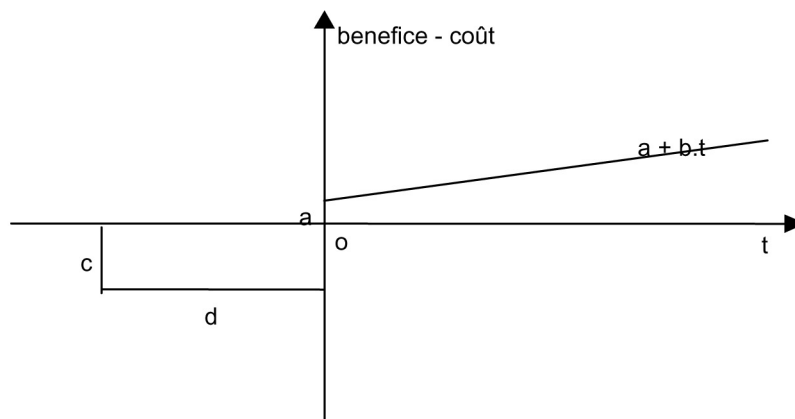
ce travail, concernant l'évolution des critères d'organisation des programmes.

1. LE BESOIN DE SUBVENTION

Pour calculer la fonction qui établit le taux de subvention nécessaire à un projet, supposons que celui-ci correspond à la chronique stylisée des coûts et des bénéfices financiers, représentée sur la figure 1. Si la mise en service est supposée réalisée à la date $t = 0$, la dépense annuelle entre les dates $-d$ et 0 est de c .

À partir de la mise en service, le bénéfice financier dégagé est supposé de la forme $(a + b \cdot t)$ ⁴.

Figure 1
La fonction bénéfice-coût



Le taux de rentabilité interne (TRI) du projet, c'est à dire le taux d'actualisation qui annule sa valeur actualisée nette financière (VAN_f), est alors une fonction des quatre paramètres c , d , a et b . Il est à comparer au taux de rentabilité qu'un opérateur (public ou privé) est en droit d'escompter et qu'il faudra, d'une manière ou d'une autre lui assurer. Ce pourrait être avec un système de péage virtuel (*shadow toll*). Nous supposons plus banalement que c'est à l'aide d'une subvention d'investissement.

4. Les calculs qui suivent se transposent sans difficulté avec une fonction exponentielle et les analyses qui en résultent n'en sont pas radicalement modifiées. En particulier,

$$\int_0^T e^{-rt} dt \approx \sum_{t=1}^T (1+r)^{-t} \text{ pour les valeurs de } r \text{ et de } T \text{ rencontrées actuellement dans les cas pratiques.}$$

Nous notons :

- α le taux d'actualisation utilisé pour calculer la VAN_f ,
- α_0 le taux d'actualisation qui annule la VAN_f du projet, c'est à dire son TRI,
- δ est le supplément de TRI que la subvention apporte à l'opérateur,
- τ est le taux de subvention de l'investissement, soit la part de c financée par subvention. Nous ferons abstraction du cas des versements de fonds publics récurrents, comme dans le cas des PPP où le versement public est en quelque sorte en viager.

Pour un taux d'actualisation α , et un bilan actualisé calculé des dates $-d$ à T , la valeur actualisée nette du projet s'écrit :

$$VAN_f = \int_{-d}^0 -c.e^{-\alpha.t}.dt + \int_0^T (a + b.t).e^{-\alpha.t}.dt \quad (1)$$

Nous supposons que l'actualisation est étendue à l'infini, ce qui est sans conséquence sur les résultats qui nous intéressent en raison du faible poids du futur lointain et, surtout, de la convergence des fonctions intégrales de l'équation (1). Cette équation devient alors⁵ :

$$VAN_f = \frac{1}{\alpha} \left[c(1 - e^{\alpha.d}) + a + \frac{b}{\alpha} \right] \quad (2)$$

Le TRI du projet, α_0 , qui annule cette VAN_f vérifie alors :

$$c(1 - e^{\alpha_0.d}) + a + \frac{b}{\alpha_0} = 0 \quad (3)$$

Un taux de subvention τ abaisse le coût annuel de construction c à $c(1-\tau)$ et porte le TRI α_0 à $(\alpha_0 + \delta)$ de sorte l'équation (3) devient :

$$(1 - \tau)c(1 - e^{(\alpha_0 + \delta).d}) + a + \frac{b}{\alpha_0 + \delta} = 0 \quad (4)$$

Dont nous pouvons déduire l'expression du taux de subvention :

$$\tau = 1 - \frac{a(\alpha_0 + \delta) + b}{c(\alpha_0 + \delta)(e^{(\alpha_0 + \delta).d} - 1)} \quad (5)$$

La relation entre τ , le taux de subvention, et δ , l'augmentation du TRI du projet qu'il s'agit d'apporter à l'opérateur, dépend, évidem-

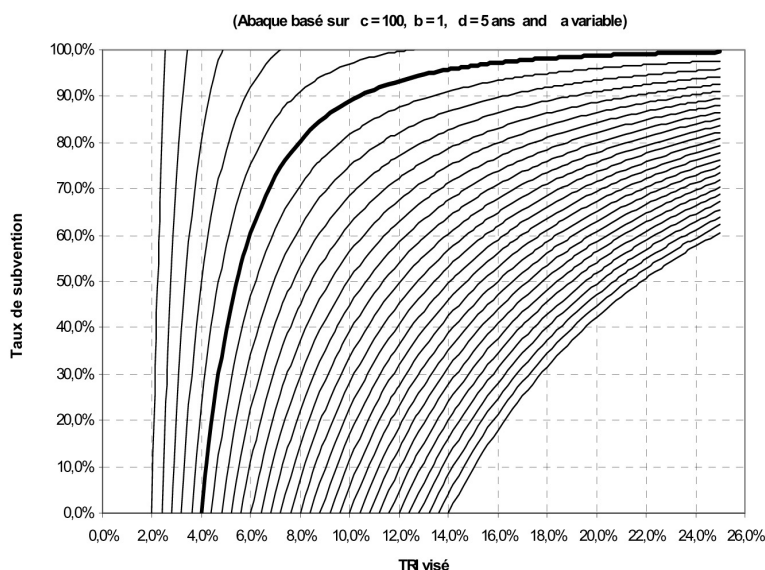
5. Les détails des calculs sont présentés dans la présentation initiale de cette formalisation (Bonnafous, 2002).

ment, des paramètres c , d , a , b et, bien sûr, de α_0 , qui caractérisent la rentabilité financière initiale du projet. Ces paramètres sont, en outre, liés entre eux par l'équation (3) qui définit le TRI du projet α_0 (ou ce qui est équivalent, $\tau = 0$ si et seulement si $\delta = 0$). Cela implique quelques difficultés dans la représentation de la fonction (5).

Si nous voulons représenter cette fonction, il nous faut figer certains de ces cinq paramètres et ne faire varier que ceux dont nous souhaitons exhiber le rôle. Nous aurons recours, pour cela, à la technique classique des abaques. Nous ne représenterons qu'un seul de ces abaques (figure 2), qui sera suffisant pour illustrer notre propos. Le coût annuel de construction c a été fixé à une valeur normée de 100, la durée de cette construction est fixée à 5 ans. L'accroissement annuel des avantages nets des coûts, b , est supposé égal à 1. Cela revient à faire varier le TRI initial du projet α_0 (ou encore a , puisque α_0 ne dépend plus que de a , bénéfice net du projet à la date de mise en service).

Figure 2

Taux de subvention et amélioration du tri



L'abaque ci-dessous présente ainsi la fonction (5) sous contrainte de l'équation (3) pour une série de valeurs de α_0 , entre 2% et 14 % avec un pas de 0,4% (ce qui correspond à des valeurs croissantes de a de gauche à droite sur le graphique). Pour chacune de ces valeurs de α_0 , qui se lit en abscisse, chaque courbe exprime le taux de subvention nécessaire pour élever le TRI au niveau $(\alpha_0 + \delta)$.

Si elles reposent sur les caractéristiques de la série chronologique des coûts et avantages du projet considéré et sur les valeurs particulières retenues pour les paramètres b , c et d , les silhouettes de ces courbes ont un caractère général. En particulier, leur concavité résulte des propriétés de la dérivée seconde de la fonction (5) qui peuvent être démontrées sous la condition que b soit positif. De cette concavité résultent des conséquences importantes quant à l'alternative entre un opérateur public et un partenariat privé qui ont été précisées par ailleurs (Bonnafous, 2002).

Ce qui nous intéresse ici est simplement le fait que si l'on accepte la silhouette stylisée des coûts et bénéfices de la figure 1, et si l'on a des estimateurs de b et de d (le paramètre c pouvant être normé à 100) nous disposons alors d'une détermination fonctionnelle du niveau de subvention nécessaire au projet pour amener son TRI financier au niveau requis.

De plus, les mêmes équations peuvent être utilisées pour les calculs socio-économiques à la condition de connaître les valeurs numériques des paramètres (ce ne sont plus les mêmes a et b , ni même peut-être le même c) ou de pouvoir les reconstituer. En particulier, l'équation (2) nous livre alors la valeur actualisée nette socio-économique du projet, soit VAN_e , pour un taux d'actualisation α qui est ici le taux d'actualisation du Plan.

Dans la section qui suit, ces paramètres étaient connus ou ont pu être reconstitués pour les 17 projets traités, tant pour le calcul financier que pour le calcul socio-économique, de sorte qu'étaient disponibles pour chacun d'eux leur TRI, leur TRE, la VAN_e dégagée et le montant de subvention.

2. QUEL CRITÈRE POUR UNE PROGRAMMATION « OPTIMALE » ?

Nous exprimerons le problème d'optimisation d'une manière fort simple : *Un nombre fini de projets étant évalués, quel en est le programme optimal, c'est à dire l'échéancier de réalisation de ces projets qui, sous une contrainte de capacité de financement public, maximise la VAN_e totale.*

L'exercice a consisté à traiter un ensemble de 17 projets d'autoroutes à péage pour lesquels les éléments nécessaires à ces simulations étaient connus et évalués selon des méthodes homogènes⁶ ou ont pu

6. La liste de ces projets est présentée en annexe. Il s'agit des projets du réseau français qui étaient en concurrence au début des années 90. Plusieurs ont été réalisés depuis avec une faible

être reconstitués à l'aide du modèle présenté dans la section 1. Ainsi, les taux de subvention ont été calculés sur la base des équations (3) et (5). Il a été supposé, que les projets étaient réalisés sous une contrainte budgétaire limitant le financement public lors de la première année du programme, ce financement étant supposé augmenter ensuite de 2,5 % par an. La contrainte initiale a été considérée à plusieurs niveaux de manière à observer si elle joue un rôle différent selon qu'elle est très forte ou, au contraire, desserrée.

Si l'on peut avancer de manière intuitive qu'un ordre de programmation fondé sur les valeurs décroissantes du ratio VAN_e par euro de subvention doit produire la meilleure VAN_e globale possible, il convient de tenter d'apporter une démonstration à ce principe d'optimisation. Comme la série des projets considérés au précédent paragraphe peut donner lieu à 17 ! permutations possibles (environ 10^{34}), les algorithmes d'exploration combinatoire peuvent être raisonnablement convoqués. Ils semblent apporter quelques résultats dans les sciences expérimentales qui sont confrontées à ce genre de problème, par exemple dans les modèles évolutionnaires de la biologie ou dans l'exploration des nanostructures. C'est à l'une de ces méthodes que nous avons eu recours. Elle a été expérimentée (Bonnafous et Jensen, 2004⁷) en retenant ce « paquet » des 17 projets.

La fonction objectif reste, bien entendu, la VAN_e globale du programme de projets réalisé sous contrainte de financement public (actualisée à l'instant zéro où l'on décide le programme). L'algorithme est issu de la méthode des recuits simulés (van Laarhoven, 1992) et consiste, en partant d'un ordre aléatoire des projets, à simuler des permutations entre deux projets, eux-mêmes étant désignés par un processus aléatoire. Toute permutation qui augmente la fonction objectif, ou qui ne la diminue que d'un faible niveau (niveau qui décroît au cours de la simulation, pour assurer la convergence), est conservée.

La contrainte de financement public F est introduite de la manière suivante. Pour chaque projet individuel, nous calculons d'abord une durée de réalisation en tenant compte de la valeur de F . Cette durée, calculée en utilisant l'équation (5), rend son financement possible dans

contribution apparente des finances publiques car ils ont été confiés à des sociétés d'autoroute qui ont pu utiliser le mécanisme d'adossement (cf. section 1).

7. Dans cet article cité en référence, le résultat correspondant à celui qui est présenté sur la figure 3 ci-après est légèrement différent en raison d'une faiblesse de l'algorithme de « remplissage » de la contrainte financière qui a été corrigé entre temps.

ce laps de temps, c'est-à-dire que le montant annuel de subvention nécessaire à la réalisation du projet est alors au plus égal à F^8 . Nous pouvons alors entamer la simulation du programme pour cette valeur de F . Le programme « prend » le premier projet. S'il reste un reliquat de subvention (cas où la subvention nécessaire à ce premier projet serait inférieure à F), il est affecté au démarrage du deuxième projet. À la fin de la première année, le premier projet avance d'une année, et le deuxième d'un pourcentage d'année, égal à la proportion de sa subvention annuelle couverte par le reliquat laissé par le premier projet. En principe, pour un projet donné, chacun des indicateurs TRE, TRI, VAN/I est fonction de la date de mise en service. Ces variables sont recalculées à chaque fois en fonction de la date à laquelle l'algorithme les désigne.

La fonction objectif se stabilise après quelques milliers d'itérations et toutes les expériences, menées avec des ordres initiaux et des paramètres variables, voient la fonction objectif converger vers la même valeur.

Nous ne retiendrons ici que le principal résultat de cet exercice qui est restitué sur la figure 3 ci-après. Il représente, en abscisse, la contrainte financière rapportée à une valeur moyenne et, en ordonnée, le surcroît de valeur actualisée nette par rapport au programme ordonné selon les TRE des projets.

Ces résultats confirment que ce gain de rendement social est nettement positif avec l'ordre du TRI et cela d'autant que la contrainte budgétaire est resserrée. Mais ils nous confirment surtout (nous démontrant au sens des physiciens) que l'ordre du ratio VAN_e / euro public investi apporte un gain de rendement social supérieur encore et que c'est bien le critère qui désigne le programme optimal.

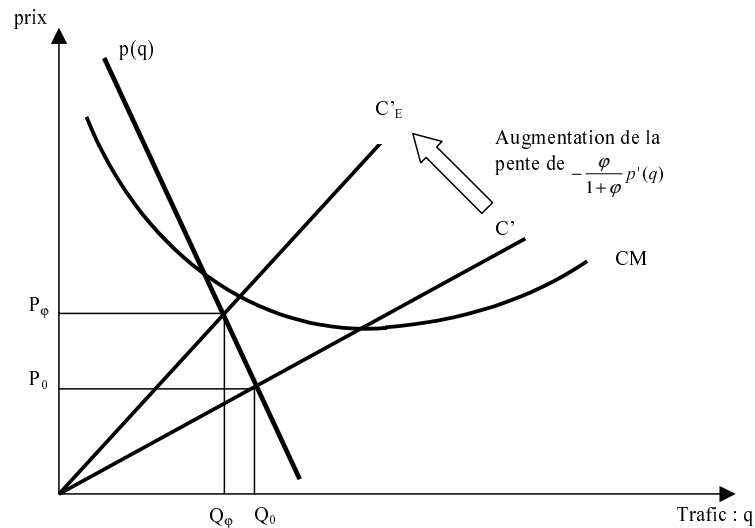
3. UNE DÉMONSTRATION FORMELLE

Dans cette section, nous montrerons la pertinence du ratio entre l'utilité sociale d'un projet et son besoin de subvention, dans le cadre d'un modèle simple. Nous mettrons notamment en évidence que c'est le coût d'opportunité des fonds publics qui fonde l'existence d'un ordre optimal de réalisation selon le critère de la VAN par euro public investi.

8. Tant que F le permet, il n'y a pas de durée pluriannuelle de réalisation imposée par la technique.

Figure 3

Rendement social comparé des programmes ordonnés selon les TRI (IRR), le ratio VAN/euro public (ϕ) et l'optimum établi par les simulations sous différentes contraintes de financement



Nous supposons dans ce qui suit que le décideur a le choix entre n projets ; chaque projet i , avec $i = 1, \dots, n$, est caractérisé par un couple de paramètres :

1) VAN_i : La valeur actualisée nette du projet i ; avec $VAN_i > 0$, $\forall i = 1, \dots, n$.

Pour évaluer les programmes d'infrastructure, il faut bien sûr être en mesure d'évaluer la valeur sociale de chaque projet. Ces évaluations, dont la base théorique est le calcul économique (Lesourne 1972, Bloy & al. 1976), conduisent à une Analyse Coûts-Avantages (ACA). En output, un « bilan socio-économique » réunit les nombreux bénéfices et les multiples préjudices et coûts de chaque projet.

Rares sont en pratique les projets parfaitement isolés de toute autre décision d'investissement. Les projets peuvent ne pas être indépendants les uns des autres et avoir des effets significatifs sur la VAN des autres. Nous supposons, dans le modèle qui suit, que ces effets sont toujours marginaux, que les ACA fournissent une assez bonne approximation de leur valeur sociale. En fait, ce problème est *a priori* traité lors de la constitution des projets eux-mêmes. Avec un peu de bon sens, les investissements fortement complémentaires seront réunis au sens du même projet. De même, l'étude des substituabilités les plus fortes se fait en amont, notamment lors du choix du tracé définitif de la

voie de circulation parmi les différentes alternatives possibles. En tout état de cause, c'est à l'étape de l'élaboration des projets qu'intervient majoritairement la question de la complémentarité et de la substituable des investissements. En pratique, il conviendra de porter une attention particulière sur ce point puisqu'il n'est pas traité par le modèle.

2) S_i : Le besoin de subvention du projet i ; avec $S_i > 0, \forall i=1, \dots, n$.

Le besoin de subvention est un montant qui représente tout ou partie des frais d'investissements (voire d'exploitation), selon la rentabilité financière du projet concerné. Il s'agit par exemple du montant nécessaire pour que l'État puisse amener une entreprise à s'engager dans un contrat BOT (*Built Operate and Transfer*) sur un segment d'autoroute. Il est pratique ici de considérer que les dépenses publiques sont concentrées l'année de mise en service, bien que ce ne soit pas systématique.

Nous supposons que la fonction objectif de la collectivité correspond à la maximisation du surplus global dégagé par les projets :

$$\underset{x}{\text{Max}} \quad W(x) = \sum_{i=1}^n x_i \text{VAN}_i$$

Les x_i ont une valeur nulle lorsque le projet n'est pas réalisé et égale à l'unité lorsque le projet est réalisé en totalité. Nous supposons qu'il est possible de réaliser partiellement un projet, le x_i correspondant étant alors compris entre 0 et 1. Pour un projet k réalisé partiellement, les valeurs $x_k \cdot \text{VAN}_k$ et $x_k \cdot S_k$ retenues sont supposées proportionnelles à VAN_k et S_k en fonction de x_k . Nous verrons plus tard que cette hypothèse tout à fait théorique, de réalisation partielle des projets, n'est pas très contraignante.

$$0 \leq x_i \leq 1, \forall i = 1, \dots, n$$

Lorsque l'optimisation se fait sans contrainte de rareté des budgets publics, la création de valeur maximale est atteinte en mettant en œuvre tous les projets dont la VAN est positive. C'est un résultat trivial.

Nous introduisons maintenant un budget (B) positif, qui plafonne les dépenses d'investissement.

$$\sum_{i=1}^n x_i S_i \leq B$$

On obtient au total le système d'optimisation sous contraintes suivant⁹.

$$\begin{aligned} \underset{x}{\text{Max}} \quad & W(x) = \sum_{i=1}^n x_i VAN_i \\ \text{s.c.} \quad & \begin{cases} B - \sum_{i=1}^n x_i S_i \geq 0 \\ x_i \geq 0, \quad \forall i = 1, \dots, n \\ 1 - x_i \geq 0, \quad \forall i = 1, \dots, n \end{cases} \end{aligned}$$

Cette modélisation possède l'avantage d'être relativement simple et de nature à faciliter l'interprétation des résultats. Elle contient une hypothèse décisive (concernant la réalisation partielle des projets) qui nous autorise à écrire un lagrangien sans postuler une relation fonctionnelle entre les paramètres VAN et S . Les $2n$ contraintes sur les x_i permettent de transformer le problème de choix discret, pour le résoudre de manière analytique.

En pratique, les projets sont soit réalisés en totalité, soit ne le sont pas du tout. La possibilité de réaliser un projet partiellement, tout en bénéficiant proportionnellement de ses caractéristiques (VAN et S), est essentiellement un outil théorique. Cela dit, la réalisation partielle d'un projet ne sera *in fine* que marginale dans la solution du programme. Un projet au plus sera dans ce cas. Le vecteur solution x^* est constitué d'un ensemble de 1 (projets réalisés), d'un ensemble de 0 (projets non réalisés), et d'une valeur comprise entre 0 et 1 pour le projet « limite » x_k . En supposant que les projets sont ordonnés selon leur priorité de réalisation, on a :

$$x^* = \left(\underbrace{1, \dots, 1}_{\text{projets acceptés}}, x_k, \underbrace{0, \dots, 0}_{\text{projets rejetés}} \right)$$

Puisque les caractéristiques attachées à chaque projet sont fixes, si la réalisation d'un projet doit se faire pour augmenter de la meilleure

9. Cette manière de formaliser le problème est proche de celle de Weingartner (1963), lorsqu'il se proposa de résoudre le problème fondamental posé par Lorie & Savage (1955). Toutefois, son travail de résolution s'est concentré sur l'approche en termes de programmation linéaire. Bertonèche & Langohr (1977) offrent une intéressante mise en perspective de ces travaux fondateurs. Pour une synthèse sur les modèles de programmation linéaire, voir Babusiaux (1990). Or si la programmation linéaire fournit un outil de résolution ponctuel efficace (pour un ensemble de projet donné), elle ne permet pas de faire apparaître un résultat théorique général.

façon le bien-être collectif, alors le projet sera fait en totalité. Le meilleur projet sera réalisé en totalité, puis le deuxième, le troisième etc... jusqu'à saturation de la contrainte budgétaire. C'est le projet « limite » qui risque de ne pas être terminé, et seulement celui-là.

Dans les cas réels, les projets d'infrastructure nécessitent plusieurs années de réalisation, ce qui est potentiellement le lieu d'un ajustement entre le budget annuel et les projets financés. On peut aussi considérer que le budget B n'est fixé par le gouvernement que dans son ordre de grandeur. Des ajustements ponctuels et marginaux sont toujours possibles. Il existe des marges de manœuvre pour éviter les absurdités qu'engendreraient les discontinuités. Pour toutes ces raisons, le choix de formaliser les x_i de manière continue entre 0 et 1 nous paraît acceptable, et le cas du projet k , ne pas poser de problèmes.

Le Lagrangien généralisé incluant les contraintes de non-négativité des x_i que nous optimisons s'écrit :

$$L(x_1, \dots, x_n, \varphi, \alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n) = \sum_{i=1}^n x_i VAN_i + \varphi \left(B - \sum_{i=1}^n x_i S_i \right) + \sum_{i=1}^n \alpha_i x_i + \sum_{i=1}^n \beta_i (1 - x_i)$$

Les conditions de Kuhn et Tucker impliquent notamment¹⁰ :

$$VAN_i - \varphi S_i + \alpha_i - \beta_i = 0, \forall i = 1, \dots, n \quad (6)$$

$$\varphi \left(B - \sum_{i=1}^n x_i S_i \right) = 0 \quad (7)$$

$$\alpha_i x_i = 0, \forall i = 1, \dots, n \quad (8)$$

$$\beta_i (1 - x_i) = 0, \forall i = 1, \dots, n \quad (9)$$

$$\varphi, \alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \geq 0$$

Les multiplicateurs de Kuhn et Tucker sont non négatifs par définition. Ils sont positifs (ou exceptionnellement nuls) lorsque la contrainte qui leur est associée est saturée et nuls lorsque que cette contrainte est non saturée. Pour le projet « limite » k , vérifiant $0 < x_k < 1$, on a d'après (6) :

$$VAN_k - \varphi S_k = 0 \Leftrightarrow \frac{VAN_k}{S_k} = \varphi \quad (10)$$

10. Les conditions suffisantes d'existence d'un maximum unique sont vérifiées sous les hypothèses du modèle.

Dans le cas d'un desserrement marginal de la contrainte budgétaire à l'optimum, c'est uniquement la réalisation du projet k qui va être affectée. On peut donc écrire les différentiels suivants à l'optimum :

$$dB = S_k dx_k \quad (11)$$

$$dW = VAN_k dx_k \quad (12)$$

En combinant les équations (10), (11) et (12) on obtient :

$$\frac{dW}{dB} = \varphi$$

φ est dans cette dernière formulation la variation de surplus collectif induite par un desserrement de la contrainte de disponibilité des fonds publics. C'est le montant maximum de surplus que la collectivité peut espérer faire correspondre à une unité budgétaire supplémentaire.

Le multiplicateur φ est associé à la contrainte budgétaire, il représente le coût de rareté des fonds publics. φ résume la contrainte budgétaire à la manière d'un prix (prix dual de la contrainte). Plus la contrainte budgétaire est forte, plus φ est élevé. C'est aussi le coût d'opportunité des ressources budgétaires.¹¹

À partir des conditions de premier ordre, on a :

$$VAN_i - \varphi S_i = \beta_i - \alpha_i$$

Pour les projets acceptés (indités A), les conditions de Kuhn et Tucker impliquent :

- la contrainte « $-x_i \leq 0$ » n'est pas saturée, on a donc $\alpha_i = 0$
- la contrainte « $x_i - 1 \leq 0$ » est saturée, on a donc $\beta_i > 0$

$$\text{D'où } VAN_A - \varphi S_A > 0 \Leftrightarrow \boxed{\frac{VAN_A}{S_A} > \varphi} \quad (13)$$

L'ensemble des projets acceptables n'est pas composé de l'ensemble de ceux disposant d'une VAN positive. Ne sont retenus que ceux ayant un ratio VAN/S supérieur au coût de rareté de l'argent public φ .

11. Il est important de ne pas assimiler ce coût de rareté des fonds publics (que l'on peut aussi interpréter comme un coût d'opportunité) au coût social des fonds publics. Le coût social des fonds publics (*shadow cost of public funds*) est formé par les coûts de la collecte de l'impôt et les distorsions de prix associés.

Pour les projets rejetés ou différés (indicés R), les conditions d'optimisation impliquent :

- La contrainte « $-x_i \leq 0$ » est saturée, on a donc $\alpha_i = 0$
- La contrainte « $x_i - 1 \leq 0$ » n'est pas saturée, on a donc $\beta_i > 0$

$$\text{D'où } \boxed{\frac{VAN_R}{S_R} < \varphi} \quad (14)$$

Les projets rejetés ont un ratio VAN/S inférieur au coût de rareté des fonds publics.

$$\text{Au total, (13) et (14) impliquent : } \frac{VAN_A}{S_A} > \varphi > \frac{VAN_R}{S_R}$$

Les projets acceptés doivent toujours avoir un ratio VAN/S supérieur à celui des projets rejetés. On privilégie les projets produisant le plus de valeur par euro public dépensé.

Si on fait varier φ (desserrement de la contrainte budgétaire) on construit une hiérarchisation totale des priorités selon le critère VAN/S. Pour optimiser le surplus global sous contrainte budgétaire, lorsqu'on ignore φ , il faut choisir en priorité les projets disposant d'un ratio « utilité collective par euro public investi » le plus élevé.

Coût d'opportunité vs coût social

Il est important de ne pas identifier le « coût d'opportunité des fonds publics » au « coût social des fonds publics ». Ce sont deux coûts aux origines économiques fondamentalement différentes. Le coût social des fonds publics (*shadow cost of public funds*) regroupe les coûts de la collecte de l'impôt et des distorsions de prix associées. Dans les procédures actuelles de constitution des bilans socio-économiques, la prise en compte du coût social des fonds publics implique une pondération des apports publics par un facteur qui est estimé dans le rapport Lebègue (CGP 2005) à 0,3. L'idée est de prendre en compte le coût total pour l'économie d'un subventionnement.

Le coût d'opportunité des fonds publics considère, pour sa part, l'utilisation alternative que pourrait avoir les fonds publics lorsqu'ils sont rares. C'est le résultat d'une optimisation avec une condition supplémentaire : la contrainte budgétaire (et d'endettement). Le coût d'opportunité est le paramètre d'un optimum de second rang, alors que le coût social des fixe les conditions d'un optimum de premier rang. Ce sont deux concepts biens différents.

Si un projet requiert 100 € de subventions publiques pour une VAN de 50 €. La prise en compte du coût social des fonds publics conduit à corriger la valeur sociale du projet qui est en fait de 20 € (en reprenant le taux du rapport Lebègue). La VAN corrigée du coût social des fonds publics est positive, ce projet doit être réalisé. Notons que ce projet a un ratio VAN/S de 0,5. Dans une optique de premier rang, ce projet doit être réalisé car le bénéfice par euro public est supérieur au coût social des fonds publics de 0,3.

Toutefois, si les fonds publics sont rares, la réalisation de ce projet peut priver de subventions un projet rapportant 1€ de VAN par euro public dépensé, c'est son coût d'opportunité. Il n'est plus souhaitable de réaliser le projet car il engendre un différentiel de VAN de -50 €.

Dans notre exemple, le coût d'opportunité généré par la contrainte budgétaire domine le coût social des fonds publics, c'est aussi le cas de figure auquel nous devons faire face. Il n'est jamais souhaitable de réaliser un projet dont le ratio VAN/S est inférieur à 0,3. Mais la prise en compte du coût social des fonds publics devient inutile lorsque la contrainte budgétaire est forte. Si le coût d'opportunité de la collectivité est de 1€, les projets dont le ratio VAN/S se situe entre 0,3 et 1 seront de toute façon rejetés.

En première approximation, la meilleure stratégie est d'abord de financer les projets apportant le plus de VAN, sans introduire le besoin de subvention. Mais c'est oublier qu'un projet rapportant beaucoup mais mobilisant la plupart des crédits limite d'autant la satisfaction procurée par de nombreuses « petites valeurs » obtenues pour un faible prix, dans un monde où les financements publics sont rares. C'est typiquement le problème soulevé par la ligne à grande vitesse (LGV) qui supportera le TGV entre Paris et Strasbourg. L'énorme besoin de subvention lié à ce projet a conduit à différer de plusieurs années d'autres projets, plus petits mais dont le ratio VAN/S est nettement meilleur comme la LGV Rhin-Rhône. Le critère normatif de la « VAN par euro public investi » est un ratio qui prend en compte la valeur créée (VAN), mais relativement à son prix (S).

4 TARIFICATION : LA PRISE EN COMPTE DU COÛT DE RARETÉ DES FONDS PUBLICS

Les sections précédentes ont mises en évidence le rôle du coût de rareté, ou coût d'opportunité, des fonds publics dans la hiérarchisation des projets. La contrainte budgétaire est fixe, mais sachant que la tarification peut affecter sensiblement la VAN des projets, comment s'assurer que la tarification est optimisée pour favoriser le critère de la VAN par euro public investi ?

Cette section met en avant l'utilité et l'éclairage que ce concept peut apporter dans la discussion sur la tarification des infrastructures. Nous verrons, par un modèle très simple¹², que l'existence d'un coût d'opportunité des fonds publics n'est pas neutre sur la tarification optimale des infrastructures. Nous souhaitons ici mettre l'accent sur les principes de financement et de tarification qu'il implique, concernant notamment la répartition du coût de l'infrastructure entre usagers et contribuables.

12. En particulier, le rationnement des fonds public n'est pas modulé dans le temps (cf. chapitre 8).

Le surplus des usagers est classiquement matérialisé par la surface V située sous la courbe de demande. Dans le cas où un péage est payé, ce surplus brut est diminué des recettes. En supposant qu'il existe une fonction de demande inverse $p(q)$, décroissante et convexe, pour tout trafic q sur l'infrastructure considérée, le surplus net des consommateurs est :

$$U(q) = V(q) - q \cdot p(q), \text{ avec } V(q) = \int_0^q p(x) dx \text{ et } V'(q) = p(q)$$

Supposons l'État participe aussi au financement de l'infrastructure pour un montant égal à la différence entre le coût et les recettes payées par les usagers. Nous postulons une fonction de coût croissante et convexe en q .

$$E(q) = C(q) - q \cdot p(q)$$

Ce montant de subventions a un coût d'opportunité φ strictement positif si les finances publiques sont contraintes, comme nous l'avons justifié précédemment. L'État est supposé maximiser le surplus global W en fonction du trafic :

$$\begin{aligned} \underset{q}{\text{Max}} W(q) &= U(q) - (1 + \varphi)E(q) = V(q) - q \cdot p(q) - (1 + \varphi)[C(q) - q \cdot p(q)] \\ &= V(q) + \varphi q \cdot p(q) - (1 + \varphi)C(q) \end{aligned}$$

Le maximum de la fonction précédente est obtenu à la condition que :

$$(1 + \varphi)p(q) + \varphi q \cdot p'(q) - (1 + \varphi)C'(q) = 0$$

Cette condition peut être réécrite sous deux formes :

$$- p(q) = C'(q) - \frac{\varphi}{1 + \varphi} q \cdot p'(q) \quad (1)$$

$$- \frac{p(q) - C'(q)}{p(q)} = - \frac{\varphi}{1 + \varphi} \frac{1}{\varepsilon_{q/p}} \quad (2)$$

où $\varepsilon_{q/p}$ est l'élasticité-prix de la demande

L'équation (2) est très ressemblante à la formule de la tarification Ramsey-Boiteux. Toutefois, il ne faut pas s'y tromper, il n'est pas ici question d'optimum de second rang ; le financement de l'activité par voie fiscale n'est pas nul à l'optimum. Si la contrainte budgétaire est inexistante, ie $\varphi = 0$, les équations impliquent bien une tarification de premier rang au coût marginal. Remarquons tout de même que la tarification Ramsey-Boiteux n'est jamais que l'expression d'une

contrainte budgétaire resserrée au maximum, c'est-à-dire un cas particulier des situations que nous étudions ici.

Ce que disent très simplement ces deux équations, c'est que le niveau du péage doit s'écartier d'autant plus du coût marginal C' , que le coût d'opportunité des fonds publics est élevé. L'utilisateur doit être d'autant plus mis à contribution que φ est élevé. Et ceci est d'autant plus vrai que l'élasticité-prix est faible.

L'équation (1) décrit l'équation d'offre « économique », qui inclut le coût d'opportunité des fonds publics. Cette formulation du « coût marginal économique » intègre le coût marginal « technologique » C' et l'impact d'une mobilisation des rares fonds publics par rapport à un financement par les usagers eux-mêmes.

$$\text{Soit } C'_E(q) \equiv C'(q) - \frac{\varphi}{1+\varphi} q \cdot p'(q) \quad (p'(q) < 0)$$

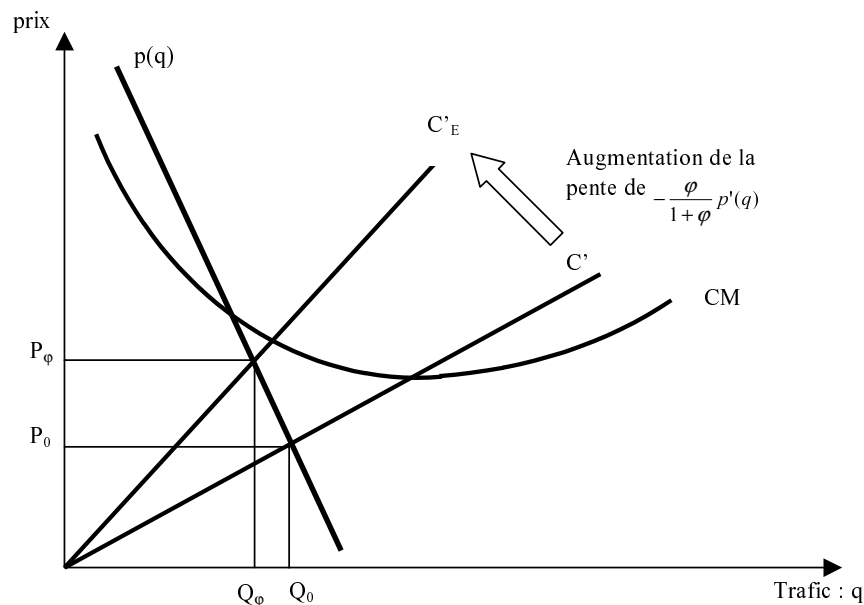
La figure 4 permet de visualiser l'arbitrage dans le financement entre usagers et contribuables, dans un cas stylisé caractérisé par :

$$- CT(q) = I + c \cdot q^2, \text{ avec } I > 0 \text{ et } c > 0$$

$$- P(q) = -a \cdot q + b, \text{ avec } a > 0 \text{ et } b > 0$$

Figure 4

Tarification optimale et rareté des fonds publics



Concluons cette rapide exploitation du concept de coût d'opportunité des fonds publics par une réflexion générale sur la tarification. La formulation que nous venons de présenter réunit deux conceptions souvent opposées par les économistes : le péage en tant qu'instrument allocatif et son utilisation comme un moyen de financement. Les résultats de ce rapide modèle montrent que le péage peut répondre aux deux types d'attente sans contradiction, que l'articulation entre ces deux objectifs doit se faire sur la base du coût de rareté des fonds publics. Le financement par péage est allocativement efficace si l'arbitrage usager-contribuable se fait selon le niveau de φ et l'élasticité de la demande au péage.

5 CONSÉQUENCES PRATIQUES : FAUT-IL ABANDONNER LE CRITÈRE DE LA RENTABILITÉ IMMÉDIATE ?

Le problème fondamental des programmes est de déterminer la date optimale de mise en service de chaque projet candidat. Sans contrainte budgétaire et dans le cas d'hypothèses standards (investissements indépendants, avantages non décroissants...), Abraham et Laure (1959) ont montré que l'année optimale de mise en service d'un projet intervient lorsque le rapport entre les avantages socio-économiques qu'il procure l'année de mise en service et le coût d'investissement (ie le taux de rentabilité immédiate) atteint le niveau du taux d'actualisation. Ce critère a été largement mis en œuvre pour les praticiens français, qui approximent la date optimale par le taux de rentabilité immédiate¹³.

La démonstration de ce résultat qui nous semble la plus élégante est proposée par Quinet (1998, p. 237), et s'obtient en supposant une actualisation continue. À la date de mise en service t^* , on a :

$$VAN(t^*) = -I.e^{-\alpha t^*} + \int_{t^*}^{+\infty} A(t)e^{-\alpha t} dt \quad (2)$$

où $A(t)$ est une fonction des avantages nets aux dates t .

13. La méthode utilisée pour arrêter le programme du 5^e Plan était par exemple de classer les investissements par taux de rentabilité immédiate décroissant, et de s'arrêter quand on avait épuisé la contrainte budgétaire, ce qui revenait à appliquer ce principe.

La condition nécessaire de maximisation de cette fonction est :

$$\frac{dVAN(t^*)}{dt} = 0 \Leftrightarrow \boxed{\frac{A(t^*)}{I} = \alpha} \quad (3)$$

Le problème était donc relativement simple à l'époque de l'article d'Abraham et Laure : dès lors que la rentabilité immédiate d'un projet a dépassé le taux d'actualisation, il faut le réaliser dès que possible¹⁴. Et il est vrai que les hypothèses précédentes de modélisation semblent relativement acceptables dans le cas des infrastructures de transport. Le résultat a donc une certaine pertinence.

Pour autant, nous avons montré que dans un environnement de rareté des fonds publics, le ratio de la VAN par euro public investi permet la prise en compte de la contrainte financière de la collectivité. Le critère de la rentabilité immédiate n'apparaît plus comme suffisant. Si intuitivement il paraît toujours aberrant de réaliser un projet avant que sa rentabilité immédiate ait atteint le taux d'actualisation, il est aussi nécessaire de considérer le coût de rareté des fonds publics. Et ce coût s'exprime tout naturellement dans la hiérarchisation des projets par le ratio VAN/S.

Pour résoudre le problème du passage des projets au programme, le critère de la VAN par euro public investi est un outil puissant. C'est du moins ce que laisse penser la démonstration analytique précédente dans un modèle basique, et ce que permettent de constater les simulations numériques exposées. Dans le chapitre 8, Joël Maurice propose un modèle plus complexe qui permet d'enrichir cette discussion.

Références

- Abraham C. et Laure A., « Étude des programmes d'investissement routiers », *Annales des Ponts et Chaussées*, novembre-décembre, 1959.
- Babusiaux D., *Décision d'investissement et calcul économique dans l'entreprise*, Economica / Technip, coll. Économie et statistiques avancées, 1990.

¹⁴. Au passage, le résultat précédent montre qu'un projet ne doit pas être réalisé dès que la VAN devient positive, mais quand elle devient maximum. « Dans ces conditions, une opération isolée ne doit pas être réalisée dès qu'elle est juste rentable [i.e. VAN > 0], mais quelques années après seulement » (Abraham & Laure, 1959, p. 744), ce qui est un résultat trop souvent ignoré.

- Baumol W. et Quandt R., « Investment and Discount Rates Under capital Rationing, A Programming Approach », *The Economic Journal*, 75(298), 1965, p. 317-329.
- Bertoneche M. et Langohr H., « Le choix des investissements en situation de rationnement de capital. Comparaison des solutions fournies par différents modèles théoriques », *Revue Économique*, 28(5), 1977.
- Bloy E., Bonnafous A., Cusset J.-M. et Gerardin B., *Evaluer la politique des transports*, Economica / Presse Universitaires de Lyon, coll. « Économie publique de l'aménagement et des transports », 1976.
- Bonnafous A. et Jensen P., « Ranking Transport Projects by their Socio-economic Value or Financial Interest Rate of Return ? », *Transport Policy*, 12, 2005, p. 131-136.
- Bonnafous A. et Jensen P., (2006), « Le cofinancement usager-contribuable et le partenariat public-privé changent les termes de l'évaluation des programmes d'investissement public », *Économie et Prévision*, à paraître.
- Bonnafous A., « Infrastructures publiques et financement privé : le paradoxe de la rentabilité financière », *Revue d'Économie Financière*, 51, 1999.
- Bonnafous A., « Les infrastructures de transport et la logique financière du partenariat public-privé : quelques paradoxes », *Revue Française d'Économie*, 17(1), 2002.
- Bonnafous A., « Les infrastructures de transport et la logique financière du partenariat public-privé : quelques paradoxes », *Revue Française d'Économie*, 17(1), 2002.
- Bonnell P., *Prévoir la demande de transport*, Presses de l'École nationale des Ponts et Chaussées, Paris, 2004.
- CERTU, *Recommandations pour l'évaluation socio-économique des projets de TCSP*, référence n° 25, DTT, 2002, 143 p.
- Commissariat Général du Plan, *Transports : pour un meilleur choix des investissements*, M. Boiteux (prés.), La Documentation Française, Paris, 1994.
- Commissariat General du Plan, *Transports : choix des investissements et coût des nuisances*, M. Boiteux (prés.), La Documentation Française, Paris, 2001.
- Commissariat General du Plan, *Révision du taux d'actualisation des investissements publics*, D. Lebègue (prés.), La Documentation Française, Paris, 2005.
- Everett H., « Generalized Lagrange multiplier method for solving problems of optimum allocation of resources », *Operations Research*, 11(3), 1963, p. 399-417.

- Hayashi Y. et Morisugi H., « International comparison of background concept and methodology of transportation project appraisal », *Transport Policy*, 7(1), January, 2000, p. 73-88.
- Langlois G. & Mollet M., *Gestion financière*, Fouchet, DECF, 2^e édition, 1999.
- Lesourne J., *Le calcul économique : théorie et application*, Dunod, 1972, 459 p.
- Lorie J.H. et Savage L.J., « Three problems in rationing capital », *Journal of Business*, 28(4), 1955, p. 229-239.
- Maurice J., « Traitement des fonds publics », *Note de réflexion*, groupe de travail du CGPC sur le taux d'actualisation, 2004.
- Quinet E. et Sauvant A., (2006), « Optimisation et décentralisation des programmes d'investissements de transport », *Économie et Prévision*, à paraître.
- Quinet E., *Analyse économique des transports*, PUF, Paris, 1990.
- Quinet E., *Principes d'économie des transports*, Economica, Paris, 1998.
- Quinet E., « Evaluation methodologies of transportation projects in France », *Transport Policy*, 7(1), January, 2000, p. 27-34
- Rothengatter W., « Evaluation of infrastructure investments in Germany », *Transport Policy*, 7(1), January, 2000, p. 17-26
- Roy W., « Évaluation des programmes d'infrastructure : Ordre optimal de réalisation sous contrainte financière », Document de travail halshs-00003971, 2005.
- Vernimmen P., *Finance d'entreprise*, 6^e édition par P. Quiry et Y. Le Fur, Dalloz, 2005.
- Vickerman R., « Evaluation methodologies for transport projects in the United Kingdom », *Transport Policy*, 7(1), January, p.7-16, 2000.
- Van Laarhoven P. et Aarts E. H. L., *Simulated Annealing : Theory and Applications*, Kluwer, Dordrecht, 1992.
- Weingartner M., *Mathematical Programming and the Analysis of Capital Budgeting Problems*, Englewood Cliffs, Prentice-Hall (reissued, London, Kershaw, 1974), 1963.

CHAPITRE 8

UNE MÉTHODE D'OPTIMISATION DES PROGRAMMES D'INVESTISSEMENTS DE TRANSPORT

Emile Quinet¹ et Alain Sauvant²

Le calcul économique appliqué aux investissements d'infrastructure se trouve, dans les conditions économiques actuelles, confronté à plusieurs problèmes qui en rendent l'application délicate et qui conduisent souvent à le réduire à une simple vérification que la rentabilité de chaque projet est satisfaisante, sans recherche d'optimisation d'ensemble. Citons la pénurie des fonds budgétaires et l'incertitude sur leur évolution future, le recours au financement privé qui occulte les considérations de rentabilité socio-économique au profit de la rentabilité financière, les problèmes posés par la décentralisation des décisions et la multiplication des décideurs. La présente réflexion propose une méthode pour intégrer ces problèmes dans l'optimisation d'ensemble d'un programme.

Le cas type de projets auxquels on s'intéresse ici est celui des investissements de d'infrastructure de transport, même si les résultats obtenus peuvent s'appliquer à d'autres situations. On a affaire à des projets qui présentent plusieurs caractéristiques :

- ils sont localisés géographiquement ;
- ils ne sont pas repérables sur une échelle quantifiée susceptible de variations continues, comme par exemple l'accumulation d'un nombre à déterminer de projets identiques et de petite taille par rapport au programme d'ensemble, ou la taille modulable d'un parc de machines. On ne peut donc pas leur appliquer les méthodes d'optimisation continue classiques (maximisation de fonction, programmation dynamique) ; ne sont pas repérables sur une échelle quantifiée susceptible de

1. Professeur émérite à l'École Nationale des Ponts et Chaussées, membre de PSE-ENPC.

2. Professeur à l'École Nationale des Ponts et Chaussées.

variations continues, comme par exemple l'accumulation d'un nombre à déterminer de projets identiques et de petite taille par rapport au programme d'ensemble, ou la taille modulable d'un parc de machines. On ne peut donc pas leur appliquer les méthodes d'optimisation continue classiques (maximisation de fonction, programmation dynamique).

- ils sont soit indépendants les uns des autres soit, pour certains d'entre eux, liés dans la mesure par exemple où la réalisation de l'un a des conséquences sur la rentabilité d'un ou plusieurs autres ;

- chaque projet peut faire l'objet de variantes incompatibles ; ces variantes peuvent porter sur leur consistance technique ou les niveaux de tarification de leur usage, ou encore leur modalité de financement, public ou privé ;

- la rentabilité de chaque projet change selon son année de mise en service, car selon cette année, les trafics et donc les avantages procurés par l'investissement changent, en général augmentent.

La définition d'un programme consiste à répondre à plusieurs questions qui sont dans l'ordre logique : le choix entre les variantes incompatibles d'un même projet ; éventuellement les choix concernant des grappes de projets dont les avantages sont liés entre eux par des effets de réseaux ; la décision de faire ou de ne pas faire chaque projet ; enfin la date de réalisation si la réponse à la question précédente est positive. Une question importante est alors celle de la décentralisation : quelles sont parmi les décisions précédentes, celles qui peuvent être décentralisées au niveau des services d'étude locaux et quelles sont celles qui doivent être effectuées au niveau central ?

Ce sont ces problèmes d'optimisation et de décentralisation de la programmation qu'aborde le présent texte. Ils n'ont fait l'objet que de peu d'attention, malgré leur importance concrète. Pour citer surtout des références françaises, Abraham et Laure³ avaient montré que, dans le cas d'investissements indépendants et sous des hypothèses assez générales, il convenait de mettre chacun d'eux en service à l'année où le rapport entre les avantages apportés la première année divisés par le coût d'investissement est égal au taux d'actualisation ; ils aussi établi que, dans une situation d'insuffisance des crédits disponibles, on pouvait choisir les projets et leur ordre à partir d'une série de variables duales, une par année, par lesquelles on multiplie la dépense objet de la contrainte, telles que la programmation s'effectue comme précédemment : en somme tout se passe comme si un coefficient de

3. Abraham C. et Laure A., « Étude des programmes d'investissements routiers », *Annales des Ponts et chaussées*, novembre, 1959.

rareté des crédits différent d'une année à l'autre assurait l'ajustement aux crédits disponibles. Mais ce résultat est obtenu sous des hypothèses assez restrictives et difficiles à vérifier (existence dans l'ensemble des opérations de chaque année d'une opération dont les avantages connaissent la plus faible croissance).

Dans plusieurs publications, Bonnafous et Jansen⁴ l'ont abordé par une méthode de simulation numérique en mettant l'accent sur les problèmes posés par le financement privé des investissements et en proposant le critère de choix « bénéfice socio-économique actualisé sur euro public investi ».

Dans un papier non publié, Pozmantir (1999)⁵ a établi une méthode pour ajuster à la marge un programme pré-établi, afin de tenir compte de modifications dans les coûts et dans les crédits disponibles.

Le présent texte propose une méthode de résolution de ce genre de problème permettant de traiter de nombreuses situations englobant celles décrites plus haut. Il est organisé ainsi : La seconde partie présente la diversité des questions que la programmation des investissements doit résoudre, en mettant à chaque fois l'accent sur les possibilités de décentralisation, et propose une méthode générale de résolution de ces problèmes, fondée sur la programmation linéaire. La troisième partie explore, à partir à la fois de situations réelles et simulées, les résultats obtenus par cette méthode et teste la possibilité de décentralisation à travers notamment la robustesse des critères simples usuellement présentés tels que taux de rentabilité interne ou le critère de la VAN par euro investi. La quatrième conclut en proposant des règles pour la décision et en suggérant des pistes de recherche complémentaire.

1. L'ÉLABORATION DES PROGRAMMES D'INVESTISSEMENTS DE TRANSPORTS

Soient un certain nombre de projets ; chacun d'eux peut comporter des variantes exclusives (tracé A ou tracé B, par exemple) ; ces projets

4. Et notamment « Ranking Transport Projects by their Socioeconomic Value or Financial Interest rate of return ? », Alain Bonnafous, Laboratoire d'Économie des Transports et Pablo Jensen, Laboratoire d'Économie des Transports and Laboratoire de Physique de la Matière Condensée et des Nano structures, First Conference on Railroad Industry Structure, Competition and Investment, Toulouse, November 7-8, 2003 ou Bonnafous A., Jensen P., « Ranking transport projects by their socioeconomic value or financial internal rate of return ? », *Transport Policy*, vol. 12, n° 2, 2005.

5. « Méthode de correction des tâches et des contraintes sur un programme d'investissement de branche, compte tenu de l'état financier de la branche », W. Pozmantir, Working Paper, Institute of Studies of Complex Systems, 1999.

peuvent ou non être liés (par exemple les avantages du projet 1 diffèrent ou non selon que le projet 2 est ou non réalisé) ; chaque projet peut être financé avec plus ou moins de participation privée ; chaque projet peut être mis en service à n'importe quelle date du futur ; enfin chaque année il y a différentes contraintes financières (contraintes sur les fonds publics, contraintes sur les volumes d'emprunts, ...).

Dans toute sa généralité, la question du programme optimal se pose de la manière suivante : Quelle est la combinaison optimale, c'est-à-dire celle qui maximise la valeur actualisée totale, en respectant les contraintes qui lient les projets entre eux et les limitations de budget ? La démarche proposée ici est fondée sur la programmation linéaire. On va la présenter en examinant des situations de plus en plus complexes.

1.1. La diversité des problèmes posés par la programmation des investissements

En allant du plus simple au plus compliqué, on rencontre les situations suivantes.

1.1.1. Choix entre variantes alternatives d'un même projet sans contrainte budgétaire avec financement uniquement public

Soient différentes variantes d'un même projet indicées par n (entre 0 et N) entre lesquelles se produit le choix ; ces variantes i diffèrent par leur consistance technique et leur mode de financement ; leur coût (supposé indépendant de la date de réalisation, mais cette hypothèse pourrait être levée) est I_n leur Bénéfice socio-économique actualisé B dépend de l'année de réalisation et s'exprime sous la forme : $B_n(t_0^n)$

Sans perte de généralité, on supposera que la durée de réalisation est concentrée sur une seule année t_0^n ; il serait facile de lever cette hypothèse au prix seulement d'une complication d'écriture. On supposera également que les seules dépenses d'infrastructure sont les dépenses de premier investissement, et que celles-ci sont réalisées sur une année seulement. Il serait possible de se dégager de ces hypothèses simplificatrices, mais au prix d'une plus grande complexité d'écriture et sans changement fondamental des résultats.

Avec :

$$B_n(t_0^n) = -\frac{I_n}{(1+r)^{t_0^n}} + \sum_{t=t_0^n+1}^{t=\infty} \frac{a_n(t)}{(1+r)^t}$$

Dans cette expression, r est le taux d'intérêt ou d'actualisation (supposé constant pour simplifier, mais son évolution selon les années serait aisée à prendre en compte), $a_n(t)$ représente les avantages de la variante j pour l'année t ; on suppose que ces avantages dépendent de l'année mais pas de l'année de mise en service⁶ et qu'ils s'étendent jusqu'à l'infini (d'autres hypothèses compliquent l'écriture, sans changer fondamentalement les résultats). En présence d'un coût d'opportunité des fonds publics tel que défini par l'instruction-cadre du 27 mai 2005 et notamment par son annexe III (§3), les fonds publics intervenant dans cette formule doivent être multipliés par ce coefficient, fixé à 1,3 par la même instruction. Le choix à faire porte sur le fait de faire ou non ce projet, sur la variante éventuelle à choisir et sur l'année de sa mise en service.

En l'absence de contrainte budgétaire, il est aisé de voir que, pourvu que l'investissement I soit indépendant (en euros constants) de la date de réalisation et que les $A_i(t)$ soient non décroissants en t , l'année optimale de réalisation de la variante i , si elle est choisie, est telle que (voir Abaham et Laure, 1959 ou Quinet, 1998⁷) :

$$a_n(t_0^n)/r = I_n$$

Si cette égalité ne peut pas être satisfaite, le projet doit être fait immédiatement si le premier terme est toujours supérieur au second, et ne doit jamais être réalisé dans le cas inverse. C'est en effet ainsi que le bénéfice actualisé tiré du projet et défini par la relation précédente est maximale.

Le critère du Bénéfice actualisé socio-économique permet de choisir entre variantes d'un même projet : on détermine pour chacune sa date optimale et on choisit celle qui, placée à sa date optimale, rapporte le bénéfice actualisé le plus élevé. C'est en effet ainsi que l'on maximise le bénéfice tiré du projet.

Il faut bien voir les limites des enseignements à tirer du bénéfice actualisé ainsi compris : il ne dit pas que la variante choisie est optimale, mais simplement qu'elle est préférable à toutes celles sur lesquelles a porté la comparaison, mais il ne dit pas que la variante choisie est la bonne. D'où l'importance à la fois de bien choisir la situation de référence⁸ et d'ouvrir au maximum l'éventail des projets

6. Cette hypothèse simplificatrice exclut les mécanismes dits de « build-up » par lesquels le trafic d'une nouvelle infrastructure met quelques années à atteindre sa valeur finale.

7. Quinet E., *Principes d'économie des transports*, Economica, 1998.

8. Voir à ce sujet les recommandations du rapport Boiteux (Boiteux, 2000).

dont on compare le bénéfice actualisé. Un biais fréquent des analystes qui veulent faire passer un projet est de choisir une situation de référence médiocre, pour accroître d'autant la rentabilité du projet proposé

1.1.2. *Établissement d'un programme d'investissements indépendants sans contrainte budgétaire avec financement uniquement public*

Si toutes les variantes optimisées sont indépendantes (c'est-à-dire que la réalisation de l'une d'elles n'a pas d'influence sur la rentabilité des autres), la règle de programmation est alors la suivante : on réalise chacune d'entre elles à sa date optimale de mise en service. En effet le programme à traiter revient à maximiser la somme de tous les bénéfices actualisés tirés de chacun des projets s'écrit :

$$\sum_n B_n(t_n^0) = \sum_n \left[-\frac{I_n}{(1+r)^{t_n^0}} + \sum_{t=t_n^0+1}^{t=\infty} \frac{a_n(t)}{(1+r)^t} \right]$$

Dans cette expression les inconnues sont les dates de mise en service t_n^0 . Les différentes expressions entre crochets du côté droit de la relation précédente comportent une et une seule de ces variables, la fonction de droite est séparable en chacune d'elles, et l'optimisation se fait pour chaque crochet séparément.

Alors la procédure de choix peut être totalement décentralisée (c'est la conséquence de la fixation au bon niveau du taux d'actualisation dans une économie « optimale »). Chaque échelon décentralisé peut isolément déterminer la bonne variante technique et sa date de réalisation. L'échelon central auquel il demandera les crédits correspondants n'aura pas de raison de les lui refuser, sauf bien sûr problèmes d'asymétrie d'information et de divergences d'objectifs entre échelon central et échelons décentralisés, situation exclue de l'analyse.

1.1.3. *Établissement d'un programme d'investissements liés sans contrainte budgétaire avec financement purement public*

L'affaire est un peu plus compliquée lorsque les investissements sont liés. C'est le cas par exemple de deux infrastructures concurrentes (respectivement : complémentaires) sur un même axe : lorsqu'on les réalise ensemble, l'avantage annuel total est inférieur (respectivement : supérieur) à la somme des avantages annuels des deux infrastructures prises isolément. Lorsqu'il en est ainsi, il faut énumérer la

combinaison des séquences possibles⁹, pour chacune d'entre elles déterminer la date de réalisation de chaque projet, en déduire le bénéfice actualisé de la séquence, et retenir finalement la séquence qui procure le bénéfice actualisé le plus élevé. Là aussi la procédure peut être décentralisée à condition que chaque grappe de projets liés soit examinée par une même entité décentralisée.

1.1.4. *Établissement d'un programme d'investissements liés sans contrainte budgétaire avec participation du secteur privé*

Alors chaque projet i se caractérise non seulement par les avantages annuels qu'il apporte, mais aussi par le montant de financement privé qu'il est possible de trouver ; ce montant résulte de négociations entre les financeurs possibles et les promoteurs du projet et se traduit par un échelonnement de transferts entre la puissance publique et le partenaire privé et/ou de dépenses qui sont à la charge de l'un ou de l'autre contractants. Dans le cadre classique de la concession du type BOT (Build Operate and Transfer), le contrat de concession se traduit par le fait que le concessionnaire effectue pour le compte de la puissance publique des dépenses initiales (l'investissement) et des dépenses d'entretien et perçoit des recettes de péage tout au long de la concession ; après la fin de la concession, c'est en principe la puissance publique qui récupère l'ouvrage, qui effectue des dépenses d'entretien et qui perçoit les péages. Dans le partenariat public privé actuellement en cours de mise en œuvre, c'est le partenaire privé qui effectue tout ou partie des dépenses d'investissement et d'entretien, et la puissance publique verse au partenaire privé, outre une éventuelle subvention au départ, des sommes annuelles. On ne considérera ici que le cas de la concession et on en représentera le mécanisme de manière sommaire : soient $EBE_n(t)$ l'EBE (recettes nettes des dépenses d'exploitation) obtenu chaque année par la redevance d'infrastructure (on y ajoutera les versements annuels de la puissance publique en cas de subvention d'exploitation ou de financement PPP). Ces recettes varient dans le temps. En fonction de la durée du partenariat T_f et du coût moyen pondéré du capital (WACC) i que le financeur privé peut obtenir¹⁰,

9. Cette énumération comporte quelques difficultés : on ne connaît pas *a priori* l'ordre des deux projets ; il faut énumérer les ordres possibles (A puis B, B puis A, A et B en même temps), dans chaque ordre, trouver la date optimale de réalisation de chaque investissement, puis calculer le bénéfice actualisé de chaque combinaison ainsi optimisée quant aux dates, et choisir la combinaison qui conduit au bénéfice actualisé le plus élevé.

10. Dans la réalité, ce coût moyen pondéré n'est pas une donnée exogène, mais résulte des caractéristiques du projet et de celles de l'investisseur.

celui-ci est susceptible de financer un montant $I_f(t_n^0)$ du coût du projet i tel que :

$$I_f(t_n^0) = \sum_{t=t_n^0+1}^{t=t_n^0+T_f} \frac{EBE_n(t)}{(1+i)^t}$$

C'est en effet pour ce niveau de financement que la VAN retirée du projet par le concessionnaire est nulle, et c'est ce à quoi une mise aux enchères efficace aboutirait. Ce montant dépend bien sûr de la date de réalisation et en général, comme les trafics augmentent au cours des années, il croît avec cette date. Alors le coût des fonds publics s'appliquera à la différence entre le coût total d'investissement et la part privé $I_f(t_n^0)$. Corrélativement d'ailleurs, les recettes d'infrastructure qui, dans une situation de financement public, doivent être affectées du coût des fonds publics, ne le sont plus tant qu'elles sont perçues par le partenaire privé, c'est-à-dire tant que dure la concession.

Là aussi la décentralisation des décisions peut être totale.

1.1.5. Optimisation en situation de contrainte budgétaire

Dans une situation « optimale », le taux d'actualisation devrait être tel que tous les projets « rentables » peuvent être réalisés dans le cadre des enveloppes budgétaires disponibles. La pratique fait apparaître des écarts entre les besoins à satisfaire et les enveloppes. Ces écarts peuvent résulter de désajustements transitoires, ils peuvent aussi résulter de sauts dans la fixation du taux d'actualisation. Ainsi, le récent rapport Lebègue¹¹, qui abaisse le taux d'actualisation du Plan de 8 % à 4 %, devrait entraîner sur plusieurs années une augmentation sensible des projets rentables et donc des enveloppes financières qui leur sont consacrées, augmentation dont on peut douter qu'elle se réalise, compte tenu des tensions que connaissent les budgets publics actuellement.

Alors le problème est identique au précédent, en tenant compte de ce que chaque année le budget de dépenses – en général le budget de dépenses sur fonds publics – ne doit pas dépasser un montant fixé par année. Avec la simplification faite ici de négliger les dépenses d'entretien, les dépenses en cause ne concernent que les dépenses d'investissement, mais dans une optique plus générale il faudrait y inclure les dépenses d'entretien.

11. Commissariat Général du Plan, *Révision du taux d'actualisation des investissements publics*, Rapport d'un groupe de travail présidé par D. Lebègue, rapporteur L. Baumstark, La Documentation Française, Paris, 2005.

Il faut noter que dans ces conditions chaque décision sur un projet particulier influe sur la programmation des autres par l'intermédiaire de la contrainte budgétaire. Une question importante est alors de savoir si des critères simples permettent aux échelons décentralisés de prendre des décisions de programmation, à commencer par choisir entre variantes techniques, ou encore de choisir entre projets liés, ou enfin de décider la date de réalisation du projet retenu.

1.2. Une méthode générale de résolution

La méthode proposée est la mise en œuvre d'une idée présentée dans Quinet (2005)¹². Examinons d'abord le cas où sont en lice uniquement des projets définis par leur contexture technique et par leur mode de financement. Le problème peut alors être formulé de la façon suivante, qui revient à un programme linéaire :

$$\text{Max}_{x_{n,t}} \left[\sum_i \sum_t B_n(t_n^0) * x_{n,t} \right]$$

Avec les contraintes suivantes :

$$0 \leq x_{n,t} \leq 1$$

$\sum_t x_{n,t} \leq 1$: cette contrainte signifie que le projet n'est réalisé qu'une fois au plus.

$\sum_n x_{n,t} * I_n \leq C_t$: c'est la contrainte budgétaire de l'année t, et il y en a autant que d'années en considération. C_t est le budget maximum de l'année t.

Cette contrainte peut s'appliquer à l'ensemble des sources de financement des projets, ou, plus souvent, à la part publique du financement, soit, avec les notations précédemment introduites, à la différence entre le coût total et la part financée par le partenaire privé. La contrainte s'écrit alors, dans le cas envisagé au long de ce texte, qui est celui d'une concession¹³ :

$$\sum_n x_{n,t} * [I_n - If_n(t_n^0)] \leq C_t$$

Sous cette forme, il est possible que le programme trouve pour un même projet n des valeurs de x_{nt} différentes de zéro pour plusieurs

12. Quinet E., « À propos du rapport Lebègue », *Transports*, août-septembre, 2005.

13. Les contraintes seraient différentes, mais faciles à écrire, dans le cas d'un partenariat public-privé.

valeurs de t , ce qui n'a pas grand sens, sauf si cela se produit pour un seul projet dont la réalisation est à cheval entre deux années successives (on verra dans la suite que c'est cette situation qui se rencontre dans les simulations).

On peut de façon alternative supposer que les $x_{i,t}$ sont des nombres entiers. On a alors un programme linéaire en nombres entiers. Ce cas n'est pas traité dans les simulations du présent chapitre.

Dans ce programme les nombres $B_n(t_n^0)$ qui interviennent sont calculés à partir de la connaissance des avantages et des coûts annuels ; les nombres $If_n(t_n^0)$ sont également calculés à partir des caractéristiques du projet en termes de tarification et de clientèle, ainsi que des conditions financières du marché. Il est facile de voir que la solution ne peut pas s'exprimer par une règle simple générale, sauf hypothèse complémentaire forte.

1.2.1. Optimisation de variantes incompatibles

On peut aussi envisager l'optimisation de variantes. Soient par exemple deux variantes i et $i+1$ d'un même projet (variantes techniques ou dans le mode de financement). Alors il suffit d'ajouter la contrainte :

$$\sum x_{n,t} + \sum x_{n+1,t} \leq 1$$

Cette contrainte exprime que l'on fait l'une ou l'autre des deux variantes, mais pas les deux.

1.2.2. Effets de réseau et investissements liés

On peut aussi traiter le problème d'investissements liés. Si par exemple deux investissements n et p sont liés, cela signifie que la rentabilité du projet n diffère selon que le projet p est ou non réalisé. Pour traduire ce phénomène, on scindera chacun de ces projets en deux selon sa modalité de réalisation. Des contraintes d'exclusion faciles à écrire permettent de tenir compte des incompatibilités de réalisations.

Si par exemple deux projets 1 et 2 présentent des effets de réseaux, on introduira, outre ces deux projets comptés avec les avantages qui résulteraient de la réalisation isolée de chacun, deux projets auxiliaires : un projet 3 correspondant au projet 2 s'il est fait après 1, et un projet 4 correspondant au projet 1 s'il

est fait après 2. Aux variables x_{1t} , x_{2t} , x_{3t} et x_{4t} , on ajoute les variables X_{1t} (et la suite) définies par :

$$X_{1t} = \sum_{u=1}^{u=t} x_{1u}$$

Les contraintes à ajouter pour tenir compte des liaisons entre projets sont :

$X_{1t} + X_{2t} \leq 1$: ce qui signifie que si 1 est réalisé, 2 ne l'est pas selon la modalité où il rapporte des avantages isolés

$X_{3t} - X_{1t} \leq 0$: ce qui signifie que si 1 est réalisé, 2 peut l'être sous la forme (et avec les avantages) où il est fait après 1.

$X_{4t} - X_{2t} \leq 0$: même chose pour 2.

1.2.3. Cas d'incertitude sur les contraintes financières futures

On peut enfin aborder le cas où il y a une incertitude sur les contraintes budgétaires futures. Supposons par exemple qu'on connaisse la contrainte budgétaire C_1 de la période 1, mais que pour les périodes ultérieures on ait une incertitude entre une première chronique C_{th} (avec $t > 1$) avec une probabilité q et une autre chronique C_{tb} (avec $t > 1$) avec une probabilité $1 - q$, l'incertitude étant levée à la fin de la première période lorsque le choix de cette première période est effectué. On introduira pour chaque projet i :

- une variable x_i , égale à l'unité si le projet i est réalisé au cours de la première période, et à 0 autrement,
- une série de variables h_{it} (avec $t > 1$) relatives à la chronique C_{th} , et qui prendront la valeur 1 pour l'année où le projet est réalisé, 0 autrement,
- une deuxième série de variables b_{it} (avec $t > 1$) relatives à la chronique C_{tb} , et qui prendront la valeur 1 pour l'année où le projet est réalisé, 0 autrement.

Le programme s'écrit alors :

$$\text{Max}_{x_n, h_{nt}, b_{nt}} \left[\sum_n B_{n1} * x_n + q \sum_{t \geq 2} B_n(T_{n,t}^0) * h_{nt} + (1 - q) \sum_{t \geq 2} B_{n,t}(T_{n,t}) * b_{nt} \right]$$

Sous les contraintes suivantes :

$$0 \leq x_n, h_{nt}, b_{nt} \leq 1$$

$$x_n + \sum_{t \geq 2} [h_{nt}] \leq 1 : \text{cette contrainte signifie que le projet ne peut être}$$

réalisé qu'une fois dans la chronique H

$$x_n + \sum_{t \geq 2} [b_{n,t}] \leq 1 : \text{m\^eme chose pour la chronique B}$$

$$\sum_n x_n * I_n \leq C_1 : \text{c'est la contrainte budg\^etaire de l'ann\^ee 1}$$

$\sum_n h_{in} * I_n \leq C_{ht} : \text{c'est la contrainte budg\^etaire de l'ann\^ee t (avec } t > 1) \text{ dans l'hypoth\^ese de la chronique H ; la m\^eme s\^erie de contrainte op\^ere pour la chronique B en intervertissant h et b.}$

Il serait bien sur ais\^e de d\^ecrire des situations d'incertitude plus complexes.

2. R\^ESULTATS DE SIMULATIONS

La m\^ethode qui vient d'\^etre propos\^ee a \^ete utilis\^ee pour explorer par simulation un certain nombre de probl\^emes li\^es à la programmation des investissements. On pr\^esentera d'abord le mod\^ele de simulation utilis\^e avant de d\^evelopper quelques r\^esultats issus de son usage.

2.1. *Le mod\^ele de simulation utilis\^e et les param\^etres retenus*

Le mod\^ele de simulation utilis\^e a \^ete exp\^eriment\^e sous tableur (Excel) ; il permet de simuler la programmation de 20 projets dont les dates de r\^ealisation peuvent s'\^etaler de 2005 à 2034. Les contraintes pratiques du solveur Excel qui n'accepte pas plus de 200 variables am\^enent à raisonner sur des contraintes budg\^etaires tri-annuelles, et 10 p\^eriodes, ce qui permet de couvrir une dur\^ee de programmation de 30 ans, qui correspond aux horizons couramment consid\^er\^es dans les \^etudes de sch\^ema directeur et de programmation à moyen terme.

Les projets sont d\^efinis par leurs avantages socio-\^economiques, variables dans le temps selon une loi polynomiale, et par leurs co\^uts qui sont \^egalement susceptibles de varier avec l'ann\^ee selon une loi polynomiale.

Le mod\^ele pr\^evoit deux types de financements possibles : public ou priv\^e. Pour cela sont utilis\^es deux taux d'actualisation :

– le taux d'actualisation public. Celui-ci est classiquement fix\^e en France par le Commissariat G\^en\^eral du Plan. Il est longtemps rest\^e à 8 %, le rapport Leb\^egue a recommand\^e une valeur plus faible qui d\^ecro\^it avec le temps (4 % de 2005 à 2034, 3,5 % de 2035 à 2054, 3 % apr\^es 2055) en euros constants. C'est cette valeur qui est pr\^esent\^ee par d\^efaut dans le mod\^ele ;

– le coût moyen du capital du financement privé est classiquement la moyenne pondérée du taux des emprunts et du rendement exigé par les investisseurs privés pour posséder des actions :

$$r_c = r_e \frac{E}{E + D} + r_d (1 - t_c) \frac{D}{E + D}$$

Dans cette expression : E représente la part des actions et D représente la part de la dette dans le financement, r_e le taux de rendement des actions, et r_d le taux de rendement des emprunts, t_c étant l'impôt sur les sociétés. Le coût moyen du capital dépend des différents paramètres de la formule précédente. Pour un financement privé pur tel que celui qui est réalisé pour le tunnel sous la Manche, la part du capital (actions ou obligations convertibles) est de l'ordre de 10 à 30 % du total du financement, le coût nominal des emprunts lancés par un investisseur privé est actuellement de l'ordre de 5,5 à 6,5 %, et le rendement attendu du capital privé, qui dépend du risque du projet, se situe selon les cas entre 8 et 12 % (et peut monter au-delà si le projet est particulièrement risqué. Le coût moyen du capital privé se situe donc autour de 7 % en euro courants, soit à peu près 5 % en euro constants. Cette valeur peut différer selon les conditions d'emprunt, la solvabilité de l'emprunteur et la part du capital. Il apparaît cependant que le coût moyen du capital privé est en général supérieur au taux d'actualisation, mais d'assez peu.

Le modèle intègre également un coût d'opportunité des fonds publics, destiné à traduire les imperfections de l'impôt et les distorsions qu'ils impliquent : un euro d'impôt entraîne une perte d'utilité collective supérieure à un euro de fonds privé, et le coût d'opportunité des fonds publics est égal au rapport entre les deux. Le rapport Lebègue déjà cité a recommandé un coût des fonds publics de 1,3 (un euro public égale 1,3 euro privé). C'est cette valeur de coût d'opportunité des fonds publics qui a été introduite par défaut dans le programme, et qui joue de la manière suivante sur les coûts et les avantages d'un investissement réalisé par un système de concession classique¹⁴, comme le montre le tableau suivant :

14. Le jeu de ce coefficient serait bien sur différent si le financement privé intervenait sous une autre forme, par exemple sous la forme de « partenariat public privé », au sens où ce terme est entendu en France.

Part à affecter du coût des fonds publics	Public	Privé
Investissement	100 %	Subvention publique seulement
Avantages financiers pendant la durée de concession	100 %	0 %
Avantages financiers au-delà de la durée de concession	100 %	100 %

Ceci s'explique par les éléments suivants :

- en cas de financement privé, au-delà de la durée de concession, les avantages financiers retournent au public, ce qui évite une levée d'impôts du même montant, et évite donc une distorsion économique ; cette distorsion économique est actualisée au taux d'actualisation public ;
- en cas de financement public, cet effet joue dès la première année de vie du projet.

La durée de concession est par défaut supposée égale à 30 ans, et la subvention publique en cas de financement privé est également calculée de sorte que la VAN financière du concessionnaire privé (actualisée au taux WACC) soit nulle sur 30 ans¹⁵.

Le modèle détermine le programme optimal, c'est-à-dire la liste des projets retenus, la date de mise en service de chacun, la VAN de chaque projet et la VAN de l'ensemble du programme, ainsi que les valeurs des indicateurs de rentabilité usuels.

– Le taux de rentabilité interne (TRI) calculé à l'année où l'on décide le programme : c'est le taux d'actualisation qui annulerait la VAN si le projet était mis en service cette année-là. Ce calcul peut être effectué sous l'angle socio-économique (TRI SE : on prend alors en compte les avantages socio-économiques) ou sous l'angle financier (TRI FI : on prend alors en compte les conséquences financières pour le gestionnaire d'infrastructure).

– Le taux de rentabilité immédiate (Tri) calculé à l'année où l'on décide le programme : c'est le rapport entre l'avantage que procurerait l'investissement cette année et le coût de l'investissement. Là aussi, ce calcul peut être fait sous l'angle socio-économique (Tri SE : on prend alors en compte les avantages socio-économiques) ou sous l'angle

15. Ces deux durées peuvent être différentes.

financier (Tri FI : on prend alors en compte les conséquences financières pour le gestionnaire d'infrastructure).

– Le rapport « Bénéfice / Coût de l'investissement », qui est recommandé par le rapport Lebègue ; il y a autant de tels ratios qu'il y a d'années possibles de mise en service ; mais le rapport Lebègue ne précise pas l'année en question ; néanmoins, pour que le ratio apporte une information, il est nécessaire de dire à quelle année est fait le calcul ; suivant l'intuition naturelle, on le calculera aussi à l'année où l'on décide le programme.

– Le rapport « Bénéfice / Fonds publics » (B/FP) ; ce ratio est également recommandé par le rapport Lebègue dans le cas de financement privé. Pour les mêmes raisons que pour le ratio précédent, il faut définir l'année de calcul et on prendra l'année d'établissement du programme.

2.2. Description des tests effectués

On a distingué :

- des tests exploratoires relatifs aux effets isolés de certains paramètres tels que la croissance du coût d'investissement, la croissance des avantages socio-économique ou la croissance des avantages financiers ;
- des tests de choix décentralisés entre variantes techniques incompatibles et entre projets liés ;
- un test de programmation centrale comparant des projets indépendants dont les paramètres ont été engendrés de façon aléatoire ;
- un test de programmation centrale sur projets réels indépendants.

2.3. Résultats de simulation

2.3.1. Tests de sensibilité aux paramètres clés

Croissance du coût d'investissement en fonction du temps

Si l'on compare des projets à paramètres identiques sauf pour ce qui concerne la croissance dans le temps du coût d'investissement en fonction de la date de mise en service, les premiers projets à programmer sont ceux pour lesquels la croissance du coût d'investissement sera la plus forte.

Ceci se comprend car, toutes choses égales d'ailleurs, il est plus avantageux de commencer par des projets dont le coût va augmenter dans le temps, comme par exemple ceux qui sont soumis à de fortes évolutions des prix fonciers. En comparaison, des projets où une grande partie des coûts sont liés à des nouvelles technologies dont le

prix baisse tendanciellement seront, toutes choses égales par ailleurs, à réaliser plus tard.

Projets à forte élasticité au PIB

On génère des projets de même taux de rentabilité interne socio-économique (et même ratio avantages financiers/avantages socio-économiques) dont certains présentent une forte croissance des avantages en fonction du temps (en général une forte élasticité au PIB), contrebalancée par des avantages initiaux faibles (et réciproquement).

L'optimisation conduit à commencer par les projets à fort avantage socio-économique initial et faible croissance des avantages. Cela implique de faire en premier les projets dont le ratio VAN socio-économique / fonds publics investis recommandé par le rapport Lebègue est le plus bas. Ceci constitue un cas où une utilisation aveugle de ce ratio peut être trompeuse.

Projets de même TRI socio-économique mais avantages financiers différents

L'optimisation amène à commencer avec les projets à forts avantages financiers ; ceux-ci nécessitent en effet moins de fonds publics, générant moins de distorsions économiques. Cet avancement de la date de réalisation dépend du coût d'opportunité des fonds publics et du niveau de la contrainte budgétaire.

En l'absence de contrainte budgétaire, l'avancement est d'autant plus élevé que le coût d'opportunité des fonds publics est élevé. Notons toutefois que cet effet se produit que le financement des projets soit public ou non, mais son importance dépend du mode de financement, public ou privé.

En présence de contrainte budgétaire, l'avancement se produit même en l'absence de coût d'opportunité des fonds publics, et il est d'autant plus fort que la contrainte l'est. Mais de ce point de vue le ratio Bénéfice/Fonds publics fournit une indication trop brutale, puisqu'il conduit à donner une priorité absolue à des projets qui ne nécessitent aucun fonds public, un critère qui n'est rigoureusement valable que pour une contrainte budgétaire extrême.

2.3.2. Tests de programmation décentralisée

Ces tests ont porté sur le choix entre variantes incompatibles d'un même projet. Leur objet essentiel était de voir si ces choix peuvent être décentralisés au moyen de critères ou de méthodes simples, ceci afin d'éviter que les choix de ce type, qui peuvent être nombreux au cours

de l'étude d'un projet, aient tous à remonter au niveau central. Le test essentiel a porté sur la sensibilité du choix au niveau de la contrainte budgétaire, le seul paramètre qui puisse empêcher une décentralisation totale des décisions en l'absence d'informations sur les coûts de rareté des fonds publics ; lorsque l'on connaît cette chronique, une décentralisation est possible.

Les tests en cause ont été effectués dans le cadre d'un programme comprenant une vingtaine de projets indépendants, chacun d'un coût de 1 000 pour simplifier la présentation du tableau, et dont la programmation s'étale sur un espace de temps de 20 ans, soit 6 périodes de trois ans, choisies pour correspondre en l'absence de contrainte financière à des mises en service de trois projets par période. Dans ce programme on introduit un projet supplémentaire qui peut faire l'objet d'un couple de variantes techniques.

Plusieurs situations de ce type ont été testées. On présente ici les résultats de trois de ces situations.

Tableau 1

Résultats de tests sur le choix de variantes incompatibles

	Variante 1		Variante 2		Contrainte financière annuelle du programme	Projet à retenir
	Avantage socio-éco en 2005	Coût en 2005	Avantage socio-éco en 2005	Coût en 2005		
Test 1	40	1 000	36	900	Pas de contrainte	1
Test 2	40	1 000	37	900	Pas de contrainte	1
Test 3	40	1 000	37	900	75 % des crédits optimaux	1
Test 4	40	1 000	37	900	50 % des crédits optimaux	2
Test 5	40	1 000	26	500	Pas de contrainte	1
Test 6	40	1 000	26	500	Pas de contrainte	1
Test 7	40	1 000	26	500	75 % des crédits optimaux	1
Test 8	40	1 000	26	500	50 % des crédits optimaux	2

Dans la première situation, correspondant au test 1, les deux variantes sont homothétiques ; elles ont même TRI et même rapport B/fonds publics, et ce d'ailleurs, quelle que soit l'année de calcul. Elles ne sont cependant pas équivalentes car le critère du bénéfice actualisé

montre qu'en l'absence de contrainte budgétaire c'est la variante 1 (la plus coûteuse) qui doit être choisie. Ce résultat se retrouve dans la plupart des autres situations simulées.

La situation 2 correspond aux tests 3, 4 et 5. Dans ces tests on prend deux variantes d'un projet, et on analyse comment le choix entre elles se transforme lorsque la contrainte budgétaire se resserre. Il apparaît que le critère simple du bénéfice actualisé sans contrainte financière est assez robuste pour définir la bonne variante. La solution ne change que lorsque le niveau des crédits tombe à la moitié des crédits optimaux, c'est-à-dire lorsque le niveau de la contrainte budgétaire est fort. Bien évidemment, ce critère, lorsqu'il permet de trouver la bonne variante, ne donne pas sa bonne date de réalisation qui dépend étroitement de la contrainte budgétaire. D'après cette analyse, le choix technique de la variante peut être décentralisé, au moins tant que la contrainte financière n'est pas trop forte, mais la décision concernant sa date de réalisation ne peut être effectuée que par un échelon central qui aurait en main tous les projets.

La situation 3 correspond aux tests 6, 7, 8 et 9 et aboutit aux mêmes résultats.

La robustesse du critère du bénéfice actualisé pour le choix entre variantes sous contrainte budgétaire, que met en évidence cet essai, n'est bien sûr qu'une conjecture qui devrait être étayée par un balayage plus systématique des situations types (en particulier, mais de façon non exclusive, il conviendrait de faire varier les coûts des projets). De toute façon, la connaissance des prix fictifs tels qu'ils sont définis par J. Maurice (cf. chapitre 9) permet d'obtenir un critère plus fiable encore.

2.3.3. *Tests de programmation centralisée sur projets engendrés aléatoirement*

Dans ces tests, les paramètres des projets (coût d'investissements et leurs variations dans le temps, avantages socio-économiques et recettes financières ainsi que leurs variations dans le temps) ont été tirés au sort de façon aléatoire autour de valeurs moyennes représentatives des projets routiers et ferroviaires moyens.

Avec ces valeurs numériques, le taux de rentabilité interne socio-économique des projets varie entre 3,2 % et 8,6 %. La part des fonds publics si la mise en service est en 2005 varie entre 60 % et 85 %.

Ce panel de projets a été utilisé pour explorer la valeur des indicateurs approchés pour la programmation, l'effet du mode de financement, et les conséquences de la valeur du coût des fonds publics sur les indicateurs de la programmation.

Indicateurs approchés pour la programmation

Pour des projets générés aléatoirement selon les règles énoncées ci-dessus, on peut examiner les corrélations entre les dates de mise en service optimales d'une part, et les indicateurs simplifiés présentés plus haut, les TRI, Tri et ratios B/fonds publics, tous calculés à la date de programmation (ici 2005).

Le tableau suivant montre la manière dont les dates de mise en service sont corrélées avec ces indicateurs :

Tableau 2
**Corrélations entre les dates de mise en service
et les indicateurs de rentabilité usuels**

Indicateurs de rentabilité	TRI SE	Tri FI	Tri SE	B /FP
Coefficient de corrélation R ² avec la date de réalisation	0,64	0,36	0,73	0,39

Globalement les projets qui présentent le meilleur ratio « B/fonds publics » sont en général ceux qu'il faut réaliser d'abord, mais ce n'est qu'une relation approchée qui comporte une marge d'erreur importante, puisque le coefficient de corrélation entre le ratio et la date de mise en service n'est que de $R^2 = 0,39$.

La corrélation entre le TRI interne socio-économique et la date optimale de mise en service est meilleure ($R^2 = 0,64$). Celle avec le taux de rentabilité immédiat socio-économique (si la mise en service est en 2005) l'est encore plus ($R^2 = 0,73$).

En revanche le taux de rentabilité financier immédiat est faiblement corrélé avec la date de mise en service optimale ($R^2 = 0,36$).

Effet du mode de financement sur la VAN socioéconomique

Avec un même coefficient de coût des fonds publics, le bénéfice socio-économique (prenant en compte les distorsions liées aux modes de collecte des fonds publics) varie selon le mode de financement des projets et le signe de la différence dépend de la différence entre le taux d'actualisation public et le coût moyen du capital : si le taux d'actualisation public est inférieur au coût moyen du capital privé, le financement public est systématiquement préférable au financement privé, comme on le voit sur l'exemple simple suivant, où l'on compare les bénéfices issus de deux modes de financement d'un projet constitué par une dépense initiale de 100 suivi par un avantage socio-économique qui se produit la première année seulement, et ensuite le

projet est hors service. Cet avantage est de 110, constitué uniquement de recettes financières. Prenons un taux d'actualisation public de $r = 0,04$ et un coût du capital privé de $i = 0,05$. En cas de financement privé, l'investissement comme les recettes bénéficient à un agent privé. Le projet n'a pas besoin de subvention publique et le bénéfice est :

$$B_{pr} = -100 + 110/(1,04) = 106$$

En cas de financement public, l'investissement est financé par un opérateur public et les recettes bénéficient à cet opérateur public. Le bénéfice doit être calculé en fonction du coût des fonds publics. Les dépenses sont à multiplier par 1,3 mais les recettes aussi :

$$B_{pu} = -100 * 1,3 + (110/1,04) * 1,3 = 108$$

Ceci est lié au fait qu'alors les avantages socio-économiques en cas de financement public doivent tenir compte des distorsions économiques évitées du fait du retour à la sphère publique des avantages financiers du projet. Ces avantages socio-économiques supplémentaires d'exploitation sont actualisés au taux public, en général inférieur au taux financier (WACC¹⁶). En revanche, pour un financement public, l'ensemble de l'investissement entraîne une distorsion économique. Globalement, le bilan du financement privé est alors négatif, toutes choses égales d'ailleurs. On obtiendrait le résultat inverse si le coût moyen du capital est inférieur au taux d'actualisation public : alors le financement privé est systématiquement préférable au financement public.

Il ne faut pas en déduire une supériorité systématique d'un des modes de financement sur l'autre. En effet l'analyse précédente ne tient pas compte de toutes les différences entre financement public et financement privé. Pour effectuer une comparaison complète et significative, il faudrait prendre en compte explicitement les particularités de chaque mode de financement, à savoir notamment un moindre coût d'exploitation et de construction pour le financement privé, introduire le niveau de risque de chaque projet qui se traduira en financement privé par un coût moyen du capital spécifique à chaque projet et en financement public par une prise en compte explicite (à travers par exemple des valeurs d'option) de l'incertitude, et enfin tenir compte des questions non abordées ici d'asymétrie d'information, de régulation et d'incitations (Pour un traitement de ces questions, voir le programme de recherche européen REVENUE, Deliverable D2).

16. Weighted average cost of capital.

Effet du coefficient d'opportunité des fonds publics sur les indicateurs usuels de rentabilité

Le tableau ci-dessous présente l'effet de la variation du coefficient des fonds publics sur les coefficients de corrélation R2 entre la date optimale et divers indicateurs. Aucune contrainte budgétaire annuelle n'a été appliquée.

Coût d'opportunité des fonds publics	B/FP	TRI SE	Tri SE	Tri Fi
0	0,30	0,55	0,37	0,03
0,3	0,30	0,51	0,41	0,22
0,6	0,26	0,46	0,46	0,23
0,9	0,31	0,45	0,51	0,32

Dans la plupart des cas, le TRI interne socio-économique semble le meilleur indicateur approché de la date de mise en service. Le taux de rentabilité immédiat financier est d'autant meilleur que le coefficient des fonds publics est élevé.

2.3.4. Test de programmation centralisé sur projets réels

Ce test a été effectué sur les principaux projets ferroviaires relatifs à des lignes nouvelles actuellement envisagés en France ; ces projets, au nombre d'une vingtaine, sont, conformément aux dispositions légales étudiés par le gestionnaire des infrastructures ferroviaires, Réseau Ferré de France (RFF). Cet établissement public doit statutairement équilibrer son budget, et ne peut donc financer les projets qu'à hauteur des recettes qu'il est susceptible d'en tirer par les péages qu'il perçoit auprès des opérateurs ferroviaires, ou des réductions de coûts permises par les projets.

Il paraît dans ces conditions normal de considérer que RFF est un investisseur privé, c'est-à-dire que les recettes qu'il perçoit et la part du coût des investissements qu'il finance ne doivent pas être affectées du coefficient de rareté des fonds publics. Les recettes financières correspondantes permettent de gager un capital qu'il convient de déterminer en prenant un taux d'intérêt égal au taux des emprunts que le marché propose à RFF, taux d'intérêt qui se situe aux alentours de 5 % en euros courants, soit 3,3 % en euros constants : c'est ce coût moyen du capital qui a été pris dans la suite de cette section. Le reste du coût d'investissement est assuré par des subventions publiques, auxquelles doit être appliqué le coefficient de rareté des fonds publics.

Seule est prise en compte parmi les investissements et les avantages financiers la part du gestionnaire d'infrastructure.

Ces données permettent de calculer des avantages socio-économiques et financiers la première année en cas de mise en service en 2005.

Dans les cas où ces informations n'étaient pas complètes, des estimations ont été effectuées ; en particulier, en l'absence de parts des fonds publics, le ratio des avantages financiers sur les avantages socio-économiques de la première année a été fixé à 32 % environ (ratio moyen constaté sur les projets où l'on peut calculer ce ratio).

Le taux de croissance des avantages socio-économiques est supposé égal à 2,28 % (en euros constants), correspondant à une hypothèse de taux de croissance annuel moyen du PIB de 1,9 % (scénario central des prévisions de trafic établies par le SESP (Service Économie, Statistiques et de la Prospective) pour le compte du ministère des Transports) et d'une élasticité du trafic au PIB de 0,5 ainsi que d'une élasticité de la valeur du temps au PIB de 0,7, conformément aux recommandations du rapport dit « Boiteux II »¹⁷.

Le taux de croissance des avantages financiers est supposé identique à celui des avantages socio-économiques, le gestionnaire d'infrastructures étant supposé conserver au cours du temps une part constante des avantages socio-économiques¹⁸.

Indicateurs approchés pour la programmation

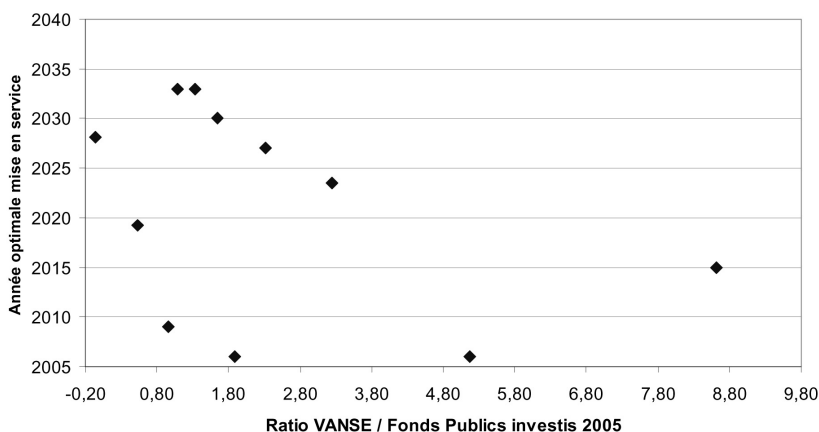
Pour des projets ferroviaires décrits ci-dessus, on peut aussi examiner les corrélations entre les dates de mise en service d'une part, et divers indicateurs simplifiés d'autre part.

Pour ce panel de projets, la relation entre la date optimale de mise en service calculée par le modèle (avec $k = 30$ % et une contrainte budgétaire annuelle correspondante), et le ratio VANSE/fonds publics

17. Commissariat Général au Plan, *Transports : choix des investissements et prise en compte des nuisances*, rapport d'un groupe de travail présidé par M. Boiteux, La Documentation Française, Paris, 2001.

18. Ceci suppose que le montant du péage du transporteur en volume (produit du péage unitaire par le nombre de trains) suit la croissance du produit du trafic par la valeur du temps. Autrement dit, le taux de croissance annuel moyen (TCAM) du péage unitaire p soit $(dp/dt)/p$ soit être égal à la somme de $(d\text{emport}/dt)/\text{emport}$ et de $(d\text{valtemps}/dt)/\text{valtemps}$, (car l'avantage temps par train = $\text{emport} * \text{valtemps} * \text{temps gagné}$) où emport est le nombre de voyageurs par train et valtemps la valeur unitaire du temps. Par exemple si la valeur du temps croît à 1,3 % par an (TCAM du PIB à 1,9 %), et si l'emport par train croît avec un TCAM de 2,5 % par an (du fait notamment de l'augmentation de la part des trains en unité multiple (UM) et à deux niveaux) (niveau estimé de 1996 à 2004 pour les TGV sur les troncs communs des LGV Sud-Est et Atlantique), le péage unitaire par train sur les LGV doit croître à environ $1,3 \% + 2,5 \% = 3,8 \%$ par an pour que cette hypothèse soit vérifiée.

Année optimale de mise en service et ratio Lebegue en 2005



en cas de mise en service en 2005 est décrite par le graphique ci-dessus. Chaque point représente un projet.

On constate que la relation entre ces deux quantités n'est qu'approximative. Ceci est lié au fait que des hypothèses de plus forte courbure de la courbe des avantages dans le temps ont été utilisées pour les projets ayant un caractère de décongestion marqué. Il en va notamment de projets comme l'augmentation de capacité de la ligne nouvelle Paris Sud-Est (LN 1) (ERTMS et investissements connexes), ou de projets visant à augmenter la capacité pour le fret ferroviaire dans des nœuds critiques.

Ainsi, par exemple, le projet augmentation de capacité de la LN1 présente le meilleur ratio Bénéfice actualisé sur fonds publics en 2005, mais sa date de mise en service optimale ne se situe que vers 2015, du fait de la croissance quadratique ou cubique des avantages avec le temps. Ceux-ci sont en effet proportionnels non pas au nombre de trains global sur cet axe, mais au nombre pour lesquels la capacité horaire de l'infrastructure en référence est dépassée. Il en irait de même pour des projets urbains dont les avantages sont essentiellement liés à la décongestion routière des voies parallèles saturées, ou bien, dans son principe, de projets de désaturation de nœuds ferroviaires de fret, même si l'horizon de cette désaturation est vraisemblablement plus éloigné.

On peut également penser que l'augmentation des coûts d'investissements en fonction de la date de mise en service n'est probablement pas la même pour un projet en rase campagne et un projet en milieu urbain, où le foncier et l'insertion environnementale vont coûter de

plus en plus cher ; une majoration des coefficients a été introduite pour ces projets en milieu urbain.

Sur ces données, la meilleure corrélation avec la date optimale de mise en service est celle donnée par le taux de rentabilité immédiat socio-économique (avec mise en service en 2005 pour tous les projets) ($R^2 = 0,83$). Celle avec le taux de rentabilité interne socio-économique (si la mise en service est en 2005) vient ensuite ($R^2 = 0,53$), puis ensuite le ratio VAN socio-économique par euro public investi ($R^2 = 0,15$).

Effet du mode de financement sur la VAN socio-économique

Avec un même coefficient (coût d'opportunité) des fonds publics, le Bénéfice socio-économique (prenant en compte les distorsions liées aux fonds publics) varie peu selon le mode de financement des projets ; elle est en cas de financement public de 30,84 milliards d'euros 2005 contre 30,89 milliards d'euros 2005 en cas de financement « privé », dans le cas d'un emprunteur AAA (avec un taux financier WACC de 3,3 % en euros constants soit 5 %).

Cette différence est liée au fait que le taux d'actualisation public (4 % en première période) est supérieur au taux retenu pour le taux de la ressource financière (3,3 % en euros constants, soit 5 % en euros courants, avec une inflation à 1,7 %).

Si le taux d'actualisation public était plus faible que le taux financier, le résultat serait inverse. Par exemple, si on prend un taux financier WACC de l'ordre de 4,8 % en euros constants, soit 6,5 % en euros courants (soit 1,5 % au-dessus d'un emprunteur AAA), le bénéfice socio-économique serait de 30,77 milliards d'euros en cas de financement « privé », contre 30,84 en cas de financement public.

L'effet du mode de financement sur le bénéfice socio-économique est donc principalement lié à d'autres facteurs (taux de la ressource financière différent selon le porteur du projet, capacité de celui-ci à mieux maîtriser les coûts et à obtenir davantage de recettes par une tarification appropriée notamment).

Comparaison avec la programmation gouvernementale

Lors du CIADT (Comité Interministériel d'Aménagement et de Développement du Territoire) de décembre 2003, le gouvernement français a affiché¹⁹ une programmation de projets dans l'ensemble des modes de transport. Pour ce qui concerne les projets ferroviaires, ont

19. [http://www.datar.gouv.fr/datar—site/datar—CIADT.nsf/\\$ID—Dossier/CLAP-5UCCRY](http://www.datar.gouv.fr/datar—site/datar—CIADT.nsf/$ID—Dossier/CLAP-5UCCRY)

été publiées une carte ainsi que pour certains projets des dates de lancement des travaux comme suit :

Haut-Bugey :	2005
LGV Rhin-Rhône Est 1 ^{re} phase :	2006
LGV Sud Europe Atlantique 1 ^{re} phase :	2008
LGV Bretagne Pays de Loire :	2009
LGV Est Européenne 2 ^e phase :	2010

Le lancement des autres projets cités lors du CIADT en question n'est pas daté dans ce document.

La programmation optimale trouvée par le modèle aboutit pour les projets en cause aux dates indicatives suivantes :

LGV Rhin Rhône Est 1 ^{re} phase :	2006
Haut-Bugey :	2006
LGV Bretagne Pays de Loire :	2009
LGV Sud Europe Atlantique :	2019

L'absence de données précises sur la LGV Est 2^e phase ne permet pas une comparaison de dates, mais le programme optimal du modèle la situe à un horizon significativement plus éloigné.

On constate donc globalement une assez grande similarité entre les résultats du modèle (programme optimal) et la programmation gouvernementale du CIADT.

3. CONCLUSION

Ce texte propose une méthode d'optimisation des programmes d'investissements de transport. Cette méthode est fondée sur la programmation linéaire, et on montre qu'on peut par son entremise résoudre la plupart des problèmes que pose la programmation des investissements : choix entre variantes d'un même projet, détermination de la date optimale de mise en service, optimisation sous contrainte budgétaire, optimisation avec contrainte budgétaire aléatoire, ...

Cette méthode est mise en œuvre d'abord pour des projets simulés afin d'explorer quelques problèmes liés à la programmation.

Les simulations présentées ici fournissent des enseignements en ce qui concerne plusieurs problèmes tels que la qualité des indicateurs approchés ou la décentralisation des décisions en situation de contrainte budgétaire (rappelons qu'en l'absence de contrainte budgétaire, la décentralisation des décisions pourrait être complète).

Il apparaît d'abord que les critères simples de rentabilité tels que le taux de rentabilité interne, le taux de rentabilité immédiate, le ratio « bénéfice actualisé sur euro investi » ou le ratio « Bénéfice actualisé sur euro public investi » (en cas de mise en service en 2005) ne sont qu'approchés. Leur corrélation avec la date de mise en service est imparfaite, et c'est semble-t-il avec le taux de rentabilité immédiate que la corrélation est la plus forte, de l'ordre de 0,8. Le ratio « VAN/fonds publics » est d'autant meilleur que l'évolution des avantages présente une courbure plus faible. Mais ces critères simples, qui peuvent être parlants en termes de communication, ne sauraient être des outils opérationnels pour l'établissement des programmes, pour lesquels la méthode de programmation proposée est à la fois simple et rapide à mettre en œuvre.

Enfin, les tests effectués montrent qu'une attention plus grande devrait être apportée à la convexité de la courbe des avantages socio-économiques dans le temps (du fait par exemple d'une fonctionnalité de décongestion) ainsi qu'aux évolutions possibles des coûts d'investissement selon la date de mise en service, facteurs qui semblent avoir un effet important sur la programmation. Les projets dont les coûts d'investissement vont le plus augmenter devraient, toutes choses égales d'ailleurs, être lancés plus tôt, pendant qu'ils ne sont pas encore trop chers.

En matière de décentralisation des procédures d'élaboration des programmes, il semble que pour les choix entre variantes incompatibles d'un même projet le critère du bénéfice actualisé maximal sans tenir compte de la contrainte budgétaire (critère décentralisable) soit assez robuste même pour des niveaux de contrainte budgétaire élevés. Il semble aussi que pour la programmation de projets liés, les critères simples usuels ne donnent pas de bons résultats. En revanche dans ce cas, le critère du bénéfice actualisé sans tenir compte de la contrainte budgétaire (critère décentralisable) apparaît également assez robuste. Mais les critères dégagés par J. Maurice, qui font intervenir les variables duales des contraintes budgétaires, sont plus décentralisables encore (cf. chapitre 9).

D'autres simulations pourraient être menées. Une extension des simulations concernerait l'optimisation des péages d'infrastructure ; à travers cette variable et en fonction de sa valeur, les avantages socio-économiques et les avantages financiers de chaque année seraient variables, et liés, et pourraient être optimisés. Il faudrait en outre bien sûr d'autres essais pour confirmer ou infirmer les conclusions provisoires tirées des simulations présentées.

Références

- Abraham C. et Laure A., « Étude des programmes d'investissements routiers », *Annales des Ponts et chaussées*, novembre, 1959.
- Bonnafeous A. et Jansen P., « Ranking Transport Projects by their Socioeconomic Value or Financial Interest rate of return ? », First Conference on Railroad Industry Structure, Competition and Investment, Toulouse, November 7-8, 2003.
- Bonnafeous A. et Jensen P., « Ranking transport projects by their socioeconomic value or financial internal rate of return ? », *Transport Policy*, vol. 12, n° 2, 2005.
- Commissariat Général du Plan, *Transports : pour un meilleur choix des investissements*, rapport d'un groupe de travail présidé par M. Boiteux, La Documentation Française, Paris, 1994.
- Commissariat Général du Plan, *Transports : choix des investissements et prise en compte des nuisances*, rapport d'un groupe de travail présidé par M. Boiteux, La Documentation Française, Paris, 2001.
- Commissariat Général du Plan, *Révision du taux d'actualisation des investissements publics*, Rapport d'un groupe de travail présidé par D. Lebègue, rapporteur L. Baumstark, La Documentation Française, Paris, 2005.
- Pozmantir W., « Méthode de correction des tâches et des contraintes sur un programme d'investissement de branche », Working Paper, Institute of Studies of Complex Systems, 1999.
- Proost S., de Palma A., Lindsey R., Balasko Y., Meunier D., Quinet E., Doll C., van der Hoofd M., Pires E., « REVENUE D2 Theoretical Framework », 2005.
- Relevé de décisions du CIADT du 13 Décembre 2003 :
« [http://www.datar.gouv.fr/datar—site/datar—CIADT.nsf/\\$ID—Dossier/CLAP-5UCCRY](http://www.datar.gouv.fr/datar—site/datar—CIADT.nsf/$ID—Dossier/CLAP-5UCCRY)
- Quinet E., *Principes d'économie des transports*, Economica, 1998.
- Quinet E., « À propos du rapport Lebègue » *Transports*, août-septembre, 2005.
- SESP, « La demande de transport en 2025 », Ministère de l'Équipement, du Logement des Transports et de la Mer.

CHAPITRE 9

CHOIX DES PROJETS SOUS CONTRAINTE BUDGÉTAIRE ANNUELLE : ESSAI DE RÉCAPITULATION

Joël Maurice¹

La présente contribution fait suite à l'introduction générale relative à la programmation des projets sous contrainte budgétaire, ainsi qu'à la contribution de Bonnafous et Roy et à celle de Quinet et Sauvant.

Son objet est de partir de ces travaux récents ou de certains travaux datant des années 80 (rapports du Plan, article de Thédié et article de Laure et Abraham) pour en proposer une analyse et une interprétation transversales, susceptibles de fournir une méthode applicable dans la plupart des cas et de façon décentralisée. Il ne s'agit pas d'une lecture nouvelle, mais d'un essai de récapitulation.

La représentation du problème utilisée ci-après recourt à la fiction du fractionnement des projets, que Laure et Abraham préconisaient et que Bonnafous et Roy ont également adoptée, tout comme Quinet et Sauvant ; des exemples simples permettront toutefois de signaler les écarts qui peuvent être engendrés par l'abandon du caractère insécable des projets. La démarche qui va être présentée est pleinement compatible avec celle de Quinet et Sauvant, qui font appel à la programmation linéaire, mais elle en interprète les résultats à la lumière du théorème de Kuhn et Tucker ; en cela, elle prolonge la démarche de Bonnafous et Roy, en la généralisant notamment à un horizon pluriannuel.

La conclusion à laquelle on parvient est qu'il serait possible de décentraliser pour chaque projet le choix de la variante optimale ainsi que de la date optimale d'achèvement, dans le respect de la contrainte budgétaire annuelle, si toutefois les pouvoirs publics étaient en mesure d'annoncer une chronique de « prix fictifs des fonds publics », comme

1. Paris-Jourdan Sciences Économique – École Nationale des Ponts et Chaussées

le proposait déjà Thédié. Dans le cas contraire, on examine la pertinence d'autres indicateurs, tel que notamment le bénéfice actualisé par euro public dépensé ; cet indicateur auxiliaire se révèle utile si un certain nombre de simplifications apparaissent admissibles, mais il n'a qu'un pouvoir séparateur imparfait.

Enfin, on explore la piste d'une optimisation de l'appel aux financements privés dont le remboursement sur fonds publics serait gagé sur la chronique supposée connue des enveloppes budgétaires disponibles. Cette démarche conduirait à une valeur unique du « prix fictif de rareté des fonds publics » au lieu d'une série de prix fictifs annuels et donnerait ainsi une justification théorique à l'usage du ratio bénéfice actualisé par euro public dépensé actualisé (à l'aide le taux d'intérêt réel du préfinancement privé), dont le pouvoir séparateur reste cependant imparfait.

Les notations sont autant que possibles les mêmes que dans l'instruction cadre des 24 mars 2004 et 27 mai 2005. Le taux d'actualisation est notamment désigné par r .

Hypothèses simplificatrices

Pour procéder à cette analyse, on adopte un certain nombre de simplifications.

- On raisonne en temps discret, à pas annuel, à partir d'une année choisie comme origine des temps².
- On suppose que l'on connaît la liste exhaustive des N projets³.
- Chaque projet est repéré par son numéro n (n variant de 1 à N), placé en *exposant*. Il est caractérisé par :
 - sa période de réalisation, qui s'achève au 31 décembre de l'année dont la date⁴ est t_0^n
 - sa durée de réalisation Θ^n , que l'on supposera pour simplifier inférieure ou égale à une seule année : la réalisation se déroule donc intégralement au cours de l'année t_0^n
 - sa date de mise en service, qui est le 1^{er} janvier de l'année ($t_0^n + 1$)
 - son coût d'investissement I^n , que l'on suppose le même quelle que soit sa date de réalisation t_0^n , et dont la valeur actualisée à l'origine des temps est donc $I^n/(1+r)^{t_0^n}$

2. L'origine des temps peut être l'année 2004, comme dans l'instruction cadre des 25 mars 2004 et 27 mai 2005 du ministère des Transports, de l'Équipement, du Tourisme et de la Mer. Mais rien ne s'oppose à ce que ce soit une année plus récente, comme par exemple 2006.

3. Hypothèse très irréaliste, car il est presque consubstantiel aux projets d'apparaître et de se développer (ou de disparaître) au fur et à mesure de l'histoire socio-économique ; il est donc paradoxal d'insérer les projets dans une liste close.

4. Par rapport à l'année origine des temps.

- sa période d'exploitation, dont l'année générique t démarre à la date $(t_0^n + 1)$ et qui est censée se prolonger⁵ à l'infini ; chacune de ces années procure un avantage socio-économique a_t^n , que l'on suppose⁶ ne dépendre que de l'année t et dont la valeur actualisée à l'origine des temps est donc $a_t^n/(1+r)^t$

- une série de dépenses publiques annuelles (qui, lorsqu'elles prennent un signe négatif, constituent en réalité des recettes publiques). À la date t_0^n , la dépense publique représente la contribution publique au financement de l'investissement I^n ; comme le coût de cet investissement, on suppose que cette dépense publique est la même quelle que soit la date de réalisation du projet ; on la note DPI^n . L'année générique t , postérieure à t_0^n , la dépense publique est la contribution publique au financement de l'exploitation ; comme l'avantage socio-économique, on suppose que cette dépense ne dépend que de l'année t ; on la note DPE_t^n ; elle peut être négative en cas de recettes nettes tarifaires ou fiscales ;

- son bénéfice socio-économique actualisé à l'année t_0^n noté $B_{t_0}^n$ et son bénéfice actualisé l'origine des temps noté $BOT_{t_0}^n$;

- Nota : dans le calcul de $B_{t_0}^n$ comme dans celui de $BOT_{t_0}^n$, en application de l'instruction cadre les dépenses ou recettes publiques DPI^n et DPE_t^n sont à multiplier par 1,3. La majoration de 0,3 est appelée « **coût d'opportunité des fonds publics** » ; elle reflète les pertes d'efficacité économique (distorsions) censées résulter du recours à l'impôt.

– Enfin, chaque année t est censé exister un plafond C_t que ne peut pas dépasser le solde des dépenses publiques, relatives à ceux des N projets considérés qui ont été réalisés depuis l'origine des temps et jusque, inclusivement, à ladite année t .

Remarque : on reviendra plus loin (§ 4) sur les particularités introduites dans ces hypothèses en cas de **partenariats public privé**.

Fiction du fractionnement des projets

Comme déjà dit en introduction, on adopte ici d'emblée la fiction consistant à supposer qu'il est possible de réaliser, à la date générique⁷ t , une fraction x_t^n (comprise au sens large entre 0 et 1) du projet n . Cette réalisation partielle est alors censée entraîner que les

5. On pourrait considérer la durée T^n de vie comme un autre paramètre du projet, mais cela ne paraît pas indispensable à ce stade, sans réelle perte de généralité.

6. Cet avantage est évidemment nul pour toute année $t \leq t_0^n$. Mais pour $t > t_0^n$, il est ici censé ne pas dépendre de t_0^n , ce qui est une hypothèse simplificatrice relativement forte. On pourrait penser en effet qu'en l'an 2020, l'avantage ne sera pas exactement le même selon que le projet aura été mis en service en 2000 ou en 2010.

7. $t \geq 0$

diverses caractéristiques susmentionnées (avantages, financements publics, etc.) de ce projet se réalisent elles-mêmes dans la même proportion x_t^n , chacune des années postérieures à l'achèvement de ce projet fractionnaire.

À chaque projet n se trouve ainsi associée la série des fractions inconnues x_t^n .

On suppose en outre que la somme des fractions ne peut excéder la réalisation intégrale du projet : $\sum_{t=0}^{\infty} x_t^n \leq 1$

Plusieurs cas peuvent alors se présenter. Le projet peut être réalisé intégralement (cas $\sum_{t=0}^{\infty} x_t^n = 1$), partiellement (cas $\sum_{t=0}^{\infty} x_t^n < 1$) ou pas du tout (cas $\sum_{t=0}^{\infty} x_t^n = 0$ qui implique $x_t^n = 0$ pour tout t).

Le recours à cette fiction ne sera pas gênant si, comme cela sera le cas quasi général, on trouve *in fine* que le projet n ou bien ne doit jamais être réalisé, ou bien doit être intégralement réalisé une année « optimale » bien déterminée \hat{t}^n . Si l'on trouve, ce qui sera exceptionnel, que le projet doit effectivement être fractionné sur plusieurs années, on devra discuter l'interprétation à donner à ce résultat dans chaque cas d'espèce ; on y reviendra plus loin (§ 6).

Modélisation de la programmation optimale

Le problème consiste à choisir la programmation des projets de façon à maximiser, sous la contrainte annuelle de disponibilité des fonds publics, le **montant total des bénéfices socio-économiques, actualisés à l'origine des temps,**

Le choix de cet objectif a été implicitement considéré comme allant de soi dans les différents travaux précités, remontant aux années 80 ou plus récents. Il suppose implicitement que le taux d'actualisation lui-même peut être déterminé indépendamment des contraintes budgétaires, qui peuvent être introduites dans une seconde étape, sans réaction sur le taux d'actualisation ; cette séparabilité ne s'impose pas comme une évidence, mais nous l'adopterons dans ce qui suit comme une approximation admissible. De même, l'objectif retenu consiste en un surplus monétarisable, dont on sait qu'il ne reflète l'utilité socio-économique que sous certaines hypothèses, satisfaites si la fonction d'utilité des consommateurs est de type semi-linéaire séparable⁸, ce que nous supposons également.

8. Voir Laffont J.-J., *Cours de théorie micro-économique*, Economica, 1982, volume 1, page 115.

Grâce à la fiction consistant à adopter pour inconnues les fractions x_t^n , le problème devient, comme Quinet et Sauvart l'ont proposé et mis en œuvre dans leur contribution, un exercice de **programmation linéaire** auquel il est possible d'appliquer **l'algorithme du simplexe**. Cette méthode très opérationnelle se suffit à elle-même, du moins au niveau central, s'il est possible de rassembler toutes les caractéristiques de tous les projets. En revanche elle ne peut pas en tant que telle être décentralisée auprès de maîtres d'ouvrages multiples.

C'est précisément pour étudier les **possibilités de décentralisation** du choix des investissements en présence de rationnement annuel des fonds publics que nous nous proposons d'approfondir l'analyse, en ayant recours au **théorème de Kuhn et Tucker** relatif à l'optimisation sous contraintes, que nous appliquerons au cas de la programmation linéaire qui vient d'être évoquée. Cette analyse va nous permettre de dégager quelques éléments interprétatifs en termes de hiérarchisation des projets.

Pour la mise en forme du problème, on pourra se reporter à **l'annexe 1**.

On en retiendra ici qu'il apparaît trois types de **variables duales** (ou « multiplicateurs de Lagrange ») **associées chacun à une contrainte**.

α_t^n qui est : nul si la totalité ou une fraction non nulle du projet n est à réaliser l'année t ; positif⁹ si aucune fraction du projet n n'est à réaliser l'année t

β^n qui est : nul si – au cours du temps – le projet n n'est pas à réaliser en entier (il en manquera toujours au moins une fraction, voire la totalité) ; positif¹⁰ si – au cours du temps – le projet est à réaliser en entier (fût-ce par fractions sur plusieurs années¹¹)

φ_t qui est : nul si les fonds publics disponibles l'année t excèdent cette année-là les dépenses publiques nettes des recettes publiques, afférentes aux projets programmés ; positif¹² si les fonds publics rationnent effectivement les dépenses publiques nettes considérées. Ce multiplicateur φ_t est la **variable duale associée à la contrainte budgétaire de l'année t** et il est ici proposé de la dénommer **prix fictif de rareté des fonds publics**, l'année t .

9. Ou éventuellement nul, mais cet événement peut être considéré comme étant de probabilité nulle.

10. Ou éventuellement nul, mais cet événement peut être considéré comme étant de probabilité nulle.

11. Dans ce cas, si ces années n'étaient pas consécutives, la solution mathématique serait dépourvue de signification socio-économique.

12. Ou éventuellement nul, mais cet événement peut être considéré comme étant de probabilité nulle.

À noter que ces « prix de rareté des fonds publics » sont distincts et indépendants du « coût d'opportunité des fonds publics » mentionné plus haut, qui est fixé à 0,3 par l'instruction cadre et qui reflète les distorsions socio-économiques liées au financement par l'impôt. On n'aura plus besoin dans ce qui suit de se référer à ce « coût d'opportunité des fonds publics » puisque, comme on l'a vu, il n'intervient qu'en amont, dans le calcul du bénéfice actualisé à l'origine des temps $BOT_{t_0}^n$, noté plus simplement BOT_t^n lorsqu'il n'y a pas d'ambiguïté.

1. PRÉCONISATION DE BASE : MAXIMISER POUR CHAQUE PROJET n LE BÉNÉFICE FICTIF ACTUALISÉ À L'ORIGINE DES TEMPS

1.1. Définition du bénéfice fictif actualisé à l'origine des temps

Considérons le projet n . Supposons le réalisé au cours de l'année¹³ t , soit B_n^t son bénéfice actualisé à la date t et BOT_t^n son **bénéfice actualisé à l'origine des temps**, ces deux versions du bénéfice actualisé étant calculées conformément aux prescriptions de l'instruction cadre précitée.

On peut montrer (voir **annexe 1**) que la programmation optimale doit respecter les équations suivantes¹⁴ :

$$BOT_t^n - \varphi_t \cdot \frac{DPI^n}{(1+r)^t} - \sum_{u=t+1}^{\infty} \varphi_u \cdot \frac{DPE_u^n}{(1+r)^u} = \beta^n - \alpha_t^n \quad \text{pour tout } n \text{ et tout } t \quad (1)$$

Supposons pour le moment (on rediscutera ce point un peu plus loin) que les pouvoirs publics centraux aient déterminé la **chronique des prix fictifs de rareté des fonds publics** φ_t et l'aient diffusée auprès des maîtres d'ouvrage publics décentralisés.

Chaque maître d'ouvrage public décentralisé serait donc en mesure de calculer pour tout projet n de sa compétence la valeur du premier membre de la relation (1), que nous proposons d'appeler « **bénéfice fictif actualisé à l'origine des temps** $BFOT_t^n$ » dudit projet :

$$BFOT_t^n = BOT_t^n - \varphi_t \cdot \frac{DPI^n}{(1+r)^t} - \sum_{u=t+1}^{\infty} \varphi_u \cdot \frac{DPE_u^n}{(1+r)^u} \quad (2)$$

13. Pour alléger les notations, l'année de construction du projet est désormais désignée dans ce qui suit par t et l'année générique d'exploitation par u , notations qui se substituent respectivement aux notations t_0 et t de l'instruction cadre.

14. Dans cette relation, u désigne l'année générique d'exploitation de la fraction de projet supposée réalisée l'année t .

Avec cette définition (2), la relation (1) devient donc :

$$BFOT_t^n = \beta^n - \alpha_t^n \quad \text{pour tout } n \text{ et tout } t \quad (3)$$

Elle est au cœur de la discussion qui va suivre. On va l'exploiter successivement pour rechercher la date optimale de réalisation d'un projet donné, ensuite pour rechercher la variante optimale d'un projet donné. On dissociera ces deux étapes pour simplifier l'analyse, sans perte de généralité.

Remarque : dans le cas où il n'y a jamais aucun rationnement des fonds publics, alors le prix fictif de rareté des fonds publics est nul quel que soit u . Il en résulte que $BFOT_t^n$ n'est alors autre que BOT_t^n , quelle que soit la date t .

1.2. Recherche de la date optimale de réalisation du projet n

On suppose ici que les N projets candidats sont **indépendants**, c'est-à-dire que de l'un quelconque d'entre n'interfère aucunement avec les autres, sauf à travers l'imputation sur l'enveloppe des fonds publics disponibles.

Le cas de projets incompatibles sera traité plus loin sous la forme particulière du choix entre variantes d'un même projet. Enfin, le cas intermédiaire de projets présentant entre eux des degrés divers de complémentarité est traité dans la contribution de Quinet et Sauvart, mais on en fera abstraction ici pour alléger la présentation, sans perte de généralité.

1.2.1. Exploitation de la relation (3) ci-dessus

Interprétons la relation (3) dans les deux cas (a) et (b) ci-après qui sont les plus courants, les cas particuliers étant évoqués ensuite.

a) Supposons en premier lieu que, selon le processus de programmation linéaire, le projet ne doit jamais être réalisé. Cela implique que $\beta^n = 0$ et que $\alpha_t^n > 0$ pour toute date de réalisation t envisagée. Selon (3), le bénéfice fictif actualisé à l'origine des temps $BFOT_t^n$ est alors négatif quelle que soit la date de réalisation t envisagée.

b) Supposons en second lieu que, selon le processus de programmation linéaire, le projet soit à réaliser intégralement. Cela implique que $\beta^n > 0$.

Supposons en outre que, selon le processus de programmation linéaire, il convienne de réaliser l'ouvrage en une seule fois, au cours de l'année \hat{t}^n . Cela implique d'une part que cette année là $\alpha_{\hat{t}^n}^n$, d'autre part que, pour toutes les autres années de réalisation $t \neq \hat{t}^n$ envisagées, $\alpha_t^n > 0$.

Il résulte alors de la relation (3) que le bénéfice fictif actualisé à l'origine des temps est égal à β^n si le projet est réalisé l'année \hat{t}^n et que toute autre année de réalisation $t \neq \hat{t}^n$ envisagée, le bénéfice fictif actualisé à l'origine des temps prend la valeur $\beta^n - \alpha_t^n$, qui est inférieure à β^n . Autrement dit : l'année optimale de réalisation est celle pour laquelle le bénéfice fictif actualisé à l'origine des temps passe par son maximum, qui est positif¹⁵.

Cas particuliers

– Par rapport au cas (b), il peut se faire que, selon le processus de programmation linéaire, la réalisation intégrale du projet n se répartisse sur deux années \hat{t}_1^n et \hat{t}_2^n (voire sur plus de deux années), ce qui implique que pour ces années là le bénéfice fictif actualisé à l'origine des temps prenne la même valeur β^n . Cette solution peut être admissible si les années en cause sont consécutives, car on peut l'interpréter comme une réalisation unique à cheval sur deux ou plusieurs exercices budgétaires successifs. Elle n'aurait pas de sens concret dans le cas d'années non consécutives et il conviendrait alors d'éliminer de la programmation au moins l'une des dates rivales et de mieux cerner l'optimum, par tâtonnement.

– Il peut aussi se faire que, selon le processus de programmation linéaire, le projet n se trouve dans une situation intermédiaire entre les cas (a) et (b) ci-dessus, en ce sens qu'il serait à réaliser en partie seulement, au cours d'une année \hat{t}^n , sans être jamais complété. Cela implique que $\beta^n = 0$ et que $\alpha_{\hat{t}^n}^n > 0$ pour toute date de réalisation t envisagée, sauf pour l'année \hat{t}^n , pour laquelle $\alpha_{\hat{t}^n}^n = 0$. Une réalisation à jamais partielle du projet n'aurait guère de sens concret et il conviendrait en pareil cas de mieux cerner l'optimum, par tâtonnement. À noter que dans leur exercice de simulation, Quinet et Sauvart n'ont jamais rencontré ce cas.

1.2.2. Règle proposée pour déterminer la date optimale de réalisation

Au terme de cette recherche de la date optimale de réalisation d'un projet n en présence de contraintes budgétaires annuelles, **il apparaît donc que, en utilisant le bénéfice fictif actualisé à l'origine des temps $BFOT_t^n$, il convient d'abord d'éliminer toute date de réalisation pour laquelle $BFOT_t^n$ serait négatif, puis, au sein des dates de réalisation restant en lice, de retenir celle qui maximise $BFOT_t^n$.**

15. Ou, exceptionnellement, nul.

Cette règle ne fait donc que **généraliser la règle habituelle de choix de la date optimale de réalisation d'un projet**, en maximisant le bénéfice fictif actualisé à l'origine des temps, qui est lui-même une généralisation du bénéfice actualisé à l'origine des temps (auquel il se ramène lorsque le prix fictif de rareté des fonds publics est nul, c'est-à-dire lorsqu'il ne s'exerce aucune contrainte budgétaire).

1.2.3. Observation sur le rôle respectif des multiplicateurs

Les multiplicateurs α_t^n et β^n jouent un rôle théorique déterminant dans la démonstration ci-dessus. Mais ils ne jouent aucun rôle pratique dans calcul du maximum de $BFOT_t^n$, qui en revanche fait intervenir les prix fictifs de rareté des fonds publics φ_t . Pour mettre en œuvre la préconisation ci-dessus, le maître d'ouvrage décentralisé n'a donc pas à se soucier des multiplicateurs α_t^n et β^n , mais il a absolument besoin de connaître la chronique des prix fictifs de rareté des fonds publics φ_t . On reviendra plus loin sur la question de l'évaluation et de la diffusion de la chronique de φ_t par les pouvoirs publics centraux.

1.2.4. Remarque sur le critère auxiliaire « taux de rentabilité immédiate »

En l'absence de contraintes budgétaires, sous un certain nombre d'hypothèses simplificatrices¹⁶,

on peut utiliser le critère auxiliaire du taux de rentabilité immédiate pour déterminer la date optimale de réalisation du projet n , qui est la première année t telle que $(a_t^n / I^n) \geq r$ (4).

Cette règle peut s'obtenir en considérant la variation du bénéfice actualisé à l'origine des temps entre l'année de réalisation $(t - 1)$ et l'année de réalisation t . Un calcul simple montre que cette variation est

telle que : $BOT_t^n - BOT_{t-1}^n = \frac{r \cdot I^n - a_t^n}{(1+r)^t}$. Il est avantageux de retarder la

date de réalisation de l'année $(t - 1)$ à l'année t tant que la variation du bénéfice actualisée à l'origine des temps augmente. Il n'est plus avantageux de retarder la date de réalisation à partir du moment où cette variation devient négative, c'est-à-dire où le taux de rentabilité immédiate (a_t^n / I^n) devient supérieur au taux d'actualisation r .

16. Coût de l'investissement indépendant de la date de réalisation, avantage lui aussi indépendant de la date de réalisation et croissant au cours du temps.

En présence de contraintes budgétaires, le calcul comparable portant sur la variation du bénéfice fictif actualisé à l'origine des temps conduit à relation suivante :

$$BFOT_t^n - BFOT_{t-1}^n = \frac{[r \cdot I^n - (\varphi_t - (1+r) \cdot \varphi_{t-1}) \cdot DPI^n] - [a_t^n - \varphi_t \cdot DPE_t^n]}{(1+r)^t}$$

La condition (4), est remplacée en cas de contrainte budgétaire par la condition :

$$\frac{a_t^n - \varphi_t \cdot DPE_t^n + (\varphi_t - (1+r) \cdot \varphi_{t-1}) \cdot DPI^n}{I^n} \geq r \quad (5)$$

Il n'est donc pas étonnant que, en présence de contraintes budgétaires, la corrélation soit assez lâche entre la date optimale de réalisation du projet calculée selon la préconisation ci-dessus et la date de réalisation tirée du critère de rentabilité immédiate selon la relation (4). C'est bien ce qui est constaté dans la contribution de Bonnafous et Roy et dans celle de Quinet et Sauvart – même si, dans leur simulation, la corrélation est moins mauvaise qu'avec d'autres critères, dont celui du bénéfice actualisé par euro public ¹⁷.

1.3. Choix entre deux variantes d'un même projet ou entre deux projets incompatibles

Comme annoncé plus haut, on va examiner maintenant la question du choix entre deux projets incompatibles entre eux, dont un cas particulier est le choix entre deux variantes d'un même projet.

1.3.1. Modélisation

On va supposer que, comme ci-dessus, tous les projets sont indépendants, sauf les projets n° 1 et n° 2, qui sont incompatibles entre eux, par exemple parce qu'ils constituent deux variantes d'un même projet.

On recourt comme précédemment à la fiction du fractionnement des projets, affectés respectivement des fractions inconnues x_t^1 et x_t^2 .

Ces fractions doivent être positives ou nulles et il est à cet égard associé à chacune une variable duale (ou multiplicateur de Lagrange) égale respectivement à α_t^1 et α_t^2 , sans changement par rapport à la présentation faite jusqu'ici.

17. Dans le cas qui sera présenté en section 4 ci-après, la relation se simplifierait et deviendrait : $\frac{a_t^n - \varphi_t \cdot DPE_t^n}{I^n} \geq r$

En revanche, on doit substituer aux relations indépendantes $\sum_{t=0}^{\infty} x_t^1 \leq 1$ et $\sum_{t=0}^{\infty} x_t^2 \leq 1$ une relation exprimant que, si l'on réalise le projet n° 1, cela exclut le projet n° 2 et réciproquement. On propose pour cela de fusionner les deux relations en une seule, à savoir :

$$\sum_{t=0}^{\infty} (x_t^1 + x_t^2) \leq 1.$$

Cette contrainte interdit que l'on réalise complètement à la fois le projet 1 et le projet 2. Elle permet que l'on réalise entièrement le projet 1 si l'on renonce entièrement au projet 2 et réciproquement. Elle laisse subsister l'éventualité par exemple de réaliser chacun des deux projets à 50 %, cas fictif qu'il faudrait éliminer s'il était fourni par l'algorithme de programmation linéaire.

À cette contrainte est associée une variable duale β unique, qui se substitue aux deux variables duales distinctes β^1 et β^2 .

Il est alors possible comme précédemment d'utiliser la programmation linéaire pour chercher à maximiser le montant total des bénéfices actualisés à l'origine des temps sous la contrainte budgétaire annuelle : c'est ce que font Quinet et Sauvent dans leur contribution. On peut de même analyser cette démarche à la lumière du théorème de Kuhn et Tucker. On parvient ainsi à la conclusion (voir **annexe 2**) que la programmation optimale doit respecter les équations suivantes :

$$BFOT_t^1 = \beta - \alpha_t^1 \quad \text{pour tout } t \quad (6.1)$$

$$BFOT_t^2 = \beta - \alpha_t^2 \quad \text{pour tout } t \quad (6.2)$$

Dans ces relations, le premier membre est, respectivement pour le projet 1 et pour le projet 2, le bénéfice fictif actualisé à l'origine des temps, conformément à la définition (2) donnée plus haut.

1.3.2. Exploitation de la relation (6.1) et (6.2) ci-dessus

Supposons par exemple que, selon le processus de programmation linéaire, la variante n° 1 doive être réalisée entièrement. Cela implique que $\beta > 0$.

Supposons que la date optimale de cette réalisation soit \hat{t}^1 . Cela implique à la fois que $\alpha_{\hat{t}^1}^1 = 0$ et que $\alpha_t^1 > 0$ pour toutes les autres années $t \neq \hat{t}^1$. Donc l'année \hat{t}^1 est celle où $BFOT_t^1$ prend la valeur β , qui est sa valeur maximale.

Mais le choix de la variante n° 1 exclut celui de la variante n° 2, dont la fraction optimale x_t^2 est nulle quel que soit t , ce qui implique que $\alpha_t^2 > 0$ quel que soit t . Dès lors, selon (6.2), $BFOT_t^2 = \beta - \alpha_t^2$ est

inférieur quel que soit t à β , c'est-à-dire à la valeur maximale de $BFOT_t^1$.

1.3.3. Règle proposée pour choisir entre des projets incompatibles et notamment entre des variantes d'un même projet

Ce qui précède conduit à proposer la procédure simple suivante :

1. On cherche pour le projet n° 1 le maximum par rapport à la date de réalisation t du bénéfice fictif actualisé à l'origine des temps $BFOT_t^1$
2. On cherche pour le projet n° 2 le maximum par rapport à la date de réalisation t du bénéfice fictif actualisé à l'origine des temps $BFOT_t^2$
3. La plus grande de ces deux valeurs désigne simultanément la variante qu'il convient de retenir et l'année où il convient de la réaliser.

1.4. Observations sur la détermination de la chronique des prix fictifs de rareté des fonds publics φ_t

La détermination de la chronique des prix fictifs de rareté des fonds publics φ_t constitue sur le plan théorique une partie de la solution du problème d'ensemble de la programmation optimale sous contraintes.

En effet, la mise en équations de ce problème d'ensemble conduit à considérer le Lagrangien (voir l'**annexe 1** et le cas échéant l'**annexe 2**, déjà évoquées), dont les inconnues sont non seulement les fractions x_t^n , mais aussi les variables duales associées aux diverses contraintes, à savoir les multiplicateurs α_t^n , β^n , et – ce qui nous concerne plus particulièrement ici – les variables duales φ_t associées aux contraintes budgétaires annuelles. On obtient le système des conditions d'optimalité du premier ordre en annulant les dérivées partielles du Lagrangien par rapport à cet ensemble d'inconnues. S'il est « bien conformé », **ce système détermine simultanément les valeurs de toutes les inconnues et notamment celles qui nous intéressent directement : x_t^n et φ_t .**

En toute rigueur, l'élaboration de ce système nécessite de centraliser toutes les informations nécessaires et notamment pour chaque projet n considéré les estimations de :

$BFOT_t^n$ qui dépend de l'année de réalisation t envisagée,

DPI^n qui est supposée la même¹⁸ quelle que soit la l'année de réalisation t envisagée,

18. Cette hypothèse simplificatrice pourrait être levée sans difficulté, comme le font d'ailleurs Quinet et Sauvent dans leur contribution.

DPE_u^n qui est supposée ne dépendre que¹⁹ de l'année d'exploitation u (pourvu qu'elle soit postérieure à la date de réalisation t envisagée).

Ainsi posé, le problème paraît être inévitablement du ressort d'une planification centralisée.

Mais il est possible d'imaginer un processus itératif de dialogue entre les autorités publiques centrales et les autorités publiques décentralisées ou décentralisées.

– Dans un tel processus, les autorités centrales peuvent partir d'hypothèses d'amorçage considérées comme de premières approximations plausibles, tant pour la liste des projets et leurs caractéristiques pertinentes (coûts, financement, bénéfice actualisé à l'origine des temps) que pour la chronique annuelle de l'enveloppe budgétaire. Elles peuvent en déduire la chronique des variables duales associées aux contraintes, et tout particulièrement une chronique des prix fictifs de rareté des fonds publics φ_t . Elles peuvent enfin diffuser cette chronique φ_t auprès de chacun des maîtres d'ouvrages publics décentralisés.

– Chaque maître d'ouvrage décentralisé peut alors calculer pour chacun de ses projets autant de valeurs du bénéfice fictif actualisé à l'origine des temps qu'il y a de variantes (notamment variantes technologiques, mais aussi variantes de financement) et qu'il y a d'années de réalisation envisageables. Il peut ensuite pour chacun desdits projets, en maximisant le bénéfice fictif actualisé à l'origine des temps, définir la variante optimale et la date de réalisation optimale (qui peut ne pas exister, si le bénéfice fictif actualisé est constamment négatif, le projet étant alors à rejeter).

Le processus peut s'arrêter-là : il aurait déjà l'avantage d'éviter des écarts fâcheux entre les projets en matière d'appel aux fonds publics.

Mais il serait aussi imaginable de poursuivre l'itération :

– Chaque maître d'ouvrage décentralisé transmettrait ses résultats aux autorités publiques centrales, qui pourraient récapituler les besoins annuels de fonds publics et les comparer aux enveloppes disponibles.

– En cas de désajustement, elles pourraient calculer une nouvelle chronique de prix fictifs de rareté des fonds publics, et recommencer une boucle itérative, et ainsi de suite.

¹⁹ Remarque identique à celle qui précède.

S'arrêter à la première itération apparaît toutefois comme plus réaliste.

On notera que cette méthode est précisément celle qui est utilisée pour déterminer le taux d'actualisation. Et de la même manière, son application à la détermination de la chronique des prix fictifs de rareté des fonds publics devrait certainement faire appel à des groupes de travail réunissant de multiples compétences.

2 CRITÈRE AUXILIAIRE DU BÉNÉFICE ACTUALISÉ PAR EURO PUBLIC DÉPENSÉ

2.1. Définition du critère auxiliaire

Supposons maintenant que les pouvoirs publics centraux ne souhaitent pas déterminer ou diffuser la série des prix fictifs de rareté des fonds publics φ_t . Est-il alors possible de recourir à un critère auxiliaire pour sélectionner les projets sous contrainte budgétaire annuelle ?

Un tel critère auxiliaire est préconisé tant par le rapport du Plan de février 2005 que par l'instruction cadre²⁰ précité : à savoir, le **ratio bénéfice actualisé par euro public dépensé**.

Ces textes ne précisent pas avec quelle hypothèse concernant l'année de réalisation ce ratio doit être calculé. La contribution de Bonnafous et Roy et celle de Quinet et Sauvart ne sont pas non plus explicites sur ce point ; de fait, ces auteurs n'ont calculé ce ratio qu'en supposant le projet réalisé immédiatement, ils n'en ont pas examiné l'évolution en faisant varier la date hypothétique de réalisation.

Compte tenu des préconisations qui ont été formulées plus haut, il semble bien que l'acception la plus appropriée de ce ratio soit de le définir comme le **bénéfice actualisé à l'origine des temps par euro public dépensé**, en considérant **qu'il prend autant de valeurs que d'année t envisagée pour la réalisation du projet**.

Nous noterons ce critère auxiliaire CA_t^n . Il répondrait ainsi à la définition suivante :

$$CA_t^n = \frac{BOT_t^n}{DPOT_t^n} \quad (7)$$

20. Rappel : Extrait de l'instruction cadre, annexe III. « Afin de tirer le meilleur parti d'un financement public limité, la règle de classement des projets doit être non pas le bénéfice actualisé induit par le projet, mais le bénéfice actualisé par euro public dépensé, prenant en compte l'ensemble des dépenses publiques et des recettes publiques éventuelles au cours de la durée de la vie du projet (actualisés au taux de 4 %) ; en conséquence, on déterminera pour chaque projet, en sus des indicateurs de rentabilité socio-économique, le bénéfice actualisé pour la collectivité divisé par la valeur actualisée nette des dépenses budgétaires de toutes les collectivités publiques (nettes des recettes fiscales éventuellement générées), pendant la durée du projet. Cet indicateur pourra faire l'objet de l'analyse de sensibilité ».

où le **dénominateur est le montant des dépenses publiques (nettes des recettes publiques) actualisées à l'origine des temps, à l'aide du taux d'actualisation r** :

$$DPOT_t^n = \frac{DPI^n}{(1+r)^t} + \sum_{u=t+1}^{\infty} \frac{DPE_u^n}{(1+r)^u} \quad (8)$$

On va dans ce qui suit explorer comment et au prix de quelles approximations ce critère peut être dérivé de la méthode à portée générale développée dans la section 2 ci-dessus.

On notera d'emblée que ce critère auxiliaire ne peut avoir qu'un pouvoir discriminant inférieur au critère de choix principal présenté au paragraphe 2, en ce sens qu'il peut conduire à des ambiguïtés, voire à des erreurs, comme l'illustre le cas simple présenté en encadré. Mais son intérêt est de fournir une approche simple, qui peut être utile « en première approximation ».

Le critère bénéfico socio-économique par euro public dépense peut être ambigu.

Exemple simple

Soit deux projets A et B. Ils ont le même coût d'investissement égal à 1 €, quelle que soit la date de réalisation. Cet investissement est intégralement financé sur fonds public et le solde annuel d'exploitation en euros publics est nul par la suite. Dans cet exemple simple le bénéfico socio-économique actualisé à l'origine des temps fournit donc directement le ratio « bénéfico socio-économique par euro public dépensé ».

Supposons que l'enveloppe budgétaire disponible soit égale à 1€ l'année 0 et à 1€ l'année 1. On ne peut donc réaliser que l'un de deux projets l'année 0 et l'autre l'année 1. Comment choisir entre le programme (A,B) et le programme (B,A) ?

Comparons tout d'abord le bénéfico actualisé à l'origine des temps dans l'hypothèse où chacun des deux projets serait réalisé l'année 0. Appelons ce bénéfico a pour le projet A et b pour le projet B. Supposons $a > b$, par exemple $a = 4$ et $b = 2$. Si l'on s'en tenait à cette comparaison, il faudrait donc réaliser le projet A l'année 0 et le projet B l'année 1, c'est-à-dire adopter l'ordre de réalisation (A,B).

Comparons maintenant le bénéfico actualisé à l'origine des temps dans l'hypothèse où chacun des deux projets serait réalisé l'année 1. Appelons ce bénéfico $a + \Delta a$ pour le projet A et $b + \Delta b$ pour le projet B. Le bénéfico actualisé total à l'origine des temps relatif à l'ordre de réalisation (A,B) est donc égal à $(a + b + \Delta b)$.

Or si l'on optait pour l'ordre inverse (B,A), le bénéfico actualisé total à l'origine des temps égal à $(a + b + \Delta a)$.

Le choix de l'ordre de réalisation (A,B) est donc correct si et seulement si $\Delta b \geq \Delta a$ (101)

– On remarque tout d'abord que cette condition ne dépend que des incréments Δa et Δb , et nullement des niveaux a et b des bénéfices actualisés relatifs à la réalisation l'année 0.

– On remarque ensuite que la condition (101) est réalisée notamment dans le cas particulier où $\Delta b > 0$ et $\Delta a < 0$, c'est-à-dire si en reportant la réalisation de l'année 0 à l'année 1, le bénéfico actualisé à l'origine des temps augmente pour le projet B et diminue pour le projet A, ce qui au regard du § 2.2 ci-dessus peut s'interpréter ainsi : ce report rapproche le projet B de sa date optimale de réalisation et au contraire éloignerait le projet A de sa date optimale de réalisation.

– La condition (101) est également remplie dans le cas particulier où $\Delta b \geq \Delta a + (a - b)$ (la parenthèse étant par hypothèse positive), ce qui implique $b + \Delta b \geq a + \Delta a$. Cette condition qui s'interprète comme suit : on établit autant de classements (présélections ou *shortlists*) des projets par ordre décroissant de bénéfice actualisé à l'origine des temps qu'il y a de dates envisageables pour la réalisation : ici une présélection relative à la date hypothétique de réalisation 0 et une autre présélection relative à la date hypothétique de réalisation 1. On peut conclure, sans ambiguïté, si les têtes de présélection sont différentes dans les deux présélections : si A est en tête de la présélection pour l'année 0 et B est en tête de la présélection pour l'année 1, il convient sans ambiguïté de retenir l'ordre de réalisation (A,B).

Mais le choix (A,B) serait incorrect si l'on avait $\Delta a > \Delta b$ (102)

Or comme on a supposé $a > b$, (102) implique $a + \Delta a > b + \Delta b$

Autrement dit : le projet A, qui par hypothèse se classait déjà en tête de la présélection de l'année 0, se classerait aussi dans le cas considéré ici en tête de la présélection pour l'année 1. Il y a alors ambiguïté, le critère de bénéfice actualisé par euro public dépensé est insuffisant pour conclure, il faut le compléter par l'analyse de la condition (101) portant sur les incréments.

Dans les deux tableaux ci-dessous, le projet A se classe en tête des deux présélections relatives aux dates de réalisation respectivement 0 et 1.

Dans le premier tableau, il convient d'adopter l'ordre de réalisation (A,B) et dans le second l'ordre de réalisation (B,A)

Tableau 1

Bénéfice actualisé à l'origine des temps

Date de réalisation	0	1
Projet A	4	5
Projet B	2	4

Remarque : $(\Delta b = 2) > (gDa = 1)$

Tableau 2

Bénéfice actualisé à l'origine des temps

Date de réalisation	0	1
Projet A	4	6
Projet B	2	3

Remarque : $(\Delta a = 2) > (\Delta b = 1)$

On aborde ces questions ci-dessous avec des hypothèses moins simplificatrices que dans l'encadré.

2.2. Notion de dépenses publiques pondérées actualisées à l'origine des temps et de critère auxiliaire pondéré

Les formules (1) ou (2) ci-dessus font intervenir ce que l'on peut appeler le montant des dépenses publiques fictives actualisées à l'origine des temps, où les flux de dépenses (ou recettes) publiques sont affectés des prix fictifs de rareté des fonds publics φ_t :

$$DPFOT_t^n = \varphi_t \cdot \frac{DPI^n}{(1+r)^t} + \sum_{u=t+1}^{\infty} \varphi_u \cdot \frac{DPE_u^n}{(1+r)^u} \quad (9)$$

On peut faire un premier pas dans la direction du critère auxiliaire en divisant par le prix fictif²¹ φ_t , ce qui conduit à la notion de **dépenses publiques pondérées actualisées à l'origine des temps**, définies par la relation suivante :

$$DPPOT_t^n = \frac{DPI^n}{(1+r)^t} + \sum_{u=t+1}^{\infty} \psi(u) \cdot \frac{DPE_u^n}{(1+r)^u} \text{ avec } \psi_u = \frac{\varphi_u}{\varphi_t} \quad (10)$$

Le qualificatif « pondérées » exprime la présence dans (10) des coefficients $\psi(u)$ qui jouent le rôle de poids.

Pour faire un deuxième pas et assimiler la relation (10) à la relation (8), il faudrait donc que les poids $\psi(u)$ soient tous égaux à 1, ce qui supposerait que les prix fictifs de rareté des fonds publics φ_t prennent tous une même valeur φ . On a déjà évoqué ce cas particulier, qui sera approfondi plus loin. On s'en tiendra pour le moment au cas où les poids peuvent être inégaux.

La condition fondamentale (1) ci-dessus peut maintenant s'écrire sous la forme équivalente suivante :

$$BOT_t^n - \varphi_t \cdot DPPOT_t^n = \beta^n - \alpha_t^n \quad (11)$$

Intéressons-nous exclusivement aux projets socio-économiquement rentables, c'est-à-dire appartenant à l'ensemble de définition tels que $BOT_t^n \geq 0$. On va examiner **plusieurs cas selon le signe de $DPPOT_t^n$** .

Cas $DPPOT_t^n \leq 0$. Dans la condition (11), le premier membre est ici positif. Cela implique que β^n soit positif, c'est-à-dire que le projet n soit intégralement réalisé (que ce soit à la date t ou à une autre date). Ainsi, **tout projet (parmi ceux dont le bénéfice actualisé est positif) dont la**

21. Rappel : le multiplicateur φ_t est positif si la contrainte budgétaire de l'année t s'exerce.

dépense publique actualisée pondérée est (au moins certaines années) négative ou nulle doit être intégralement réalisé²².

Cas $DPPOT_t^n > 0$. Ecrivons alors la condition (11) sous la forme équivalente suivante, obtenue en divisant les deux membres par $DPPOT_t^n$.

Cela nous conduit à introduire le *critère auxiliaire pondéré* CAP_t^n , défini comme suit :

$$CAP_t^n = \frac{BOT_t^n}{DPPOT_t^n} \quad (12)$$

On peut alors transcrire la condition d'optimalité (11) sous la forme :

$$CAP_t^n - \varphi_t = \frac{\beta^n}{DPPOT_t^n} - \frac{\alpha_t^n}{DPPOT_t^n} \quad (13)$$

On se propose maintenant de considérer successivement le cas d'un projet qu'il conviendrait de rejeter et celui d'un projet qu'il conviendrait de réaliser intégralement.

2.3. Cas d'un projet qu'il conviendrait de rejeter, c'est-à-dire de ne jamais réaliser

Pour un tel projet, cela implique $\beta^n =$ et $\alpha_t^n > 0$ quelle que soit la date de réalisation t envisagée. Donc il faut et il suffit qu'un tel projet vérifie la condition suivante :

$$CAP_t^n < \varphi_t \quad \text{quelle que soit la date de réalisation } t \text{ envisagée} \quad (14)$$

D'où un premier usage du ratio bénéfice par euro public dépensé (pondéré) : on peut écarter tout projet pour lequel ce ratio serait toujours inférieur au prix fictif de rareté des fonds public. Ce ratio joue donc un premier rôle de tamis²³.

2.4 Cas d'un projet qu'il conviendrait de réaliser intégralement

Pour un tel projet cela implique que $\beta^n > 0$, et que $\alpha_t^n > 0$ quelle que soit l'année de réalisation t envisagée, sauf s'il s'agit de la date optimale de réalisation \hat{t}^n , pour laquelle $\alpha_{\hat{t}^n}^n = 0$.

22. La date de réalisation optimale resterait à préciser. La méthode auxiliaire (contrairement à la méthode générale de la section 2) ne permet pas de départager d'emblée les années candidates. On reviendra plus loin sur cet aspect de l'usage du critère auxiliaire.

23. Mais on n'est guère avancé car on est ici censé ne pas connaître le prix de rareté des fonds publics.

Donc il faut que ce projet vérifie la condition suivante à la date optimale de réalisation :

$$CAP_{\hat{t}^n}^n > \varphi_{\hat{t}^n}^n \quad \text{pour l'année } t = \hat{t}^n \quad (15)$$

En revanche les autres années $t \neq \hat{t}^n$, il y a *ambiguïté*. En effet, on peut avoir :

$$\text{ou bien } CAP_t^n > \varphi_t \quad \text{si } 0 < \alpha_t^n < \beta^n \quad (16.1)$$

$$\text{ou bien } CAP_t^n \leq \varphi_t \quad \text{si } \alpha_t^n \geq \beta^n \quad (16.2)$$

On va maintenant interpréter ces relations.

2.5. *Classer chaque année de réalisation envisageable les projets par ordre décroissant du bénéfice actualisé par euro public dépensé actualisé pondéré CAP_t^n : une présélection nécessaire mais non suffisante*

Une conséquence logique des conditions (15) et (16) est que si une année t donnée l'on classe les projets n par ordre décroissant du critère auxiliaire pondéré CAP_t^n , on obtiendra cette année-là une *présélection (short list)* dont seront éliminés les projets vérifiant la condition (16.2) et où seront retenus les projets vérifiant la condition (15), qui sont ceux que l'on cherche, mais aussi les projets vérifiant (16.1), qui ne sont pas pertinents.

Il en résulte que **si un projet est présélectionné une seule année, il faut réaliser ce projet cette année-là.**

En revanche, **si un projet est présélectionné deux ou plusieurs années, il faut choisir à laquelle de ces dates candidates²⁴ il convient de le réaliser** et pour cela, un **examen complémentaire est nécessaire.**

2.6. *Un critère auxiliaire secondaire pour départager deux dates de réalisation rivales : la variation du critère auxiliaire pondéré CAP_t^n entre ces deux dates*

Soit un projet présélectionné, appartenant à la fois à deux²⁵ présélections relatives à la date t_j et à la date t_k (supposée postérieure à t_j).

À laquelle de ces deux dates convient-il de réaliser ce projet ?

Bien entendu, **la seule méthode sûre pour départager les projets présélectionnés consiste à appliquer le critère de base, c'est-à-dire à**

24. On pourrait le cas échéant réaliser le projet à cheval sur deux ou plusieurs années consécutives, mais il est indispensable de choisir si les présélections sont relatives à des années non consécutives.

25. On s'en tient à deux pour simplifier.

rechercher celle de ces deux dates qui conduit au plus fort bénéfice fictif actualisé à l'origine des temps $BFOT_t^n$ du projet, comme on l'a vu en section 2 ci-dessus.

On va cependant examiner **quel usage l'on peut faire du critère auxiliaire (pondéré) CAP_t^n** , ou plutôt de la **variation de ce ratio entre les deux dates**.

Plus précisément, soit un projet j qu'il convient de réaliser intégralement à la date optimale²⁶ t^j et soit un projet k qu'il convient de réaliser intégralement à la date optimale t^k , postérieure à t^j .

On peut (voir **annexe 3**) calculer pour le projet j le **critère auxiliaire pondéré $CAP_{t^n}^j$** à la date t^j , puis à la date t^k , et enfin la **variation $\Delta_{t^k}^{t^j} CAP^j$ de ce ratio** entre la date t^j et la date t^k . On peut calculer la variation correspondante $\Delta_{t^k}^{t^j} CAP^k$ pour le projet k .

Avec un certain nombre d'approximations (explicitées en **annexe 3**), on en déduit que, s'il convient de réaliser le projet j intégralement à la date optimale t_j et le projet k intégralement à la date t_k , alors les variations vérifient l'inégalité :

$$\Delta_{t^k}^{t^j} CAP^k > \Delta_{t^k}^{t^j} CAP^j \quad (17)$$

Interprétation : imaginons que la présélection des projets à réaliser à la date t^j se compose des deux projets j et k et que la présélection des projets à réaliser à la date t^k se compose des deux mêmes projets. Partons de la situation imaginaire où l'on envisagerait de réaliser chacun de ces deux projets à la date t^j et supposons que l'on se pose ensuite la question : s'il faut retarder l'un des projets à la date t^k , lequel faut-il retarder ? La réponse est : **il convient de retarder celui des deux projets pour lequel le ratio « bénéfice actualisé par euro public (pondéré) » connaît la variation la plus forte en passant de la date t^j à la date t^k .**

Remarques

– Il est évident que la condition (17) est remplie si le premier membre est positif et que le second membre est négatif, c'est-à-dire si, lorsque l'on passe de la date t_j à la date postérieure t_k , le ratio bénéfice par euro public (pondéré) à l'origine des temps CAP^k du projet k augmente tandis que celui CAP^j du projet j diminue.

– Il est aisé de montrer que la condition (17) est remplie si le ratio du projet j est plus élevé que celui du projet k à la date t_j et que, à l'inverse, le ratio du projet k est plus élevé que celui du projet j à la date t_k .

26. Pour alléger l'écriture, on utilise ici les notations t^j et t^k au lieu des notations t^j et t^k .

2.7. Simplifications complémentaires dans le calcul des critères auxiliaires

En toute rigueur, les formules ci-dessus font intervenir **la dépense publique pondérée actualisée à l'origine des temps** $DPOT_t^n$ qui – comme on l'a déjà signalé plus haut – nécessite de connaître la chronique des « poids » $\psi(u)$, donc en fait celle des prix fictifs de rareté des fonds publics $\varphi(u)$... dont les méthodes auxiliaires cherchent précisément à s'affranchir.

2.7.1. Approximation du critère du bénéfice par euro public (simple) actualisé à l'origine des temps

On est alors conduit à envisager l'approximation consistant à substituer à la dépense publique pondérée la dépense publique actualisée simple, telle que définie plus haut :

$$DPOT_t^n = \frac{DPI^n}{(1+r)^t} + \sum_{u=t+1}^{\infty} \frac{DPE_u^n}{(1+r)^u} \quad (8)$$

et le critère auxiliaire simple correspondant : $CA_t^n = \frac{BOT_t^n}{DPOT_t^n}$ (7)

Cette approximation revient à assimiler dans (10) les **poids** ψ_t à **1** quelle que soit l'année t .

Elle peut recouvrir deux types d'erreurs, plus ou moins discutables.

a) En toute rigueur, admettre que $\psi_t = 1$ quel que soit t implique que le prix fictif de rareté des fonds publics φ_t prend une constante φ indépendante de l'année t de réalisation envisagée. Mais φ_t ne saurait être constant dans le cas général, si l'on considère des chroniques de disponibilités budgétaires qui peuvent a priori être quelconques. On reviendra toutefois sur ce point en section 4 ci-après.

b) Si néanmoins on « fait comme si » l'on avait $\psi_t = 1$ quel que soit t tout en « faisant comme si » le prix fictif de rareté des fonds publics φ_t pouvait varier au cours du temps, on introduit des inconsistances temporelles, peu satisfaisantes mais assimilables – peut-être – à une approximation acceptable.

2.7.2 Préconisations

Avec ces approximations permettant d'utiliser le critère auxiliaire CA_t^n de bénéfice par euro public (simple) actualisé à l'origine des temps, les préconisations ci-dessus deviennent :

- Chaque année de réalisation envisagée, classer les projets par ordre décroissant du critère CA_t^n
- Eliminer les projets qui ont un ratio trop faible.

c) Examiner si un projet donné est présélectionné deux années non consécutives et en pareil cas, choisir entre les années rivales, par exemple en effectuant avec d'autres projets les permutations au vu de la variation du critère.

d) En toute rigueur, remonter au niveau central pour arrêter la sélection définitive, car il peut exister des projets rivaux, présélectionnés par deux maîtres d'ouvrages décentralisés, entre lesquels il faut arbitrer.

3. REGARD SUR LE PARTENARIAT PUBLIC PRIVÉ

Revenons maintenant sur le cas du partenariat public privé qui a été seulement évoqué ci-dessus.

Le cas type est celui où l'investissement du projet n est totalement préfinancé par le partenaire privé et où la puissance publique rembourse, avec un taux d'intérêt²⁷ i_n , selon une chronique de versements (intérêt et principal) convenue par avance. Cette solution en partenariat peut d'ailleurs être considérée comme une variante parmi d'autres envisageables pour réaliser le projet considéré.

Il est à noter que le taux d'intérêt i_n est propre au projet n , car il dépend des risques spécifiques qui lui sont attachés. On reviendra sur ce point un peu plus loin (§ 4.2).

3.1. Incidence sur les formules de programmation optimale

Soit $IPPP^n$ le coût de l'investissement dans cette variante de partenariat public privé, couramment supposé inférieur au coût des autres variantes en raison de l'efficacité plus élevée dont sont censés faire preuve les partenaires privés. Supposons cet investissement réalisé à la date t inconnue.

Si le projet est réalisé à la date t on suppose que la puissance publique rembourse sa dette au partenaire privé en lui versant l'année u , postérieure à t , un montant²⁸ $V_{t,u}^n$, sur une certaine période, que pour simplifier on suppose infinie.

Considérons pour le moment cette chronique comme exogène.

$$\text{Il est à noter qu'elle vérifie l'identité : } IPPP^n = \sum_{u=t+1}^{\infty} \frac{V_{t,u}^n}{(1+i_n)^{u-t}} \quad (18)$$

27. Il s'agit du taux d'intérêt réel, puisque tous les prix sont constants (ce que l'on a résumé par « euros constants »).

28. Toujours en euros constants.

On observe que dans ce cas le bénéfice fictif actualisé à l'origine des temps prend la forme²⁹ :

$$BFOT_t^n = BOT_t^n - \sum_{u=t+1}^{\infty} \varphi_u \cdot \frac{DPE_u^n + V_{t,u}^n}{(1+r)^u} \quad \text{pour ce projet } n \text{ et pour tout } t \quad (19)$$

Une solution de remboursement formellement simple est donnée pour toute année u postérieure à t par $V_{t,u}^n = i_n \cdot IPPP^n$ (u variant ici de 1 à $+\infty$) (20)

De sorte que la relation (19) prend la forme simple suivante :

$$BFOT_t^n = BOT_t^n - \sum_{u=t+1}^{\infty} \varphi_u \cdot \frac{DPE_u^n + i_n \cdot IPPP^n}{(1+r)^u} \quad \text{pour ce projet } n \text{ et pour tout } t \quad (21)$$

Remarque

En réalité, la chronique des versements $V_{t,u}^n$ peut elle-même faire l'objet d'une procédure d'optimisation. Pour ne pas alourdir la présentation, on n'examinera pas ici cette question. On s'en tiendra au cas plus simple abordé ci-après.

3.2. Chronique optimale des préfinancements privés

On envisage maintenant le **cas purement fictif** où les pouvoirs publics pourraient faire intervenir des partenaires privés sur tous les projets avec **un seul et même taux d'intérêt**, noté i .

Le partenariat public privé pourrait alors être utilisé pour **déconnecter** la chronique annuelle Y_t des financements de projet et la chronique annuelle C_t des rationnements de fonds publics, pourvu que la puissance publique puisse faire face à la charge de sa dette (intérêts et principal) en versant **exactement** le montant annuel C_t . Cette condition est remplie si et seulement si ces deux chroniques sont équivalentes lorsqu'on les escompte avec le taux consenti par les partenaires financiers, ce qui se traduit par l'identité :

$$\sum_{t=0}^{\infty} \frac{C_t}{(1+i)^t} - \sum_{t=0}^{\infty} \frac{Y_t}{(1+i)^t} = 0 \quad (22)$$

29. Il convient de préciser que le coût fictif de rareté des fonds publics utilisé ici est bien le même que dans les sections 2 et 3 ci-dessus ; il n'y a pas de valeur particulière de ce coût fictif pour les projets PPP. La valeur de ce coût fictif est en effet censée découler (dans son principe, voir § 2.4) du processus d'optimisation par programmation linéaire à portée générale qui a déterminé l'optimum en prenant en compte tous les projets et pour chacun d'eux toutes les variantes en lice, y compris les variantes PPP.

Le préfinancement privé, s'il était réalisable, permettrait de substituer cette contrainte intertemporelle unique (22) à la suite (infinie) des contraintes budgétaires annuelles. Une telle simplification permet de dégager d'importants degrés de liberté, donc d'atteindre un optimum sensiblement plus avantageux.

On peut alors montrer (voir **annexe 4**) que la réalisation de cet optimum intertemporel implique que le prix fictif de rareté des fonds public soit de la forme :

$$\frac{\varphi_t}{(1+r)^t} = \frac{\lambda}{(1+i)^t} \quad (23.1) \quad \text{équivalente à} \quad \varphi_t = \lambda \cdot \left(\frac{1+r}{1+i}\right)^t \quad (23.2)$$

Si la relation (23.2) était vérifiée, il en découlerait, pour le calcul de la dépense publique pondérée de la formule (10), que le poids ψ_u serait donné par :

$$\psi_u = \frac{\varphi_u}{\varphi_t} = \left(\frac{1+r}{1+i}\right)^{u-t} \quad \text{avec } u > t : \quad (24)$$

Remarque : Dans cette formule, la variable duale λ est fonction de la chronique C_t

On peut alors examiner ce que deviennent les résultats des sections précédentes dans l'hypothèse où l'on pourrait appliquer ces relations (23) et (24). On va procéder à cet examen d'abord dans le cas particulier où le taux d'intérêt du privé et le taux d'actualisation seraient égaux, puis dans le cas plus général où ces deux taux seraient distincts.

3.2.1. Cas où l'on aurait $r = i$

Quelle que soit l'année considérée, le prix fictif de rareté des fonds public φ_t **prendrait alors une seule et même valeur** (égale au scalaire³⁰ λ) et le poids ψ_u **serait constamment égal à 1**. On se trouverait donc dans le **cas très simplifié** déjà évoqué plusieurs fois plus haut.

3.2.2. Cas où l'on aurait $i > r$

Cependant la situation précédente dans laquelle il y aurait égalité entre le taux d'actualisation applicable dans le public et le taux d'intérêt réel requis par le privé apparaît assez peu probable. Il semble a priori plus vraisemblable que les prêteurs privés demandent un taux d'intérêt réel supérieur au taux d'actualisation public, qu'ils aient une

30. Le calcul complet de la maximisation du Lagrangien permettait de déterminer la valeur de ce scalaire.

aversion au risque plus forte ou une surface financière moins large que la puissance publique, ou pour toute autre raison.

Quinet et Sauvart dans leur contributions indiquent à cet égard que l'écart $(i - r)$ serait de l'ordre de 1 à 2 points de pourcentage.

a) Avec cette hypothèse $i > r$, la relation (23.2) montre que, à l'optimum intertemporel, les prix fictifs de rareté des fonds publics $\varphi(t)$ formeraient une **suite géométrique décroissante**, et selon la relation (20), il en irait de même pour les poids $\psi(u)$. Ces suites auraient pour « raison » $\frac{1+r}{1+i}$

b) Mais on peut franchir un pas de plus en considérant cette fois l'expression (23.1). À l'optimum intertemporel, on peut en effet remplacer $\frac{\varphi_t}{(1+r)^t}$ par l'expression qui lui est égale $\frac{\lambda}{(1+i)^t}$.

Dès lors, la dépense publique fictive actualisée à l'origine des temps définie par (9) peut se transcrire sous la forme :

$$DPFOT_t^n = \lambda \cdot \left(\frac{DPI^n}{(1+i)^t} + \sum_{u=t+1}^{\infty} \frac{DPE_u^n}{(1+i)^u} \right) \quad (9)$$

Dans cette relation, la parenthèse s'analyse comme la « **dépense publique escomptée à l'origine des temps** » définie comme suit :

$$DPEOT_t^n = \frac{DPI^n}{(1+i)^t} + \sum_{u=t+1}^{\infty} \frac{DPE_u^n}{(1+i)^u} \quad (25)$$

On a ainsi remplacé le terme « actualisée » par le terme « **escomptée** » qui fait implicitement référence au taux d'escompte privé, égal à $1/(1+i)$

• **Simplification du critère de base : bénéfice fictif actualisé à l'origine des temps**

Le critère fondamental du bénéfice fictif actualisé à l'origine des temps se simplifie lui-même et devient, à l'optimum intertemporel :

$$BFOT_t^n = BOT_t^n - \lambda \cdot DPEOT_t^n \quad (26)$$

Cette expression est remarquable car **elle ne fait plus intervenir que la seule variable duale** λ , quel que soit le projet n et quelle que soit la date de réalisation t envisagée.

À noter que le taux d'actualisation reste sous-jacent dans l'expression (26) ; c'est lui (et non le taux du privé i) qui est utilisé pour calculer le bénéfice actualisé à l'origine des temps.

La décentralisation de l'optimum serait dans ce cas particulièrement simple : les autorités centrales auraient à diffuser, en plus du taux d'actualisation r , d'une part le taux d'intérêt réel i relatif aux préfinancements privés, d'autre part la variable duale unique λ . Chaque maître d'ouvrage décentralisé n'aurait plus qu'à calculer à l'aide de la relation (26) le bénéfice fictif actualisé à l'origine des temps pour chacun de ses projets et choisir variante optimale et date optimale de réalisation en cherchant le maximum (positif) correspondant.

• **Simplification du critère auxiliaire : bénéfice par euro public**

De la forme simplifiée (26) on peut tirer un critère auxiliaire lui-même plus simple à calculer, à savoir le bénéfice par euro public escompté ainsi défini :

$$CAPE_t^n = \frac{BOT_t^n}{DPEOT_t^n} \quad (27)$$

Dans cette relation, le dénominateur est la dépense publique escomptée à l'origine des temps définie par (25) et pour simplifier on supposera ici qu'il est positif.

On peut alors reprendre le raisonnement développé à la section 3 ci-dessus en remplaçant la relation (13) rappelée ci-après par :

$$CAPE_t^n - \lambda = \frac{\beta^n}{DPEOT_t^n} - \frac{\alpha_t^n}{DPEOT_t^n} \quad (28)$$

On peut dès lors chaque année t :

- rejeter tout projet tel que $CAPE_t^n < \lambda$
- constituer avec les projets tels que $CAPE_t^n \geq \lambda$ une présélection
- rechercher les projets présélectionnés deux (ou plusieurs) années non consécutives et étudier leur cas de plus près (par exemple selon le critère secondaire de permutation).

La mise en œuvre est grandement simplifiée par le fait que la variable duale λ est ici la même chaque année (alors qu'à la section 3 il existe en toute rigueur une variable duale φ_t annuelle).

Mais ce critère auxiliaire reste ambigu et ne se suffit toujours pas à lui-même. Il permet de rejeter un projet à l'échelle décentralisée, mais pas de le retenir à l'échelle décentralisée – car il peut être rival d'un autre projet relevant d'un autre maître d'ouvrage : il faut donc théoriquement remonter au niveau central pour effectuer la sélection définitive.

4. CONCLUSION

En cas de rationnement annuel des fonds publics, il existe en théorie chaque année un **prix fictif de rareté des fonds publics** φ_t .

4.1. Méthode de base : maximiser le bénéfice fictif actualisé à l'origine des temps *BFOT*

La démarche éprouvée du calcul économique avec usage de prix fictifs s'offre dès lors comme une solution présentant le double avantage d'être aisée à appliquer et entièrement décentralisable, si du moins les pouvoirs publics centraux estimaient possible et souhaitable de déterminer la chronique de φ_t et de la diffuser auprès des maîtres d'ouvrages décentralisés.

À l'aide de la formule suivante donnant le « bénéfice fictif actualisé à l'origine des temps »

$$BFOT_t^n = BOT_t^n - \varphi_t \cdot \frac{DPI^n}{(1+r)^t} - \sum_{u=t+1}^{\infty} \varphi_u \cdot \frac{DPE_u^n}{(1+r)^u} \quad (2)$$

chaque maître décentralisé pourrait alors calculer pour chaque projet autant de valeurs du *BFOT* qu'il y a de variantes (notamment variantes technologiques, mais aussi variantes de financement) et qu'il y a d'années de réalisation envisageables.

Il pourrait ensuite pour chacun desdits projets, en maximisant *BFOT*, déterminer la variante optimale et la date de réalisation optimale (qui peut ne pas exister, si le *BFOT* est constamment négatif, le projet étant alors à rejeter).

S'ils décidaient d'adopter cette démarche, les pouvoirs publics centraux pourraient, à l'instar de ce qui a été fait pour déterminer le taux d'actualisation, charger un groupe de travail de personnalités qualifiées d'élaborer des propositions relatives à la chronique du prix fictif de rareté des fonds publics φ_t .

4.2. Critère auxiliaire du bénéfice par euro public CA_t^n

Si la méthode de base ci-dessus ne pouvait pas être suivie, que peut-on attendre du recours au critère auxiliaire constitué par le ratio « bénéfice socio-économique par euro public » que l'instruction cadre demande de calculer ?

Ce critère peut rendre service, mais il comporte certains risques d'erreur.

– Ainsi, il peut être utile pour jouer le rôle de filtre, en éliminant les projets présentant un ratio trop faible (encore que la limite soit floue).

– Il peut aussi servir à classer les projets par ordre décroissant du ratio, ce qui donne une présomption d'ordre de préférence, mais ce classement ne peut se suffire à lui-même, car il convient d'établir un classement pour chaque année de réalisation envisageable et un même projet peut être bien classé deux (ou plusieurs) années non consécutives : il faut alors une procédure complémentaire pour choisir entre les années rivales.

Une conséquence de cette ambiguïté résiduelle est que le critère ne peut pas être décentralisé en toute sécurité, car deux projets dépendant de deux maîtres d'ouvrage décentralisés différents peuvent être rivaux et l'arbitrage ne peut alors être élucidé qu'au niveau central.

Ce critère auxiliaire n'est donc qu'un substitut aisé mais imparfait à la méthode de base.

4.3. *Quelques éléments relatifs au partenariat public privé (PPP)*

L'examen des variantes de financement de projet, et notamment le cas échéant d'un recours au PPP, est justiciable des démarches qui viennent d'être présentées, au même titre que toute autre variante.

À titre purement exploratoire, on peut cependant examiner le cas entièrement fictif où les pouvoirs publics, connaissant la chronique des enveloppes budgétaires annuelle C_t , pourraient obtenir du secteur privé une chronique d'enveloppes annuelles de préfinancement Y_t , donnant lieu à un taux d'intérêt i , de telle sorte que la charge de remboursement correspondante (intérêt et principal) coïncide exactement avec les enveloppes budgétaires annuelles C_t .

Cette approche aurait l'avantage de substituer aux contraintes budgétaires annuelles une contrainte financière intertemporelle unique :

$$\sum_{t=0}^{\infty} \frac{C_t}{(1+i)^t} - \sum_{t=0}^{\infty} \frac{Y_t}{(1+i)^t} = 0 \quad (22)$$

Elle ouvrirait la possibilité d'une meilleure allocation intertemporelle des ressources financières.

En outre, les calculs mentionnés ci-dessus seraient simplifiés :

– une variable duale λ unique se substituerait à la chronique des prix fictifs de rareté des fonds publics φ_t ;

– les euros publics seraient à escompter à l'origine des temps à l'aide du taux d'intérêt i , au lieu du taux d'actualisation r .

La méthode de base conserverait toute sa pertinence, mais la méthode auxiliaire serait plus aisée à pratiquer (sans pour autant être affranchie de ses autres inconvénients).

Pour autant, cette modélisation reste très abstraite, pour de multiples raisons : le taux d'intérêt des PPP dépend des risques de chaque projet ; les incertitudes sur la chronique C_t compromettent sans doute radicalement la possibilité de préfinancements assimilables à la relation (22).

Mais cette idéalisation sous-jacente pourrait éventuellement faciliter les travaux du groupe d'expert précité et conduire à une estimation non déraisonnable d'un coût fictif de rareté des fonds publics constant³¹ qui simplifierait grandement la mise en œuvre du calcul économique avec contrainte budgétaire.

4.4. Voies d'approfondissement

Précisément, la question de l'incertitude et du risque, que Quinet et Sauvart abordent dans leur contribution, est absente du présent document. Elle constitue une voie d'approfondissement pour un éventuel prolongement du présent essai de récapitulation.

5. ÉPILOGUE

Dans tout ce qui précède, on a eu recours à la fiction des projets fractionnables.

Cette approche permet en particulier chaque année t de déterminer un « **projet partiel p** », réalisé partiellement cette année là. On peut

montrer que le ratio $\frac{BOT_t^p}{DPOT_t^p}$ fournit³² la valeur du prix fictif φ_t .

Mais cette fiction n'est pas pleinement satisfaisante, car en toute rigueur, il faudrait raisonner en **nombre entiers** : un projet est à réaliser intégralement ou pas du tout. Cela conduirait à déterminer chaque année t toutes les combinaisons des projets finançables sur fonds publics dans la limite du rationnement $C(t)$. Il faudrait alors parmi toutes ces combinaisons déterminer celle qui maximise la valeur nette actualisée combinée. Il faudrait ensuite, si un projet est présélectionné plusieurs années, tâtonner pour l'affecter à une année précise. C'est d'ailleurs une approche mise en œuvre dans la contribution de Bonnafoy et Roy.

31. Mais révisable à intervalles réguliers, comme préconisé pour le taux d'actualisation par le rapport Lebègue.

32. Il faut distinguer en toute rigueur le cas où le projet partiel finit ou non par être complété à 100 % au cours d'un ou plusieurs autres années.

On peut sur des cas simples (voir **annexe 5**) mettre en évidence les erreurs de sélection que peut induire la fiction du fractionnement. Il nous semble néanmoins que cette méthode a le grand mérite d'être relativement simple et maniable, au prix d'erreurs acceptables.

Annexe 1. Application du théorème de Kuhn et Tucker à l'optimisation par programmation linéaire

Remarque préliminaire : les notations utilisées sont celles définies dans l'introduction générale. Elles visent à être aussi proche que possible des notations de l'instruction-cadre des 25 mars 2004 et 27 mai 2005.

On suppose que l'année générique t on réalise et achève une fraction du projet n égale à x_t^n , comprise au sens large entre 0 et 1.

On se propose de rechercher les valeurs des fractions inconnues x_t^n en vue de maximiser la somme des bénéfices des projets, actualisés à l'origine des temps, sous trois types de contraintes :

$$x_t^n \geq 0 \quad \text{pour tout projet } n \text{ et toute date d'achèvement } t$$

$$1 - \sum_{t=0}^{\infty} x_t^n \geq 0 \quad \text{pour tout projet } n$$

$$C_t - \sum_{n=1}^N \left[x_t^n \cdot DPI^n + \left(\sum_{u=0}^{t-1} x_u^n \right) \cdot DPE_t^n \right] \geq 0 \quad \text{pour tout } t$$

Cette formulation conduit à considérer le Lagrangien suivant :

$$\Lambda = \sum_{n=1}^N \sum_{t=0}^{\infty} [x_t^n \cdot BOT_t^n + \alpha_t^n \cdot x_t^n] + \sum_{n=1}^N \beta^n \cdot \left(1 - \sum_{t=0}^{\infty} x_t^n \right) + \dots \dots \dots \sum_{t=0}^{\infty} \frac{\varphi_t}{(1+r)^t} \cdot \left(C_t - \sum_{n=1}^N \left[x_t^n \cdot DPI^n + \left(\sum_{u=0}^{t-1} x_u^n \right) \cdot DPE_t^n \right] \right)$$

qui fait intervenir les multiplicateurs de Lagrange (ou variables duales) suivants :

α_t^n qui est : nul si la totalité ou une fraction (non nulle) du projet n est à réaliser l'année t ; positif³³ si aucune fraction du projet n n'est à réaliser l'année t

β^n qui est nul : si le projet n n'est pas à réaliser en entier (il en manquera toujours au moins une fraction, voire la totalité) ; positif³⁴ si

33. Ou éventuellement nul, mais cet événement peut être considéré comme étant de probabilité nulle.

34. Ou éventuellement nul, mais cet événement peut être considéré comme étant de probabilité nulle.

le projet est à réaliser en entier (fût-ce par fractions sur plusieurs années³⁵).

φ_t qui est³⁶ : nul si les fonds publics disponibles l'année t excèdent cette année-là les dépenses publiques nettes des recettes publiques, afférentes aux projets programmés ; positif³⁷ si les fonds publics rationnent effectivement les dépenses publiques nettes considérées. Ce multiplicateur φ_t s'analyse comme un **prix fictif de rareté³⁸ des fonds publics** à la date t .

Nota : Les « prix fictif de rareté des fonds publics » φ_t sont **distincts et indépendants** du « coût d'opportunité des fonds publics » fixé à 0,3 par l'instruction cadre, qui reflète des pertes d'efficacité économique (distorsions) censée résulter du recours à l'impôt . Ce coût d'opportunité intervient exclusivement dans le calcul du bénéfice actualisé BOT_t^n et n'apparaît donc pas dans la problématique du rationnement budgétaire qui fait l'objet du présent examen.

La maximisation du Lagrangien, ou du moins les conditions du premier ordre de cette maximisation (annulation des dérivées partielles par rapport aux fractions inconnues x_t^n), conduisent alors aux résultats suivants :

$$BOT_t^n - \varphi_t \cdot \frac{DPI_t^n}{(1+r)^t} - \sum_{u=t+1}^{\infty} \varphi_u \cdot \frac{DPE_u^n}{(1+r)^u} = \beta^n - \alpha_t^n \quad \text{pour tout } n \text{ et tout } t$$

Ce système, complété³⁹ par les équations de rationnement annuel des fonds publics, détermine simultanément les fractions inconnues x_t^n et les prix fictifs de rareté des fonds publics φ_t .

35. Dans ce cas, si ces années n'étaient pas consécutives, la solution mathématique serait dépourvue de signification socio-économique.

36. Le multiplicateur est plus précisément $\varphi_t/(1+r)^t$: la présence du dénominateur $(1+r)^t$ est une commodité qui permet de simplifier les calculs ultérieurs.

37. Ou éventuellement nul, mais cet événement peut être considéré comme étant de probabilité nulle.

38. Autre appellation : variable duale associée à la contrainte budgétaire de l'année t .

39. Plus exactement : le système complet s'obtient en annulant les dérivées partielles du Lagrangien y compris par rapport à tous les multiplicateurs de Lagrange.

Annexe 2. Choix entre deux projets incompatibles

On reprend les notations de l'annexe 1.

Supposons que les projets n° 1 et n° 2 soient incompatibles entre eux et qu'il existe d'autres projets n tous compatibles entre eux ainsi qu'avec le projet. n° 1 et le projet n° 2.

La formule du Lagrangien de l'annexe 1 devient :

$$\Lambda = \sum_{t=0}^{\infty} [x_t^1 \cdot BOT_t^1 + \alpha_t^1 \cdot x_t^1 + x_t^2 \cdot BOT_t^2 + \alpha_t^2 \cdot x_t^2] + \beta \cdot \left[1 - \sum_{t=0}^{\infty} (x_t^1 + x_t^2) \right] \\ + \sum_{n=3}^N \sum_{t=0}^{\infty} [x_t^n \cdot BOT_t^n + \alpha_t^n \cdot x_t^n] + \sum_{n=3}^N \beta^n \cdot \left(1 - \sum_{t=0}^{\infty} x_t^n \right) + \dots \\ \dots \sum_{t=0}^{\infty} \frac{\varphi(t)}{(1+r)^t} \cdot \begin{pmatrix} C(t) & \dots & \dots \\ - \left(x_t^1 \cdot DPI^1 + \left(\sum_{u=0}^{t-1} x_u^1 \right) \cdot DPE_t^1 \right) \\ - \left(x_t^2 \cdot DPI^2 + \left(\sum_{u=0}^{t-1} x_u^2 \right) \cdot DPE_t^2 \right) \\ - \sum_{n=3}^N \left(x_t^n \cdot DPI^n + \left(\sum_{u=0}^{t-1} x_u^n \right) \cdot DPE_t^n \right) \end{pmatrix}$$

La grande différence est que le multiplicateur β de la première ligne s'applique simultanément aux (deux) variantes incompatibles.

Dérivons le Lagrangien successivement par rapport à x_t^1 et x_t^2 . On obtient :

$$BOT_t^1 - \varphi(t) \cdot \frac{DPI^1}{(1+r)^t} - \sum_{u=t+1}^{\infty} \varphi(u) \cdot \frac{DPE_u^1}{(1+r)^u} = \beta - \alpha_t^1 \quad \text{pour tout } t \\ BOT_t^2 - \varphi(t) \cdot \frac{DPI^2}{(1+r)^t} - \sum_{u=t+1}^{\infty} \varphi(u) \cdot \frac{DPE_u^2}{(1+r)^u} = \beta - \alpha_t^2 \quad \text{pour tout } t$$

S'il faut réaliser intégralement la variante 1 (par exemple) une année \hat{t}^1 , on aura cette année là $\alpha_{\hat{t}^1}^1$ et pour toutes les autres années $t \neq \hat{t}^1$ on aura : $\alpha_t^1 > 0$.

La variante 2 étant censée n'être jamais à réaliser, quel que soit t (y compris pour $t = \hat{t}^1$) on aura : $\alpha_t^2 > 0$

Il en résulte une procédure simple :

- on cherche le maximum par rapport à t de $BFOT_t^1$
- on cherche le maximum par rapport à t de $BFOT_t^2$
- la plus grande de ces deux valeurs désigne simultanément la variante qu'il convient de retenir et l'année où il convient de la réaliser.

Annexe 3. Variation du critère auxiliaire pondéré CAP_t^n « bénéfique par euro public dépensé pondéré, actualisé à l'origine des temps » lorsque l'on passe de la date t_j à la date t_k

Rappels. Pour un projet n que l'on envisage de réaliser à la date t , on définit les notions suivantes :

– dépense publique pondérée actualisée à l'origine des temps

$$DPPOT_t^n = \frac{DPI^n}{(1+r)^t} + \sum_{u=t+1}^{\infty} \psi(u) \cdot \frac{DPE_u^n}{(1+r)^u} \quad \text{avec } \psi_u = \frac{\varphi_u}{\varphi_t} \quad (10)$$

Nota : on suppose ici $DPPOT_t^n > 0$. Voir discussion plus haut dans le texte.

– critère auxiliaire pondéré : bénéfique actualisé à l'origine des temps par euro public pondéré :

$$CAP_t^n = \frac{BOT_t^n}{DPPOT_t^n} \quad (11)$$

– condition d'optimalité du premier ordre issue de la maximisation sous contrainte budgétaire annuelle⁴⁰ :

$$CAP_t^n = \varphi_t + \frac{\beta^n}{DPPOT_t^n} - \frac{\alpha_t^n}{DPPOT_t^n} \quad (12)$$

On suppose ici que le projet j doit être réalisé à la date optimale \hat{t}^j , que l'on notera t^j pour alléger l'écriture.

On a alors (propriété de la date de réalisation optimale) :

$$CAP_{t^j}^k = \varphi_{t^j} + \frac{\beta^k}{DPPOT_{t^j}^k}$$

On suppose de même que le projet k doit être réalisé à la date optimale \hat{t}^k , que l'on notera t^k pour alléger l'écriture. On supposera sans perte de généralité $t^k > t^j$.

On a alors (propriété de la date de réalisation optimale) :

$$CAP_{t^k}^k = \varphi_{t^k} + \frac{\beta^k}{DPPOT_{t^k}^k}$$

40. On suppose ici $DPPOT_t^n > 0$. Voir discussion plus haut dans le texte.

En revanche, le critère auxiliaire pondéré du projet j calculé à la date t^k , qui est supposée postérieure à sa date de réalisation optimale t^j , est tel que :

$$CAP_{t^k}^j = \varphi_{t^k} + \frac{\beta^j}{DPPOT_{t^k}^j} - \frac{\alpha_{t^k}^j}{DPPOT_{t^k}^j}$$

De même, le critère auxiliaire pondéré du projet k calculé à la date t^j , qui est supposée antérieure à sa date de réalisation optimale t^k , est tel que :

$$CAP_{t^j}^k = \varphi_{t^j} + \frac{\beta^k}{DPPOT_{t^j}^k} - \frac{\alpha_{t^j}^k}{DPPOT_{t^j}^k}$$

Calculons la **variation du critère auxiliaire** (pondéré) relative au projet j lorsque l'on passe de la date t^j à la date t^k , en posant :

$$\Delta_{t^j}^k CAP^j = CAP_{t^k}^j - CAP_{t^j}^j$$

Donc :

$$\Delta_{t^j}^k CAP^j = \varphi_{t^k} - \varphi_{t^j} + \beta^j \cdot \left(\frac{1}{DPPOT_{t^k}^j} - \frac{1}{DPPOT_{t^j}^j} \right) - \frac{\alpha_{t^k}^j}{DPPOT_{t^k}^j}$$

Calculons de même la variation du critère auxiliaire (pondéré) relative au projet k lorsque l'on passe de la date t^j à la date t^k ,

$$\Delta_{t^j}^k CAP^k = \varphi_{t^k} - \varphi_{t^j} + \beta^k \cdot \left(\frac{1}{DPPOT_{t^k}^k} - \frac{1}{DPPOT_{t^j}^k} \right) + \frac{\alpha_{t^j}^k}{DPPOT_{t^j}^k}$$

Écart entre la variation du projet k et la variation du projet j est alors :

$$\begin{aligned} \Delta_{t^j}^k CAP^k - \Delta_{t^j}^k CAP^j &= \frac{\alpha_{t^k}^j}{DPPOT_{t^k}^j} + \frac{\alpha_{t^j}^k}{DPPOT_{t^j}^k} \dots\dots \\ &\dots\dots + \beta^k \cdot \left(\frac{1}{DPPOT_{t^k}^k} - \frac{1}{DPPOT_{t^j}^k} \right) - \beta^j \cdot \left(\frac{1}{DPPOT_{t^k}^j} - \frac{1}{DPPOT_{t^j}^j} \right) \end{aligned}$$

On observe qu'au second membre les deux termes en α s'ajoutent, tandis que les termes en β sont multipliés par des parenthèses constituées par des différences. On peut alors avoir l'intuition que ces parenthèses sont **négligeables**.

Si on adopte cette **simplification**, la relation ci-dessus se réduit à :

$$\Delta_{t^j}^k CAP^k - \Delta_{t^j}^k CAP^j = \frac{\alpha_{t^k}^j}{DPPOT_{t^k}^j} + \frac{\alpha_{t^j}^k}{DPPOT_{t^j}^k}$$

Les multiplicateurs $\alpha_{t^k}^j$ et $\alpha_{t^j}^k$ étant l'un et l'autre positifs, la relation ci-dessus implique alors : $\Delta_{t^j}^k CAP^k > \Delta_{t^j}^k CAP^j$

Interprétation : imaginons que la présélection des projets à réaliser à la date t^j se compose des deux projets j et k et que la présélection des projets à réaliser à la date t^k se compose des deux mêmes projets. Partons de la situation imaginaire où l'on envisagerait de réaliser chacun de ces deux projets à la date t^j et supposons que l'on se pose ensuite la question : s'il faut retarder l'un des projets à la date t^k , lequel faut-il retarder ? La réponse est : **il convient de retarder celui des deux projets pour lequel le critère « bénéfice actualisé par euro public pondéré » connaît la variation (algébrique) la plus forte en passant de la date t^j à la date t^k .**

Annexe 4. Partenariat public privé

Chroniques de préfinancement

On suppose ici que les pouvoirs publics connaissent exactement la chronique des enveloppes budgétaires annuelles C_t .

On suppose en outre qu'elles peuvent demander à des partenaires privés de préfinancer les dépenses publiques annuelles à concurrence d'enveloppes annuelles de fonds privés Y_t .

On suppose que ces deux chroniques sont équivalentes sur le plan bancaire avec le taux d'intérêt i , ce qui revient à dire que les deux chroniques de flux actualisés à l'origine des temps à l'aide du d'intérêt i sont égales :

$$\sum_{t=0}^{\infty} \frac{C_t}{(1+i)^t} - \sum_{t=0}^{\infty} \frac{Y_t}{(1+i)^t} = 0 \quad (22)$$

Il existe une infinité de chroniques Y_t possibles et il s'agit de chercher la plus avantageuse, c'est-à-dire celle qui permet de maximiser la

somme de tous les bénéfices actualisés à l'origine des temps. Le Lagrangien de l'annexe prend ici la forme suivante :

$$\Lambda = \sum_{n=1}^N \sum_{t=0}^{\infty} [x_t^n \cdot BOT_t^n + \alpha_t^n \cdot x_t^n] + \sum_{n=1}^N \beta^n \cdot \left(1 - \sum_{t=0}^{\infty} x_t^n\right) + \dots \dots$$

$$\dots \dots \sum_{t=0}^{\infty} \frac{\varphi_t}{(1+r)^t} \cdot \left[Y_t - \sum_{n=1}^N \left(x_t^n \cdot DPI^n + \left(\sum_{u=0}^{t-1} x_u^n \right) \cdot DPE_t^n \right) \right] + \lambda \cdot \left[\sum_{t=0}^{\infty} \frac{C_t}{(1+i)^t} - \sum_{t=0}^{\infty} \frac{Y_t}{(1+i)^t} \right]$$

La recherche de la chronique optimale des financements privés $Y(t)$ conduit, en annulant les dérivées partielles par rapport à $Y(t)$, aux conditions du premier ordre suivantes :

$$\boxed{\frac{\varphi_t}{(1+r)^t} = \frac{\lambda}{(1+i)^t}} \Leftrightarrow \boxed{\varphi_t = \lambda \cdot \left(\frac{1+r}{1+i} \right)^t}$$

Ce qui implique notamment, pour $u > t$:

$$\psi_u = \frac{\varphi_u}{\varphi_t} = \left(\frac{1+r}{1+i} \right)^{u-t}$$

Il apparaît ainsi que le coût fictif de rareté des fons publics φ_t est une suite géométrique de raison $\frac{1+r}{1+i}$ et il en va de même pour la suite des poids ψ_u

Si le taux d'intérêt privé i est supérieur au taux d'actualisation r , ces deux suites sont décroissantes.

Dans le cas où l'on aurait $i = r$, le coût fictif de rareté des fons publics serait constant (et égal à λ) et le poids serait constant et égal à 1.

Il subsiste alors une seule inconnue : λ . On détermine sa valeur en écrivant que la contrainte intertemporelle de financement est serrée.

Annexe 5. Fractionnement versus programmation en nombres entiers

Plaçons nous dans le cas d'école où tous les projets auraient le même coût d'investissement, supposé financé entièrement sur fons publics, et où l'exploitation n'aurait aucune incidence en termes de recettes ou de dépenses publiques.

Supposons alors que par exemple : $I^n = 1$ million d'euros (1 M€) quel que soit n .

Supposons en outre que la contrainte de rationnement des fonds publics soit égale à C_t l'année t , avec $C_t > 0$, et soit nulle toutes les autres années : on ne peut financer d'infrastructure que l'année t .

On va d'abord recourir à la programmation en nombres entiers, puis on examinera ce que produirait la méthode de programmation linéaire.

a. Programmation en nombres entiers

Plusieurs cas se présentent selon la valeur de l'enveloppe budgétaire C_t

– Si $0 \leq C_t < 1$, on ne peut réaliser aucun projet.

– Si $1 \leq C_t < 2$, on peut réaliser au maximum 1 projet. Comment le choisir ? Il faut évidemment commencer par calculer pour chaque projet n le bénéfice actualisé à l'origine des temps qu'il dégage, en supposant que sa date de mise en service est t (qui, sauf peut-être fortuitement pour certains projets, diffère de sa date optimale de réalisation \hat{t}^n) ; soit BOT_t^n cette valeur. Il faut ensuite parmi les N projets sélectionner le projet qui dégage le maximum de cette valeur, c'est-à-dire chercher la solution de $Max_n BOT_t^n$.

Une méthode simple consiste à classer les projets par ordre de BOT_t décroissant et à retenir⁴¹ le projet classé $n^{\circ}1$.

– Si $2 \leq C_t < 3$, on applique la méthode précédente et l'on sélectionne les projets classés $n^{\circ}1$ et $n^{\circ}2$.

– Et ainsi de suite : on sélectionne les c_t premiers projets, avec :

$c_t =$ partie entière de $C(t)$

b. Recours au fractionnement et à la programmation linéaire

Affectons maintenant à chaque projet n d'un multiplicateur (variable continue) x_t^n , comprise entre deux limites : $x_t^n \geq 0$ et $x_t^n \leq 1$

Prenant alors les N variables x_t^n comme inconnues, cherchons à résoudre le problème de programmation linéaire suivant :

$$Max_{x_t^n} \sum_{i=1}^N x_t^n \cdot BOT_t^n \text{ sous les contraintes :}$$

$$x_i(t) \geq 0 \text{ et } x_i(t) \leq 1 \text{ ainsi que } \sum_{n=1}^N x_t^n \leq C_t$$

On trouve alors les résultats suivants, qui se réfèrent au classement des projets par ordre décroissants de BOT_t

41. On simplifiera en supposant que toutes les $VAN_i(t)$ sont différentes (pas d'ex æquo).

- Si $0 \leq C_t < 1$, il faut réaliser une fraction C_t du projet classé ci-dessus $n^{\circ}1$
- Si $1 \leq C_t < 2$, il faut réaliser intégralement le projet classé ci-dessus $n^{\circ}1$ et réaliser une fraction égale à $(C_t - 1)$ du projet classé ci-dessus $n^{\circ}2$.
- Et ainsi de suite.

Il est clair que la mise en œuvre d'une fraction de projet n'a pas de sens concret, en tout cas dans le cas d'école considéré ici où la disponibilité budgétaire se limite à une seule année, ce qui interdit de compléter ledit projet l'année suivante.

Ce résultat est donc différent de celui fourni par la programmation en nombre entier.

Mais la programmation linéaire peut être utile si on la complète par la règle suivante : on corrige le résultat en ne sélectionnant que les projets qu'elle conduit à retenir en entier, et en supprimant tous les projets qu'elle conduit à retenir partiellement. On peut ainsi procéder par itération.

TROISIÈME PARTIE

L'AIDE À LA DÉCISION PUBLIQUE

Le calcul économique et la décision publique se trouvent confrontés à de nouvelles pratiques.

- L'une, qui n'est pas franchement une nouveauté, mais qui risque de se développer, consiste à ignorer les résultats des évaluations *a priori*. Comme le calcul économique ne peut aider à fourbir un argumentaire, la décision est prise sur des critères exclusivement politiques, éventuellement rattachés à des considérations d'aménagement du territoire. Or, pour les TGV comme pour les autoroutes, développer les tronçons non rentables du réseau est évidemment source de baisse de l'efficacité globale. Pourtant, les élus continuent à réclamer des autoroutes, si possible à péage, car ils souhaitent ainsi être déchargés des questions de financement et de maintenance.

- Une autre difficulté, nouvelle mais de plus en plus nette, réside dans les procédures de type débat public. Pour beaucoup, le débat public est fait pour s'abstraire des recommandations données par le calcul économique. Qu'il s'agisse de s'opposer à un projet annoncé comme rentable pour la collectivité, ou que l'objectif soit au contraire de promouvoir un projet annoncé comme peu pertinent, le calcul économique est vécu comme antidémocratique. De plus, le débat public est souvent confisqué par ceux qui souhaitent lui dicter leurs conclusions. Ainsi, confiné dans un milieu étroit, rejeté dans le « monde » des experts, le calcul économique est mis sur la touche.

Face à cette nouvelle donne, l'objectif du réseau de recherche a été d'abord de comprendre ces évolutions, puis de montrer comment le calcul économique peut retrouver, dans ce nouveau contexte, une certaine pertinence.

- Dans cette perspective, Charles Raux, Stéphanie Souche et Aurélie Mercier nous rappellent dans le chapitre 10 ce que peut et ce que ne peut pas le calcul économique. En introduisant les questions liées à l'équité, on entre dans le champ de l'acceptabilité de la décision publique, un champ beaucoup plus mouvant que le calcul économique.

- Une des façons d'entrer de plain pied dans les questions liées à l'équité consiste à s'intéresser aux impacts des infrastructures de transport. Comme le rappelle Stef Proost dans le chapitre 11, il est possible, grâce aux modèles d'équilibre général calculable d'évaluer ce que pourraient être les gains et les pertes liés à un projet. Ces résultats peuvent ensuite être utilisés dans la négociation qui accompagne la décision publique, par exemple pour établir les éventuelles indemnités.

- Mais cet outil ne peut être considéré comme une panacée. Ainsi que le montre Jacques-François Thisse dans le chapitre 12, la localisation des équipements collectifs et l'évaluation de leurs impacts restent des sources de tensions dans le corps social et donc dans le processus politique. Il n'est d'ailleurs pas impossible qu'émergent de l'analyse des situations d'indécidabilité.

- Le chapitre 13, rédigé par Alain Trannoy, aboutit à un constat proche sinon similaire. Même si les outils proposés par l'auteur procèdent à un tri parmi les raisonnements en termes d'équité spatiale, entre ceux qui peuvent s'appuyer sur une démarche rationnelle et ceux qui en sont dépourvus, la solution consensuelle ne va pas de soi. Même en dégageant un langage et des raisonnements qui peuvent être partagés par les parties prenantes, la voie du compromis reste difficile à trouver.

- Pour comprendre combien sont grandes les difficultés à résoudre, il ne faut pas hésiter à entrer dans les dédales de l'acceptabilité. Cette notion « fourre-tout » renvoie en effet à des comportements variés, souvent observables, qui ont une forte tendance à se résumer à une préférence pour le *statu quo*. Dans le chapitre 14, Charles Raux, Stéphanie Souche et Lucie Vsakova s'intéressent aux figures de l'acceptabilité et aux éléments qui, dans le prolongement du chapitre précédent, pourraient contribuer à ce que se construise au moins un langage commun entre les participants aux débats publics.

CHAPITRE 10

DE LA MODÉLISATION DES COMPORTEMENTS AU CALCUL ÉCONOMIQUE : L'ÉQUITÉ DES POLITIQUES DE TRANSPORT

Charles Raux, Stéphanie Souche et Aurélie Mercier¹

En matière de débat sur les politiques de transport, qu'il s'agisse par exemple de projets d'infrastructure ou de changements tarifaires, les arguments d'équité et de justice prennent souvent le pas sur ceux de l'efficacité au sens économique. Cette dernière est la raison d'être du calcul économique, lequel semble en apparence démunie quand il s'agit d'équité.

L'équité en matière de politique de transport peut être déclinée selon trois dimensions (Raux et Souche, 2004) : l'équité verticale au sens de l'attention à porter aux personnes les plus fragiles économiquement ; l'équité horizontale au sens de la répartition de la charge entre contribuables et usagers des transports, et entre usagers des différents modes ; enfin et surtout, l'équité spatiale qui se réfère à la garantie d'accessibilité aux aménités et à la liberté d'aller et venir.

Dans ce chapitre, nous montrons que le calcul économique peut apporter des éléments de diagnostic quant à l'équité des politiques de transport. Les indicateurs de surplus sont calculés sur la base de modèles de demande, capturant les variations de flux de déplacements ou des comportements des consommateurs, en réaction à des variations de l'offre de transport. Il importe donc de s'assurer que ces indicateurs produits à l'aide de modèles de demande sont cohérents avec les principes du calcul économique.

Un bref rappel théorique sur les liens entre théorie du surplus et fonction de demande est donc un préalable nécessaire (section 1). Ensuite nous exposons l'état de l'art de la cohérence entre la modéli-

1. Laboratoire d'Économie des Transports, CNRS, Université Lyon 2, ENTPE

sation de la demande de transport – pour les modèles les plus couramment utilisés – et le calcul de surplus (section 2). Vouloir traiter de la thématique des gagnants et des perdants, et plus généralement de l'équité, amène logiquement à aborder la question de la comparabilité des utilités et celle de l'agrégation des utilités individuelles (section 3).

1. FONCTION DE DEMANDE DE TRANSPORT ET THÉORIE DU SURPLUS

En comparant des options d'utilisation de ressources rares, telles que l'argent public, l'espace ou les ressources environnementales disponibles, le calcul économique vise à aider à la décision des collectivités publiques. Son corpus théorique de base est celui de la micro-économie : aussi, pour établir cette base théorique, il nous faut rappeler brièvement les concepts d'utilité du consommateur et d'utilité collective. Ces concepts sont ensuite appliqués au cas particulier de la variation d'utilité collective dans le cas du transport, pour déboucher enfin sur le surplus des usagers et la fonction de demande de transport.

1.1. L'utilité du consommateur

Soit un individu i qui dispose d'un revenu r_i pour acheter différents biens de consommation dénotés j ($j = 1, m$) de prix unitaire p_j en quantités q_j^i . Sous les hypothèses classiques de rationalité², l'individu est supposé capable de classer les différents paniers de biens, classément représenté par un indice U_i appelé utilité (ou satisfaction) de l'individu i , soit :

$$U_i(q_1^i, q_2^i, \dots, q_m^i)$$

Cette utilité est à maximiser sous la contrainte de revenu de i , soit

$$\sum_{j=1}^m p_j q_j^i = r_i$$

A l'optimum

$$\frac{\partial U_i}{\partial q_j^i} = \lambda_i p_j$$

où λ_i est le multiplicateur de Lagrange associé à la contrainte de revenu du consommateur i . Les utilités marginales pour l'individu i de

2. Pour un exposé complet de la théorie micro-économique voir (Mas-Colell *et al*, 1995).

la consommation des biens j sont proportionnelles aux dépenses marginales que représentent les prix unitaires de ces biens.

En outre, à l'optimum nous avons

$$\lambda_i = \frac{\partial U_i}{\partial r_i}$$

Ce multiplicateur de Lagrange s'interprète comme l'utilité marginale du revenu de l'individu i .

1.2. L'utilité collective

Le modèle général du calcul économique suppose l'existence d'une fonction de bien-être collectif \mathbf{U} , fonction des satisfactions individuelles³.

Autour de l'optimum, où l'on est supposé se trouver, la variation marginale d'utilité collective résultant de variations marginales des quantités de biens disponibles s'écrit :

$$d\mathbf{U} = \sum_{i=1, n, j=1, m} \sum \frac{\partial \mathbf{U}}{\partial U_i} \frac{\partial U_i}{\partial q_j^i} dq_j^i = \sum_{i=1, n, j=1, m} \sum \frac{\partial \mathbf{U}}{\partial U_i} \lambda_i p_j dq_j^i$$

On introduit l'hypothèse forte de répartition optimale des revenus, ce qui signifie que du point de vue de la collectivité, une variation élémentaire de revenu a la même valeur quel que soit l'individu considéré : formellement, l'utilité collective marginale du revenu individuel est identique pour tous les individus, i.e.

$$\lambda_i \frac{\partial \mathbf{U}}{\partial U_i} = 1 \quad \text{à une constante près.}$$

On obtient :

$$d\mathbf{U} = \sum_j p_j dq_j \quad \text{avec} \quad q_j = \sum_{i=1}^n q_j^i$$

en supposant que les prix p_j ne dépendent pas de la consommation des agents (hypothèse de concurrence parfaite).

3. Cette hypothèse d'existence suppose donc résolu le problème de l'agrégation des préférences individuelles. Pour un exposé complet voir le chapitre I de Bloy *et al* (1976).

1.3. La variation d'utilité collective dans le cas d'un projet de transport

La démarche consiste à introduire le bien particulier qu'est le service de transport, et à travers cette particularité, le temps nécessaire d'une part au transport, d'autre part à la consommation des autres biens⁴.

Nous supposons dans un premier temps une transformation marginale des quantités analysées pour passer ensuite à une transformation structurelle.

Chaque unité de bien j nécessite un temps de consommation t_j , supposé constant au cours de la transformation. Nous considérons le transport comme un $(m+1)^{\text{ème}}$ bien dont la quantité q_{m+1}^i représente le transport consommé par l'individu i (exemple, kilomètres parcourus ou choix de se déplacer ou non). Son prix p_{m+1} est le prix payé par l'usager : par exemple, pour la route il inclura les coûts d'amortissement et d'entretien du véhicule, les frais de carburant et les éventuels péages et frais de stationnement. Enfin, le temps de consommation unitaire de transport t_{m+1} peut, à la différence des autres biens j , varier au cours de la transformation (augmentation ou diminution du temps de parcours générée par le projet de transport).

Alors la maximisation de U_i par l'individu i se fait sous les deux contraintes de revenu r_i et de temps θ_i disponibles, soit

$$\sum_{j=1}^{m+1} p_j q_j^i = r_i, \text{ et}$$

$$\sum_{j=1}^{m+1} t_j q_j^i = \theta_i$$

La condition d'optimum devient

$$\frac{\partial U_i}{\partial q_j^i} = \lambda_i (p_j + v_i t_j)$$

où v_i est la valeur du temps de l'individu i . On retrouve la condition de proportionnalité entre utilité marginale et dépense marginale : ici la dépense marginale agrège le prix unitaire du bien et la valeur du temps unitaire de sa consommation.

4. Cette section s'appuie sur la section 1 du chapitre II de Bloy *et al* (1976), en adaptant la présentation et les notations à notre propos.

Il s'ensuit, toujours avec l'hypothèse de répartition optimale des revenus, que la variation d'utilité collective devient

$$dU = \sum_{j=1}^{m+1} \sum_{i=1}^n (p_j + v_i t_j) dq_j^i$$

Cependant, pour chaque individu le temps total disponible est une donnée physique identique. En rappelant que le temps unitaire de consommation des biens j est constant, tandis que celui du transport varie, il s'ensuit que, au cours de la transformation,

$$\sum_{j=1}^{m+1} t_j dq_j^i + q_{m+1}^i dt_{m+1} = 0$$

En outre, on peut considérer que la valeur du temps est identique pour tous les individus⁵, soit $v_i = v$.

d'où

$$dU = \sum_{j=1}^{m+1} \sum_{i=1}^n p_j dq_j^i - v dt_{m+1} \sum_{i=1}^n q_{m+1}^i = \sum_{j=1}^m p_j dq_j + p_{m+1} dq_{m+1} - v dt_{m+1} q_{m+1}$$

Nous supposons que l'on se trouve en situation de plein emploi des ressources, dans une économie fermée, et que, dans le cadre de la transformation marginale étudiée ici, cette dernière a des effets négligeables sur les entreprises ($k = 1, p$) autres que celle produisant le service de transport ($p+1$). L'hypothèse d'économie fermée permet de réécrire le premier terme de l'équation précédente du point de vue des entreprises consommatrices et productrices de biens et services, sous la forme

$$\sum_{j=1}^m p_j dq_j = \sum_{j=1}^m \sum_{k=1}^p p_j dq_j^k + \sum_{j=1}^m p_j dq_j^{p+1}$$

L'hypothèse d'effets négligeables implique que $\sum_{j=1}^m \sum_{k=1}^p p_j dq_j^k = 0$

$$\text{D'où} \quad \sum_{j=1}^m p_j dq_j = \sum_{j=1}^m p_j dq_j^{p+1}$$

5. Cette hypothèse peut être posée sans perte de généralité, puisque l'on peut étendre les résultats qui suivent à une multitude de groupes d'individus de valeur du temps identiques, les valeurs du temps (et donc les courbes de demande, cf. *infra*) différant entre les groupes.

Ce terme n'est autre que l'opposé de la dépense totale de l'entreprise productrice du service de transport.

La variation d'utilité collective peut se réécrire ainsi

$$dU = \sum_{j=1}^m p_j dq_j^{p+1} + p_{m+1} dq_{m+1} - v_i dt_{m+1} q_{m+1}$$

ou encore, en introduisant le terme $q_{m+1} dp_{m+1}$

$$dU = -q_{m+1}(dp_{m+1} + v_i dt_{m+1}) + \left(q_{m+1} dp_{m+1} + p_{m+1} dq_{m+1} + \sum_{j=1}^m p_j dq_j^{p+1} \right)$$

La variation d'utilité collective est donc la somme de deux variations de surplus, la première relative aux consommateurs (usagers de l'infrastructure), la deuxième relative à l'entreprise de transport.

En effet, la variation des surplus des consommateurs se calcule sur la base d'une variation de prix dp_{m+1} et d'une variation de temps de parcours dt_{m+1} , cette dernière valorisée par une valeur du temps v_i , variations rapportées au trafic total q_{m+1} .

La variation de surplus de l'entreprise de transport est la somme de gains provenant de la variation de prix appliquée au trafic, soit $q_{m+1} dp_{m+1}$, de la variation de trafic acquittant le prix, soit $p_{m+1} dq_{m+1}$,

et des dépenses, soit $\sum_{j=1}^m p_j dq_j^{p+1}$

Le passage à une transformation structurelle ne pose pas problème, si on considère celle-ci comme une suite de transformations marginales, et si on maintient l'hypothèse que ces transformations marginales successives ont des effets négligeables sur les entreprises autres que celle produisant le service de transport. Soit, en passant d'une situation initiale 1 à une situation finale 2 :

$$\Delta U = \int_1^2 \left[-q_{m+1}(dp_{m+1} + v_i dt_{m+1}) + \left(q_{m+1} dp_{m+1} + p_{m+1} dq_{m+1} + \sum_{j=1}^m p_j dq_j^{p+1} \right) \right]$$

1.4. Le surplus des usagers et la fonction de demande de transport

Intéressons-nous maintenant de plus près à la variation de surplus des usagers, soit

$$\Delta S = \int_1^2 \left[-q_{m+1}(dp_{m+1} + v_i dt_{m+1}) \right]$$

La combinaison d'une composante monétaire et d'une composante temporelle du déplacement, cette dernière valorisée par une valeur du temps, permet d'introduire la notion de coût généralisé d'un déplacement : ce dernier s'écrira en toute généralité pour un individu et sur une liaison donnés

$$c = p + vt$$

où p est le coût monétaire du déplacement, t sa durée et v la valeur du temps de l'individu.

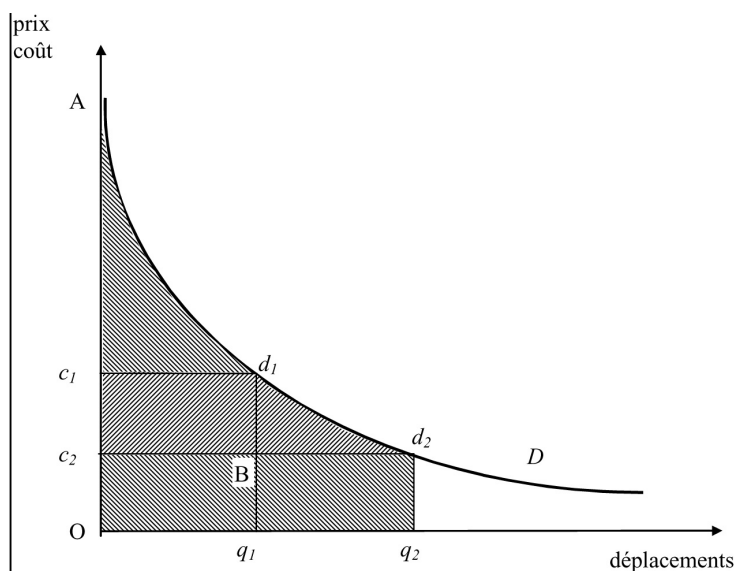
En nous affranchissant de l'indice du bien transport pour alléger la notation, nous obtenons

$$\Delta S = \int_1^2 -qdc$$

Jules Dupuit a donné une définition opératoire de la courbe de demande dans son célèbre mémoire au sujet « de la mesure de l'utilité des travaux publics » (1844). Même si cette loi de demande reste pour une bonne part inconnue, il énonçait quelques caractéristiques générales, à savoir que la consommation augmente quand le prix diminue, et que l'augmentation due à un abaissement de prix est d'autant plus considérable que le prix est déjà lui-même plus bas, d'où une forme convexe attribuée à la courbe de demande (cf. Figure 1).

Figure 1

Courbe de demande et surplus des usagers



On peut étendre cette approche au cas plus général de la demande qui s'exprime sur un axe de transport (exemple d'une liaison routière, ferroviaire ou aérienne), en fonction du coût généralisé du déplacement.

En situation initiale, quand le coût généralisé s'établit à c_1 , un nombre q_1 d'usagers de l'infrastructure sont prêts à supporter un coût supérieur ou égal à c_1 . L'utilité que retirent ces usagers de leur déplacement est égale à l'aire Oq_1d_1A . Le coût généralisé qu'ils doivent supporter est égal à l'aire $Oq_1d_1c_1$. Le surplus (avantage positif) qu'ils retirent de leur déplacement est donc égal à l'aire c_1d_1A .

En situation finale, si le coût généralisé s'établit à c_2 inférieur à c_1 , les q_1 usagers déjà utilisateurs de l'infrastructure profitent de l'abaissement de ce coût, d'où une variation de surplus positive égale à l'aire $c_2Bd_1c_1$. En outre, des usagers supplémentaires (soit l'augmentation de trafic $q_2 - q_1$), qui n'étaient pas prêts à supporter le coût c_1 , trouvent désormais un intérêt à emprunter l'axe⁶, puisque l'avantage qu'ils retirent de leur déplacement est supérieur ou égal à c_2 . Le surplus qu'ils retirent de leur déplacement est égal à l'aire Bd_2d_1 . La variation de surplus total est donc égale à l'aire $c_2d_2d_1c_1$.

Formellement, l'utilité des consommateurs s'évalue comme l'intégrale de l'aire sous la courbe de demande ordinaire (ou marshallienne). En pratique comme cette courbe de demande est inconnue, on utilise l'approximation triangulaire introduite par Dupuit, par interpolation entre les deux points d_1 et d_2 observés (ou modélisés).

2. DU MODÈLE DE COMPORTEMENT AU CALCUL DE SURPLUS

De nombreux indicateurs d'accessibilité ont été développés pour rendre compte d'une variation de qualité apportée par une variation de l'offre de transport. Ces indicateurs ont parfois été interprétés par leurs auteurs comme des indicateurs de surplus. Dans cette section nous rappelons l'état de l'art de la cohérence entre la modélisation de la demande des consommateurs et le calcul de surplus. L'analyse est menée pour deux catégories de modèles les plus couramment utilisés, à savoir le modèle gravitaire de distribution spatiale et les modèles de choix discret à utilité aléatoire. Etant donné qu'une variation de l'offre de transport se traduit avant tout par une variation des conditions

6. Cette augmentation de la consommation de transport, que l'on appelle aussi parfois « trafic induit », inclut aussi bien les nouveaux usagers que les usagers initiaux qui intensifient leurs déplacements.

d'accès aux différents lieux d'activités, nous focalisons notre analyse sur l'équité spatiale.

2.1. Du modèle gravitaire à la théorie de l'accessibilité

Le modèle gravitaire⁷ a été développé pour modéliser la distribution spatiale des déplacements entre origines et destinations. Il suppose que l'espace où se déroulent ces déplacements ait été découpé en zones. La loi de distribution des déplacements s'exprime ainsi :

$$T_{ij} = kO_iD_jf(c_{ij})$$

où :

T_{ij} est le flux de déplacements de la zone i vers la zone j ,

k un paramètre du modèle,

O_i la « masse » émettrice de la zone origine i (exemple, le nombre d'actifs dans la zone),

D_j la « masse » attractive de la zone de destination j (exemple, le nombre d'emplois dans la zone),

$f(c_{ij})$ une fonction de résistance au déplacement, avec c_{ij} le coût généralisé de déplacement de la zone i vers la zone j .

Les fondements théoriques de ce modèle ont été posés par Wilson (1970).

On introduit en outre des contraintes de cohérence, relatives aux marges de la matrice des T_{ij} . Avec une contrainte telle que $\sum_j T_{ij} = O_i$ on obtient le modèle contraint aux origines, ce qui donne

$$T_{ij} = O_i \frac{D_j f(c_{ij})}{\sum_k D_k f(c_{ik})} \quad (1)$$

La partie de droite de ce modèle, soit $\frac{D_j f(c_{ij})}{\sum_k D_k f(c_{ik})}$, peut aussi être

interprétée comme un modèle de probabilité de choix de destination à partir d'une zone origine i donnée.

7. Cette présentation a été adaptée et enrichie à partir de Bonnafous et Masson (2003).

Si l'on rajoute la contrainte $\sum_k T_{kj} = D_j$ on obtient le modèle doublement contraint qui se reformulerait en :

$$T_{ij} = k_i O_i k_j D_j f(c_{ij})$$

$$\text{avec : } k_i = \frac{1}{\sum_j k_j D_j f(c_{ij})} \text{ et } k_j = \frac{1}{\sum_i k_i O_i f(c_{ij})}$$

La résolution de ces équations passe par un calcul itératif (méthode de Furness). Les méthodes sont maintenant éprouvées et convergent correctement en général.

Le calage économétrique s'appuie en général sur une fonction de résistance de forme exponentielle négative, soit :

$$f(c_{ij}) = \exp(-\beta c_{ij}) \quad (2)$$

$$\text{On définit } A_i = \sum_j D_j \exp(-\beta c_{ij}) \quad (3)$$

indicateur d'accessibilité le plus courant (Hansen, 1959) : on retrouve le dénominateur du modèle (1).

La variation de surplus, dérivée du modèle gravitaire, entre une situation 1 (avant) et une situation 2 (après), associée à une zone particulière i , peut se calculer ainsi

$$\Delta S_i = - \sum_j \int_{c_{ij}^1}^{c_{ij}^2} T_{ij} dc_{ij} \quad (4)$$

Si l'on se réfère à la Figure 1 et son commentaire, le surplus couvre entièrement les deux types de comportements des usagers en réaction à une baisse de coût généralisé c_{ij} , par exemple un gain de temps : d'une part ceux qui utilisent ce temps gagné sur les déplacements dans d'autres activités que la mobilité, d'autre part ceux qui réinvestissent ce temps gagné dans une intensification de leur mobilité (augmentation du nombre de déplacements ou de la portée de ceux-ci). Ce dernier type de comportement désigne bien le « trafic induit », celui qui profite des opportunités nouvelles de destination offertes par une amélioration de l'offre de transport.

Neuberger (1971) a montré qu'avec le modèle (1) la variation de surplus de la zone i s'écrivait :

$$\Delta S_i = \frac{O_i}{\beta} \ln \left(\frac{\sum_k D_k \exp(-\beta c_{ik}^2)}{\sum_k D_k \exp(-\beta c_{ik}^1)} \right) = \frac{O_i}{\beta} (\ln A_i^2 - \ln A_i^1) \quad (5)$$

Le calcul de variations d'accessibilité spatiale à partir d'un modèle de distribution spatiale de type gravitaire trouve donc sa cohérence avec le calcul économique classique de surplus.

2.2. Le modèle de choix discret à utilité aléatoire

Comme, en matière de déplacements, le consommateur est le plus souvent face à un choix discret, i.e. se déplacer ou non, choisir une destination, un mode de transport ou un itinéraire parmi un ensemble d'options possibles, les modèles de choix discret se sont progressivement imposés en raison de leurs avantages théoriques et empiriques.

On considère⁸ une population d'individus faisant face au même ensemble d'options A (exemple, un ensemble de modes de transport) et on cherche à estimer la fraction de la population qui choisit une option donnée. La population peut être divisée en sous-populations homogènes par rapport à des facteurs socio-économiques observables. Dans chaque sous-population, on modélise le comportement d'un individu représentatif comme un comportement déterministe. Néanmoins, l'observation montre que, même à l'intérieur d'une sous-population homogène, les individus ne font pas tous le même choix, d'où une composante aléatoire dans le comportement. Les individus vont donc, en cohérence avec la théorie néoclassique, choisir l'option qui maximise l'utilité qu'ils en retirent. Cette utilité a deux composantes, l'une déterministe, l'autre aléatoire. Soit

$$U_j^i = u_j^i + \epsilon_j^i$$

représentant l'utilité pour l'individu i du choix de l'option $j \in A$, u_j^i étant la composante déterministe (observable et mesurable) de l'utilité et ϵ_j^i une variable aléatoire de moyenne nulle.

On suppose qu'au sein d'une sous-population homogène, les individus ne diffèrent les uns des autres que selon les caractéristiques et

8. Cette présentation est adaptée de Anderson *et al* (1992).

facteurs non-observés qui influencent leur choix. On peut donc prédire au mieux la probabilité de choix de l'option j par l'individu i , soit :

$$P_j^i = \Pr(U_j^i = \max_{k \in A} U_k^i)$$

2.2.1. Le modèle logit multinomial

Une large classe de modèles de choix discrets présente des propriétés remarquables et une relative facilité de calage, ce qui en a généralisé l'usage malgré certaines limites qu'il n'y a pas lieu de discuter ici. Le modèle logit multinomial dérive de l'hypothèse que les aléas ϵ sont distribués indépendamment et identiquement selon une loi double-exponentielle. La probabilité de choix de l'option j par l'individu i suit alors une loi logistique, soit :

$$P_j^i = \frac{\exp\left(\frac{u_j^i}{\mu}\right)}{\sum_{k \in A} \exp\left(\frac{u_k^i}{\mu}\right)}$$

Ben Akiva et Lerman (1979) ont proposé comme mesure de l'attractivité de l'ensemble d'options A (et donc comme mesure d'accessibilité pour un ensemble A de destinations), l'espérance du maximum des utilités. Dans le cas du logit (Anderson *et al*, 1992) cette espérance est égale à

$$\mu \ln \sum_{j \in A} \exp\left(\frac{u_j^i}{\mu}\right)$$

2.2.2. Le modèle logit emboîté

Une extension du modèle logit est le modèle logit emboîté (ou hiérarchique) développé par Ben-Akiva (1985) – qui est une réponse au problème IIA⁹.

Considérons en effet un processus de choix emboîté à deux étapes (Anderson *et al*, 1992). L'ensemble d'options Ω est partitionné en sous-ensembles A_l qui regroupent chacun des options ayant plusieurs caractéristiques observables en commun (par exemple l'ensemble A_1 des voitures particulières, l'ensemble A_2 des bus quelle que soit leur

⁹. Independence of Irrelevant Alternatives ou « paradoxe bus bleu / bus rouge » (Anderson *et al*, 1992, p. 24).

couleur, etc.). On suppose que l'individu sélectionne le sous-ensemble A_i dans Ω avec une certaine probabilité, puis une option du sous-ensemble A_i avec une probabilité dépendant de l'utilité de cette option, soit :

$$P_{j \in A_i}^i = \frac{\exp\left(\frac{u_j^i}{\mu_2}\right)}{\sum_{k \in A_i} \exp\left(\frac{u_k^i}{\mu_2}\right)}$$

Le choix du sous-ensemble A_i par l'individu i dépend de l'utilité globale de ce sous-ensemble. Pour cela, Ben-Akiva propose l'espérance du maximum des utilités de A_i , soit :

$$S_i^i = \mu_2 \ln \sum_{j \in A_i} \exp\left(\frac{u_j^i}{\mu_2}\right)$$

d'où la probabilité de choix de A_i par l'individu i :

$$P^i(A_i) = \frac{\exp\left(\frac{S_i^i}{\mu_1}\right)}{\sum_{m \in \Omega} \exp\left(\frac{S_m^i}{\mu_1}\right)}$$

Ce modèle est cohérent avec la maximisation d'une fonction d'utilité aléatoire si $\mu_1 > \mu_2$ (Mc Fadden, 1978). Autrement dit, les préférences au niveau supérieur doivent être plus hétérogènes que celles du niveau inférieur.

Le modèle logit emboîté s'applique de manière identique¹⁰ à toute hiérarchie possible de choix, par exemple le choix de destination puis le choix du mode.

2.2.3. Théorie de l'accessibilité et choix discret

En développant la « théorie de l'accessibilité urbaine », Koenig (1974) arrive à un résultat identique à celui de Neuburger en matière de surplus lié à l'accessibilité, mais sur la base d'hypothèses fondées dans la théorie micro-économique du consommateur : 1) à chaque

10. Sous réserve de ses autres limites (cf. Ben-Akiva et Lerman, 1985).

destination (exemple, l'emploi) le consommateur associe une utilité nette, différence entre l'utilité brute de la destination (exemple, le salaire) et le coût généralisé du déplacement ; 2) l'utilité brute d'une destination est une variable aléatoire de loi de distribution exponentielle négative. Les déplacements des consommateurs se répartissent alors selon la loi de distribution gravitaire.

En fait, le modèle fondé sur la théorie de l'accessibilité peut être réécrit très simplement sous la forme d'un modèle de choix discret à utilité aléatoire. En effet la probabilité de choix d'une destination j par l'individu i s'écrit :

$$P_j^i = \frac{D_j \exp(-\beta c_{ij})}{\sum_{k \in A} D_k \exp(-\beta c_{ik})}$$

La composante déterministe du modèle de choix discret est donc :

$$u_j^i = \frac{1}{\beta} \ln D_j - c_{ij}$$

2.2.4. Modèles de choix discret et calcul du surplus

Small et Rosen (1981) ont explicité les conditions dans lesquelles le calcul économique pouvait être adapté au cas des modèles de choix discrets. Dans les modèles de choix continu, le surplus peut être calculé à partir de l'aire située sous la courbe de demande compensée, dont la valeur est approchée en pratique et sous réserve d'hypothèses additionnelles, par l'aire située sous la courbe de demande ordinaire (cf. plus haut le calcul de surplus à la Dupuit).

Cette pratique reste valable pour une classe particulière de modèles de choix discret¹¹, ceux pour lesquels la fonction d'utilité est séparable de manière additive sous la forme de trois composantes, soit

$$u_j^i = V^i(r^i) + W_j(p_j, q_j, r^i; C^i) + \varepsilon_j^i \quad \text{où}$$

$V^i(r^i)$ ne dépend que des caractéristiques de l'individu i et notamment son revenu r^i ,

W_j est la fonction d'utilité individuelle déterministe et observable dont la forme est identique pour tous les consommateurs,

11. Sous des hypothèses additionnelles et notamment celle que la fonction de demande compensée est approchée de manière adéquate par la fonction de demande marshallienne.

p_j, q_j les prix et quantité du bien j ,

C^i un vecteur de caractéristiques observables du consommateur i et de caractéristiques du bien qui peuvent varier selon les consommateurs.

Il est à noter au passage que, quand on compare deux options, la différence d'utilités fait que le terme $V^i(r^i)$ disparaît.

En supposant que chaque individu i représente un groupe de N^i individus identiques quant à leurs revenus et leurs caractéristiques mesurées, que l'utilité marginale du revenu λ est constante dans le groupe, et que les effets revenus sont négligeables, on trouve dans le cas du modèle logit que la variation de surplus s'écrit

$$\Delta E = -\left(\frac{N}{\lambda}\right) \left[\ln \sum_{j \in A} \exp(W_j) \right]_{W^0}^{W^f} \quad (6)$$

où \underline{W} est le vecteur de composantes W_j qui prend les valeurs initiale \underline{W}^0 et finale \underline{W}^f .

Comme $\sum_{j \in A} \exp(W_j)$ est le dénominateur du modèle logit, cette quantité est aisée à calculer.

Pour cette classe de modèles de choix discret à utilité « séparable-additive » le calcul de surplus donné par l'expression (6) est donc cohérent avec le calcul du bien-être collectif. C'est le cas notamment de tous les modèles *LRUM* (*Linear Random Utility Model*) et en particulier du modèle de choix discret fondé sur la théorie de l'accessibilité.

Compte tenu de leurs emboîtements de « logsums », la règle de cohérence de calcul du surplus s'applique aussi aux modèles logit emboîtés, respectant cette condition d'utilité « séparable-additive » dans leurs spécifications d'utilité à chaque étape.

2.2.5. Extensions du concept d'accessibilité

Le courant de la « time-geography », issu des travaux des géographes suédois Hägerstrand puis Lenntorp dans les années 70, marque un changement du paradigme de représentation de la mobilité : dans cette nouvelle représentation, l'individu réalise un ensemble d'activités en différents lieux, dont le passage de l'un à l'autre nécessite des déplacements, le tout dans un environnement spatial et temporel contraint. Ce courant se caractérise par un champ de recherche actif

depuis plus de vingt ans, qui commence à déboucher sur des modèles opérationnels de microsimulation des schémas d'activités-déplacements (sur une période d'une journée, voire plus).

Cette approche d'activités-déplacements ou « fondée sur les activités » (*activity-based*) a été reprise dans le cadre des modèles de choix discret à utilité aléatoire, en considérant non plus le choix élémentaire d'un déplacement, mais l'ensemble des choix réalisés dans le schéma d'activités. Ce type de modèle a donné lieu au développement d'un indicateur d'accessibilité fondé sur le schéma d'activités (*ABA, Activity Based Accessibility* ; Ben-Akiva et Bowman, 1998). L'ABA est l'espérance du maximum des utilités attachées aux différents schémas d'activités quotidiens possibles pour l'individu et donc, là encore, un « logsum » des utilités attachées à chacun de ces schémas d'activités.

Une autre approche (Miller, 1998) combine l'approche espace-temps des géographes et la théorie des choix discrets des économistes, pour aboutir à des indicateurs *STAM (Space Time Accessibility Measures)*. Parmi les différents STAM proposés par Miller, certains se réfèrent à des logsums d'utilités.

On retiendra de cette brève revue de ces indicateurs étendus d'accessibilité, qu'une condition suffisante pour que ces derniers soient cohérents avec le calcul économique est qu'ils soient fondés sur la théorie micro-économique de l'utilité, et qu'ils respectent la règle énoncée plus haut de spécification d'utilité « séparable-additive ».

3. POTENTIALITÉS ET LIMITES DU CALCUL ÉCONOMIQUE

L'un des thèmes principaux dans le débat sur l'acceptabilité est l'identification des gagnants et des perdants dans le cadre d'une politique particulière de transport. Pour le traiter il convient toutefois d'éclaircir certaines questions sous-jacentes : peut-on comparer des variations de surplus selon des zones (i.e. l'équité spatiale avec les accessibilités) ou selon des catégories économiques ou sociales (i.e. l'équité horizontale ou verticale) ? Ceci est l'objet de la première section.

Remarquons que nous posons d'abord la question de la comparabilité de ces mesures de surplus et non la question de la pondération de ces surplus en vue d'une agrégation en un bien-être collectif : cette question est abordée dans la section suivante.

3.1. Le problème de la comparabilité des utilités

D'une manière générale, la mesure du surplus à partir d'une fonction de demande variera en valeur absolue selon l'étendue des biens considérés dans le modèle de demande, ce qui résulte en des valeurs particulières des paramètres du modèle. Cela pose certes la question de la valeur de référence de l'utilité mais cette difficulté est contournée si l'on se contente d'analyser des *variations* de surplus. La deuxième question, inhérente à la comparabilité des utilités, a trait aux différences éventuelles d'échelle ou d'unité entre les utilités mesurées sur des catégories différentes.

Pour répondre à ces deux questions, nous analyserons séparément le cas du modèle gravitaire et celui du modèle de choix discret.

3.1.1. Le cas du modèle gravitaire

Suivant la spécification retenue pour le modèle gravitaire, on sera amené à comparer des accessibilités telles que décrites par l'équation (3) pour différentes zones. Cela est licite si les accessibilités sont effectivement liées à un modèle gravitaire avec une spécification unique de la fonction de coût généralisé de déplacement, entre autres le même paramètre β .

Donc il est possible de comparer des variations de surplus telles que décrites par l'équation (4), suite à la mise en place d'une politique de transport particulière, selon différentes zones, pour une catégorie donnée d'opportunités : exemple, les emplois, les commerces, etc.

Cela revient bien sûr à ignorer au sein de chaque zone l'hétérogénéité éventuelle selon d'autres caractéristiques sociales ou économiques des agents. La différence d'accessibilité prise en compte dans la variation de surplus dans l'équation (5) pèse d'autant plus lourd que O_i est élevé. Si l'on veut identifier les groupes d'individus gagnants – ou perdants – (repérés par leur zone commune d'origine de leurs déplacements), rien n'interdit de se limiter au deuxième facteur de cette expression i.e.

$$(\ln A_i^2 - \ln A_i^1)$$

Cette approche a été appliquée notamment à l'analyse de l'une des dimensions de l'acceptabilité (cf. chapitre sur l'acceptabilité ; Raux *et al*), à savoir l'équité spatiale, dans le cas du boulevard périphérique nord de Lyon à péage (cf. Encadré 1).

Encadré 1 : L'équité spatiale dans le cas boulevard périphérique nord de Lyon

Le boulevard périphérique nord de Lyon est un ensemble d'ouvrages d'une longueur de 11 km ouvert en 1997, dans sa version initiale sous la dénomination de TEO, avec péage et restriction des voies parallèles gratuites.

L'impact du péage et de cette restriction des voies parallèles a pu être évalué selon les différentes dimensions de l'équité, à savoir l'équité spatiale, verticale et horizontale. En termes d'équité spatiale, ce sont des mesures d'accessibilité entre les différentes zones de l'agglomération qui ont été calculées. Ces calculs ont été effectués à l'aide du « modèle stratégique de déplacements de l'agglomération lyonnaise » développé par le LET et la SEMALY en 1997. Le modèle repose sur un découpage spatial de l'agglomération en 25 zones. L'accessibilité retenue est celle de l'équation (3) et le modèle de distribution gravitaire est un modèle doublement contraint aux marges.

Le coût généralisé c_{ij} du déplacement de i à j agrège le temps et le prix, y compris le péage, avec une valeur du temps moyenne.

Les variations d'accessibilité aux emplois (entre avant et après la mise en service de TEO) à partir de chacune des 25 zones, ont été calculées sous la forme variations relatives, soit $\Delta A_i / A_i$ ce qui assure une certaine cohérence avec le calcul de surplus¹², bien que l'on ne tienne pas compte de la « masse » de la zone représentée par O_i . C'est une manière de ramener la variation de surplus de la zone à une variation de surplus d'un individu résidant dans cette zone, ce qui n'est pas sans intérêt dans un débat sur l'acceptabilité.

Ces variations d'accessibilité ont pu être rapportées à l'état des accessibilités des différentes zones avant l'ouverture de TEO. Les résultats des simulations effectuées ont montré qu'en termes d'accessibilité, la distribution des avantages et des inconvénients apportés par la nouvelle infrastructure s'opérait de manière inégalitaire entre les différentes zones de l'agglomération. Ces résultats venaient donc conforter la perception de l'opinion considérant que l'équité spatiale était gravement mise en cause (pour plus de détails, voir Raux et Souche, 2004).

3.1.2. *Le cas du modèle de choix discret à utilité « séparable-additive »*

Dans ces modèles, les probabilités de choix sont indifférentes aux valeurs absolues des utilités et ne dépendent que des différences entre utilités. Là encore, la question de la valeur de référence de l'utilité ne se pose pas.

Par contre, les unités des utilités des individus varient selon les individus. Une normalisation est préconisée (Ben-Akiva et Lerman,

12. En effet ces variations relatives s'assimilent à des variations élémentaires de logarithmes.

1985) en convertissant l'accessibilité en unités de l'une des variables x de la fonction d'utilité (en général le temps ou le coût de transport), soit :

$$\Delta A_i^x = \frac{\Delta A_i}{\partial A_i / \partial x}$$

Il est donc possible là aussi, de comparer des variations de surplus entre différentes catégories d'agents.

3.2. *Le problème de l'agrégation des utilités individuelles*

Dans l'approche utilitariste, la justice est un sous-produit de la recherche du plus grand bonheur pour le plus grand nombre, qui ignore les inégalités éventuelles dans la distribution des utilités. En réaction à cette approche, les théories contemporaines de la justice ont cherché à intégrer le concept d'équité. Un apport majeur est la théorie de la justice de Rawls (1971) reposant sur plusieurs principes, dont le fameux principe de différence qui s'appuie sur le critère du maximin : ce dernier vise à maximiser le surplus minimal. L'approche du maximin suppose de retenir un projet qui favorise l'intérêt des plus pauvres, à l'inverse du calcul économique traditionnel qui privilégie la plus grande variation de surplus collectif. Ce critère du maximin peut être formalisé de la façon suivante :

$$\text{Max}[\text{Min}_i \Delta S_i]$$

tandis que l'approche traditionnelle du calcul économique avance le critère d'efficacité optimale, soit :

$$\text{Max} \left[\sum_i \Delta S_i \right]$$

Face aux questions d'équité et de justice sociale, Wolfelsperger (1995) avance une conception néo-utilitariste de la justice, en supposant l'existence d'un « observateur idéal » chargé de représenter les préférences des individus. La fonction de justice sociale cherchant à maximiser le bien-être collectif, peut s'écrire sous la forme suivante :

$$W = \frac{1}{\alpha} \sum_i (U_i)^\alpha \quad \text{si } \alpha \neq 0 \quad \text{avec } U_i, \text{ l'utilité de l'individu } i \text{ et}$$

$$W = \sum_i \log(U_i) \quad \text{si } \alpha = 0$$

Le paramètre α permet de pondérer les utilités individuelles en fonction de leur niveau et représente ainsi le degré de prise en considération des inégalités. Wolfelsperger distingue trois cas principaux pour qualifier l'aversion pour l'inégalité :

Une préférence pour l'inégalité se caractérise par $\alpha > 1$. Cela sous-entend que le poids accordé à un individu, dans l'évaluation de la justice de la répartition, est d'autant plus important que cet individu bénéficie d'un niveau initial d'utilité élevé.

Une vision utilitariste se traduit par $\alpha = 1$. Dans ce second cas, qui renvoie à la célèbre formule de Bentham « le plus grand bonheur du plus grand nombre », tous les individus sont considérés égaux et ce, indépendamment de leur niveau initial d'utilité.

Le troisième cas caractérisé par $-\infty < \alpha < 1$ fait référence à une aversion pour l'inégalité. La fonction de justice accorde à une variation d'utilité un poids d'autant plus important que le niveau initial de cette utilité est faible. Comme le rappelle Wolfelsperger, plus la valeur de α est faible, plus on se rapproche d'une situation où seule est prise en compte la variation d'utilité de l'individu dont le niveau de bien-être initial est le plus faible. Ce troisième cas présente des similitudes avec l'approche du maximin de Rawls évoquée ci-dessus.

Ce raisonnement en termes d'aversion pour l'inégalité se retrouve également dans les travaux de Thisse (1994) concernant la localisation d'un établissement de service public. Il met en évidence le conflit entre équité et utilitarisme et propose de le résoudre en introduisant un paramètre α comme un exposant à la distance dans l'équation suivante :

$$G(s) = \sum_{i=1, m} n_i d^\alpha(x_i, s)$$

avec $d(x_i, s)$ la distance entre l'utilisateur et la localisation de l'établissement et n_i le nombre d'utilisateurs à distance x_i .

Pour Thisse, il revient au décideur public d'évaluer le paramètre α pour arbitrer entre efficacité et équité. Un paramètre égal à l'unité privilégie l'efficacité dans la mesure où le décideur accorde une même importance à toutes les variations d'utilité. À l'inverse, plus ce paramètre tend vers l'infini, plus le décideur se place dans une optique d'équité spatiale (correspondant au principe du maximin) en favorisant les zones les plus éloignées de l'établissement public.

L'importance des pondérations est également soulignée par Proost (2001). Il illustre, à partir d'une fonction de bien-être social, l'impact de différentes pondérations du surplus et ce, pour diverses catégories

sociales, sur des scénarios variés de politiques de transport et de financement et redistribution. Il évalue alors les effets de chacun des scénarios selon plusieurs critères tels que le changement parétien, l'efficacité ou l'équité.

La question de l'agrégation des utilités individuelles ne peut donc être tranchée en l'état. Diverses pistes sont ainsi présentées au décideur politique sans qu'une solution ne s'impose d'elle même.

En outre, rien n'est moins sûr qu'un indicateur unique convienne au débat sur l'acceptabilité d'une politique de transport : une certaine désagrégation est nécessaire pour identifier les gagnants et les perdants. Il semble possible, en cohérence avec le calcul économique, de calculer des variations de surplus selon des catégories spatiales (i.e. accessibilité, équité spatiale), sociales (i.e. équité verticale) ou d'usage des modes (i. e. équité horizontale). Ce calcul se fait bien entendu sous réserve d'hypothèses d'homogénéité : homogénéité des habitants d'une zone donnée pour l'équité spatiale, homogénéité des usagers d'un mode par rapport aux usagers d'un autre mode, etc.

4. CONCLUSION

Ici s'arrête provisoirement notre exploration des contributions potentielles du calcul économique au processus d'acceptabilité. Faut-il aller plus loin dans la recherche de cohérence théorique ? On trouvera par exemple dans un article de Foster et Neuberger (1974) une discussion détaillée sur les différences entre les mesures marshalliennes et hicksiennes de variations de bien-être. Il n'est pas sûr que ces subtilités intéressent le débat sur l'acceptabilité, dans lequel la remise en cause du calcul économique semble plus fondamentale.

Nous avons montré, à partir de la littérature, comment le calcul de variations d'accessibilité spatiale à partir d'un modèle de distribution spatiale de type gravitaire trouvait sa cohérence avec le calcul économique classique de surplus. Le mode de calcul de surplus, cohérent avec le calcul économique, a également été rappelé pour une classe particulière de modèles de choix discret qui respectent une spécification d'utilité « séparable-additive ». Ont ensuite été précisées les conditions dans lesquelles ces variations d'utilités pouvaient être comparables entre zones ou entre catégories socio-économiques.

L'illustration de notre propos par un exemple d'application, nous permet de conclure qu'il est possible de mener des analyses comparatives, en cohérence avec les principes du calcul économique, en termes de gagnants et de perdants potentiels face à un projet, et ce selon leur localisation ou selon leurs profils socio-économiques.

Rien n'interdit sur le plan théorique d'aller plus loin, vers une identification plus fine de ces gagnants et perdants potentiels, afin de préparer des mesures éventuelles de compensation. Cette analyse suppose toutefois :

- une meilleure capture des phénomènes de moyen et long termes découlant de modifications de l'offre de transport, donc le développement de modèles liant transport et usage des sols ;
- le développement de modèles de microsimulation permettant d'analyser finement les comportements d'activités et de déplacements d'individus hétérogènes dans l'espace construit ;
- des données statistiques fines géolocalisées, permettant de croiser les données économiques (dont les revenus pour lesquels certains paramètres statistiques sont rendus disponibles depuis peu) et les données de comportement de mobilité spatiale provenant des enquêtes déplacements.

Références

- Anderson S.P., de Palma A. et Thisse J.-F., *Discrete choice theory of product differentiation*, The MIT Press, 1992, 423 p.
- Ben Akiva M.E. et Lerman S., « Disaggregate travel and mobility choice models and measures of accessibility », in Hensher D.A. et Stopher P.R. (eds), *Behavioural travel modelling*, Croom Helm, London, 1979, p. 698-710.
- Ben Akiva M.E. et Lerman S., *Discrete Choice Analysis : Theory and Application to Predict Travel Demand*, The MIT Press, 1985.
- Ben-Akiva M. et Bowman J. L., « Integration of an activity-based model system and a residential location model », *Urban Studies*, Vol. 35, n° 7, 1998, p. 1131-1153.
- Bloy E., Bonnafous A., Cusset J.M. et Gérardin B., *Évaluer la politique des transports*, Economica, Paris, PUL, Lyon, 1976, 103 p.
- Bonnafous A. et Masson S., « Evaluation des politiques de transports et équité spatiale », *Revue d'Économie Régionale et Urbaine* n° 4, 2003, p. 547-572.
- Dupuit J., « De la mesure de l'utilité des travaux publics », *Annales des Ponts et Chaussées*, 2^e série, Mémoires et Documents, n° 116, t.VIII, 1844, p. 332-375.
- Foster C.D., Neuburger H.L.I., « The Ambiguity of the Consumer's Surplus Measure of Welfare Change », *Oxford Economic Papers*, vol 26, n° 1, 1974, p. 66-77.
- Hansen W.G., « How accessibility shapes land use », *Journal of the American Institute of Planners*, 25, 1959, p. 73-76.

- Koenig, G., « Théorie économique de l'accessibilité urbaine », *Revue Economique*, n° 2.1974.
- Mas-Colell A., Whinston M.D. et Green J.R., *Microeconomic Theory*, Oxford University Press, 1995, 981 p.
- Miller H.J., *Measuring Space-Time Accessibility Benefits Within Transportation Networks : Basic Theory and Computational Procedures*. Paper presented at the 37th annual meeting of the Western Regional Science Association, Monterey, California, 1998.
- Mc Fadden, « Modelling the Choice of Residential Location », in Karlqvist A., Lundqvist A., Snickars F. et Weibull J. (eds), *Spatial Interaction Theory and Planning Models*, North-Holland, 1978, p. 75-96.
- Neuberger H., « User benefit in the evaluation of transport and land use plans », *Journal of Transport Economics and Policy*, 5, 1971, p. 52-75.
- Poulit J., *Urbanisme et transport : les critères d'accessibilité et de développement urbain*. SETRA, Division urbaine, Ministère de l'Équipement, Paris, 1974, 55 p.
- Proost S., « Achieving Equity Through Urban Transport Pricing ? » CUPID-PROGRESS workshop, 2001. Article disponible sur www.transport-pricing.net/proost.doc.
- Raux C. et Souche S., « The acceptability of urban road pricing : A theoretical analysis applied to experience in Lyon », *Journal of Transport Economics and Policy*, Vol. 38, Part 2, 2004, p. 191-216.
- Rawls J., *A Theory of Justice*, Harvard University Press, 1971, 607 p.
- Small K.A. et Rosen, H.S., « Applied Welfare Economics with Discrete Choice Models », *Econometrica*, vol. 49, n° 1, 1981, p. 105-130.
- Thisse J.-F., « L'équité spatiale », in Auray J.P., Bailly A., Derycke P.H. et Huriot J.M. (éd.), *Encyclopédie d'économie spatiale*, Economica, Paris, 1994, p. 225-231
- Wilson A. G., *Entropy in urban and regional modelling*, Pion, London, 1970.
- Wolfelsperger A., *Économie publique*, Collection Thémis Economie, PUF, Paris, 1995.

CHAPITRE 11

L'ANALYSE DES PROJETS D'INFRASTRUCTURE DE TRANSPORT DANS UN CADRE D'ÉQUILIBRE GÉNÉRAL

Stef Proost¹

Dans l'analyse Coût-Bénéfice classique d'un projet d'infrastructure de transport² on se concentre le plus souvent sur l'estimation des effets directs sur le marché de transport. La simple somme des bénéfices pour les utilisateurs, moins les coûts du projet, moins les coûts externes, plus l'accroissement net des revenus publics sur le marché des transports est le point de départ pour une telle analyse. Théoriquement, cette approche n'est justifiée que moyennant des hypothèses héroïques sur l'étendue du projet et sur le fonctionnement du reste de l'économie : absence de distorsions de marché (taxes, effets externes, concurrence imparfaite) et une distribution équitable des revenus.

Dans ce chapitre on explore comment étendre l'approche classique pour mieux tenir compte de ces distorsions de marché et des problèmes de distribution des revenus. On procédera en quatre étapes. On commencera par préciser la nature des distorsions de marché et pourquoi celles-ci sont importantes. Ceci se fera à l'aide de quelques exemples. Le cœur de ce chapitre sera un ensemble de règles d'évaluation de projet qui sera plus complet que les règles classiques. On étendra progressivement le champ d'action des règles énoncées en utilisant des modèles économiques de plus en plus complexes. Ensuite on passera en revue brièvement la littérature scientifique et les méthodologies utilisées en France, aux Pays-Bas et au Royaume-Uni.

1. Center for Economic Studies, Faculty of Economics, KULeuven, Belgium

2. Notre approche se veut valable aussi bien pour les projets d'infrastructure que pour l'évaluation de réformes de tarification et de taxation.

1. LE CHOIX ENTRE LES MODÈLES D'ÉQUILIBRE GÉNÉRAL ET D'ÉQUILIBRE PARTIEL COMPLÉTÉ PAR DES PRIX FICTIFS

En principe un modèle d'équilibre général de l'économie, s'il contient une modélisation détaillée du secteur du transport, permet de répondre à la plupart des questions posées dans l'analyse d'un projet. Un tel modèle a deux avantages : il permet de calculer les effets d'activité induits dans le reste de l'économie et deuxièmement, il permet de tenir compte des distorsions de marché dans le reste de l'économie. Prenons par exemple un projet d'infrastructure qui fait baisser le coût du transport des matières premières d'une usine d'acier. Ceci peut donner lieu entre autre à une baisse du prix de l'acier, une demande accrue d'acier et une baisse de la demande de produits substitués comme l'aluminium. Les changements d'activité des différents secteurs engendrés par le nouveau projet peuvent être calculés à l'aide des matrices Input-Output contenues dans les modèles d'équilibre général. Cette information sur les changements d'activité est en soi utile pour estimer la demande de services de transport. Le deuxième avantage de tels modèles est de pouvoir estimer les effets induits sur les autres marchés, sur les plus ou moins values de recettes fiscales et sur l'utilisation qui en est faite.

Il se fait cependant que de tels modèles d'équilibre général sont très rares et s'ils existent, ils vont souvent être trop agrégés dans l'espace³ pour pouvoir donner des informations utiles sur la demande de transport. C'est la raison principale pour laquelle nous proposons dans ce chapitre une approche qui part d'une analyse classique du marché de transport (utilisant des modèles d'« équilibre partiel ») et corrige celle-ci à l'aide de prix fictifs qui tiennent compte des effets induits sur les marchés autres que le transport.

Ces prix fictifs peuvent être calculés avec des modèles d'équilibre général simplifiés. Tenir compte des effets induits sur les autres marchés est important aussi bien pour des petits que pour des grands projets et aussi bien pour des projets d'infrastructure que pour des grands projets de réforme de tarification ou de taxation dans le secteur des transports. En effet, chaque projet, aussi petit soit-il, a des effets secondaires sur les autres marchés qui doivent être pris en compte dans le calcul via des facteurs correctifs même si chacun de ces projets est trop petit pour être mesuré avec un modèle macro-économique classique au niveau d'un pays ou d'une région.

3. La plupart des modèles d'équilibre général multisectoriels sont des modèles nationaux.

2. POURQUOI DOIT-ON ALLER PLUS LOIN QU'UNE ANALYSE AU NIVEAU DU MARCHÉ DE TRANSPORT ?

Prenons un exemple analytique simple. Considérons une économie avec des individus identiques qui disposent d'un revenu R exogène et ne consomment que trois biens : un bien général acheté sur le marché (quantité x_1), un bien de transport (quantité x_2) et un bien produit à la maison (quantité x_3). Les biens 1 et 3 sont parfaitement substituables et donnent donc la même utilité par unité. Le coût marginal (exprimé en temps de travail) du bien marchand 1 est égal à 1, le coût marginal du bien 3 est égal à $1+a$. Produire les biens à la maison est par hypothèse moins efficace et ceci se reflète dans un paramètre $a > 0$.

Moyennant un coût d'investissement I par individu⁴, le prix de transport (par hypothèse égal au coût marginal) peut être réduit de p à $p-\Delta p$. Les individus ont une fonction d'utilité quasi-linéaire et une contrainte de budget de la forme suivante :

$$\begin{aligned} U &= x_1 + x_3 + v(x_2) \\ x_1 + (1+a)x_3 + px_2 &= R \end{aligned} \quad (1)$$

Cette fonction d'utilité nous facilite la vie parce qu'elle implique que la demande pour le bien de transport 2 sera indépendante du niveau de revenu, que tout revenu supplémentaire est dépensé au bien 1 ou 3 et que l'utilité marginale du revenu (dépensé au bien 1) est constante et égal à 1. Pour l'instant, le ménage ne consomme pas le bien 3 parce que le bien 1 lui coûte moins cher par unité d'utilité. Les fonctions de demande pour les biens 1, 3 et 2 peuvent être écrites comme :

$$\begin{aligned} x_1(R, p_2, p_3) &= R - p_2 x_2 \\ x_3(R, p_2, p_3) &= 0 \\ x_2 &= x_2(p_2) \end{aligned} \quad (2)$$

Nous voulons évaluer le projet d'investissement de transport et montrer l'importance des distorsions sur les marchés non transport pour cette évaluation. On commencera par l'évaluation du projet dans l'absence d'une telle distorsion.

Dans l'absence de distorsions sur le marché du bien 1, le prix de marché du bien 1 est égal au coût marginal ($p_1=1$). Le changement d'utilité de l'individu représentatif à la suite du projet d'investissement

4. Il s'agit ici d'une économie stationnaire où le coût d'investissement a la dimension d'un coût de location qui comprend l'annuité du coût du bien de capital et les coûts d'entretien.

est donné par l'expression (3). Cette expression a été obtenue en utilisant la définition de la fonction d'utilité (1) et des fonctions de demande (2). Le premier terme dans (3) représente le gain d'utilité réalisé sur le bien 2 parce que la consommation aura augmenté (ou diminué) à la suite de la diminution du coût et du prix de p à $p - \Delta p$. Le second terme dans (3) représente le gain d'utilité réalisé par la consommation accrue du bien 1. Cette consommation accrue est égale à la diminution nette des dépenses sur le bien 2 moins le coût d'investissement de transport payé par chaque ménage. La consommation du bien 3 reste égale à 0 parce qu'il reste plus intéressant pour les ménages de consommer le bien 1 plutôt que le bien 3.

$$du = [v(x_2(p - \Delta p)) - v(x_2(p))] + [px_2(p) - (p - \Delta p)x_2(p - \Delta p) - I] \quad (3)$$

Cette expression peut être réécrite dans la forme (4) en utilisant l'intégration par parties pour obtenir la formulation classique du surplus du consommateur sur le marché de transport utilisé par les praticiens de l'analyse coût-bénéfice.

$$du = \int_{p-\Delta p}^p x_2(p) dp - I \quad (4)$$

Dans l'absence de distorsions sur les autres biens, nous obtenons bien qu'il suffit d'analyser uniquement les changements sur le marché du transport. La consommation des autres biens dans l'économie va changer à la suite des effets revenu et prix mais les effets sur le bien-être de ces changements sont déjà inclus dans la mesure de surplus (4) utilisée. Cette expression est bien commode et est utilisée dans quasi toutes les analyses coûts-bénéfices.

Introduisons maintenant une distorsion sur le marché du bien 1. Le bien 1 « marchand » est frappé par une accise t de sorte que le prix du bien 1 devient égal à $1+t$ ⁵. Prenons une accise $t > a$ de telle sorte que tous les ménages préfèrent utiliser le bien produit à la maison plutôt que le bien marchand 1. Les coûts marginaux n'ont pas changé, il n'y a que les prix aux consommateurs qui ont changé. Le résultat sera une inefficacité au niveau de l'économie globale qui est égale au surcoût de production du bien 3 (« a ») fois la consommation du bien 3. Ceci peut être démontré aisément. Les fonctions de demande deviennent main-

5. Une distorsion est ici définie comme une différence entre le coût marginal et le prix au consommateur. Au lieu d'une taxe on peut introduire une marge monopolistique sur le bien 1. Au lieu des recettes fiscales on va alors redistribuer les profits à la population.

tenant :

$$\begin{aligned} x_1(R, p_2, p_3) &= 0 \\ x_3(R, p_2, p_3) &= \frac{R - p_2 x_2}{1 + a} \\ x_2(p_2) \end{aligned} \quad (5)$$

Refaisons maintenant l'évaluation du même investissement I. Nous ne devons pas tenir compte des recettes publiques provenant de la taxe parce qu'on se trouve dans le cas extrême où la consommation du bien 1 est tombée à zéro. L'expression (3) qui mesure l'avantage net pour la société de l'investissement de transport dans l'absence de distorsions devient maintenant :

$$du = [v(x_2(p - \Delta p)) - v(x_2(p))] + \frac{(p - \Delta p)x_2(p - \Delta p) - I - px_2(p)}{1 + a} \quad (6)$$

Le premier terme dans (6) représente encore toujours le gain d'utilité réalisé sur le bien 2, le second terme représente le gain d'utilité venant de la consommation accrue du bien 3. Ce dernier est égal à l'économie de dépense sur le bien 2 divisé par le prix du bien 3. Cette expression peut être réécrite de la même façon que (4) pour obtenir :

$$du = \int_{p - \Delta p}^p x_2(p) dp - I - \frac{a}{1 + a} dx_3 \quad (7)$$

Dans une économie avec des distorsions sous la forme d'une taxe, il ne suffit plus de mesurer le surplus sur le marché de transport pour connaître les effets d'un projet d'investissement. En effet, le projet de transport peut aggraver (diminuer si $dx_3 < 0$) les distorsions dans le reste de l'économie et le projet de transport aura dans ce cas un avantage net plus petit (grand). L'inefficacité consiste à consommer plus du bien 3 dont on sait qu'il est une fraction « a » plus coûteuse pour l'économie que le bien 1. L'inefficacité sera d'autant plus grande que le facteur de surcoût « a » qui permet d'échapper à la taxe, est grand et ceci sera d'autant plus vraisemblable que la taxe sur le bien 1 est élevée.

Un projet transport peut faire changer la consommation sur les autres marchés qui connaissent des distorsions pour trois raisons. D'abord parce que le revenu disponible pour ces autres biens change (financement de l'investissement par un prélèvement forfaitaire I sur le

revenu comme dans notre exemple ci dessus), deuxièmement parce qu'il y a des effets de substitution associés au changement de prix du bien transport, troisièmement, dans le cas où l'investissement serait financé par une augmentation d'une taxe non forfaitaire (p.ex. sur le travail), il y a un effet de prix supplémentaire. Dans notre exemple on pourrait imaginer que la taxe t sert à financer l'investissement de transport et que le résultat serait une plus grande substitution de biens marchands taxés par des biens non marchands, avec comme résultat une perte d'efficacité supplémentaire et donc un projet de transport moins intéressant.

Passons maintenant à un exemple plus concret. Une ligne de chemin de fer qui relie un village A et un centre urbain B est utilisée surtout par des navetteurs. Imaginons que la vitesse de la ligne est améliorée et que ceci amène les ménages à augmenter leur offre de travail de 10 % et que cette offre conduit à un emploi additionnel de 10 %. L'analyse classique de projet d'investissement considère comme bénéfiques de ce projet les gains en temps des passagers existants de la ligne plus le surplus réalisé par les 10 % de voyageurs additionnels. Ceci mène à une sous estimation des bénéfices du projet parce que l'emploi additionnel est un bien taxé à près de 100 %⁶. En termes de l'expression (7), le bien 1 est ici la quantité de travail dont les revenus sont dépensés en biens marchands, le bien 3 est ici la production domestique qui est trop grande à cause de la taxe sur le travail. Le projet de transport augmente ici le bénéfice direct pour le ménage du bien 1. Le bénéfice direct pour les ménages est inclus dans le surplus du marché de transport. L'autre moitié du bénéfice sociale, le bénéfice indirect (via l'apport de recettes fiscales) est typiquement oubliée dans une analyse limitée au marché de transport.

Poussons notre exemple encore plus loin et analysons le financement des projets de transport. Si le financement de l'amélioration de la ligne de chemin de fer se fait par une taxation accrue de l'emploi dans le pays, on peut s'attendre à une diminution de l'offre de travail et de l'emploi dans les autres régions. Cet effet sera différent si on utilise d'autres sources de financement public comme une diminution des transferts sociaux. Un même projet d'investissement dans le secteur du transport, engendrant le même surplus pour les utilisateurs mais financé par des taxes différentes, aura donc un bénéfice net différent pour la collectivité.

Ces exemples simples nous ont montré qu'un projet de transport doit tenir compte des effets induits sur d'autres marchés quand ces

6. Le total des taxes directes, indirectes et contributions à la sécurité sociale représentent près de 100 % du salaire net perçu par le salarié ; le coût salarial payé par l'employeur est ainsi près du double dudit salaire net.

marchés ont des distorsions. Les canaux par lesquels les autres marchés sont affectés sont multiples : des effets de substitution par les prix, des effets de revenu ou encore des augmentations de taxes nécessaires pour financer le projet d'infrastructure.

Les distorsions du marché recouvrent en général toutes les différences entre la volonté marginale à payer du consommateur (égal au prix au consommateur) et le coût d'opportunité du bien. Il s'agit de taxes mais aussi de marges monopolistiques. Les déséquilibres de marché (ou un rationnement est nécessaire comme par exemple pour le chômage involontaire) peuvent aussi être traités dans un cadre proche de celui qu'on propose ici.

3. UN PREMIER MODÈLE STATIQUE SIMPLE

Le premier modèle simple qu'on utilise pour dériver des règles d'analyse coût- bénéfice est une généralisation de Mayeres et Proost (1997, 2001) et de Calthrop, De Borger et Proost (2007) et a été utilisé pour un projet de recherche européenne (Proost *et al.*, 2007). Il s'agit d'un modèle à une période où le transport de personnes et de marchandises sont intégrés dans un modèle d'équilibre général. Dans ce modèle statique, la capacité de transport peut être louée à un prix donné. Notre objectif est double. Premièrement il est de dériver des expressions algébriques générales qui montrent la structure des effets à considérer. Deuxièmement on va déterminer quels effets peuvent être calculés à l'aide de modèles de transport et quels effets doivent être importés à partir de modèles économiques plus généraux.

Il y a N individus $i=1, \dots, N$ qui diffèrent dans leur productivité horaire de travail e_i . Cette productivité n'est pas observable par le planificateur et est la source de l'inégalité de la distribution des revenus⁷. Chaque individu maximise une fonction d'utilité $U(C, T, D, \ell)$ définie sur quatre biens. Il y a un bien C (le numéraire non taxé) qui peut être produit sans avoir recours au facteur transport de marchandises, un bien de consommation D qui requiert une grande quantité de transport de marchandises, un bien T représentant le transport de personnes et finalement le loisir ℓ . Le transport de personnes et de marchandises produisent une externalité négative : la congestion. On utilise donc un modèle avec un seul mode de transport, même une seule route sur lequel il y a un flux de camions et de voitures distribué de manière uniforme dans le temps. Ajouter plus de modes de trans-

7. Ceci est la procédure standard pour introduire l'inégalité des revenus (voir Mirrlees, 1971).

port, plus de localisations et d'autres types d'externalités comme la pollution environnementale et le bruit est possible mais notre modèle perdrait en clarté. Dans ce modèle simple, il n'y a pas d'incertitude et la localisation de toutes les activités est fixée.

L'individu fait face à deux contraintes. La première est une contrainte budgétaire : $C_1 + q_T T_i + q_D D_i = w e_i L_i + G$ où G représente le subside forfaitaire de l'État et où L_i représente l'offre de travail. La deuxième est une contrainte de temps $L_i + \ell_i + T_i \phi(\sum_{i=1}^N T_i + F, I) = \bar{L}$. La fonction de congestion $\phi(\sum_{i=1}^N T_i + F, I)$ nous donne le temps nécessaire par unité de transport T ; il dépend du volume du transport de passagers ($\sum_{i=1}^N T_i$), du volume total F de fret mais aussi de la capacité de l'infrastructure de transport I ⁸. Afin de simplifier au maximum la présentation, nous avons aussi supposé que la contribution à la congestion d'une voiture et d'un camion sont identiques.

Passons à présent à la structure de la production. Elle est de nouveau simplifiée afin d'obtenir des expressions algébriques simples. Nous utilisons une fonction de production agrégée linéaire (8) qui relie la production du transport de personnes (T) et de marchandises (F), la production des biens de consommation C et D , la production d'un bien intermédiaire (X) ainsi que la capacité du réseau de transport I à un input primaire dans la production : le travail mesuré en unités à productivité identique dont le total est L . Les unités utilisées pour les quantités des biens ont été ajustées de telle sorte que tous les coefficients dans l'expression (8) sont égaux à 1 :

$$\sum_{i=1}^N (T_i + C_i) + F + X + I \leq \sum_{i=1}^N e_i L_i \quad (8)$$

Le membre de droite de (8) nous donne le nombre d'unités équivalentes de travail nécessaire pour produire les différents biens qui se trouvent dans le membre de gauche. Le bien de consommation D est le seul à nécessiter un input de transport de marchandises. Nous supposons, pour un niveau de congestion ϕ donné, que cette production a des rendements d'échelle constants ce qui donne la fonction (9).

$$\sum_{i=1}^N D_i = S(F, X; \phi) \quad (9)$$

8. L'interprétation naturelle du bien transport est l'utilisation d'une route avec de la congestion. On peut aussi utiliser ce cadre théorique pour les transports publics où la qualité d'utilisation diminue quand il y a plus d'utilisateurs.

Plus le niveau de la congestion $\phi(\sum_{i=1}^N T_i + F, I)$ est élevé, plus on aura besoin d'autres inputs (F,X) pour produire une unité de D. Finalement, on suppose la concurrence parfaite dans l'ensemble du secteur de la production. Sous ces hypothèses, les prix à la production d'équilibre pour les biens T, C, F, X, I et L sont tous égaux à l'unité. Le coût marginal du bien D, par contre, n'est pas constante et dépendra du niveau de la congestion.

Un planificateur bienveillant a pour objectif de maximiser la fonction de bien-être W, donné en (10), définie ici comme la simple somme des utilités individuelles. Comme toutes les fonctions d'utilité sont supposées identiques, la concavité de la fonction d'utilité déterminera l'utilité marginale sociale relative (ou poids distributif) λ_i d'une unité de revenu (ou d'une unité du bien C non taxé dont le prix est égal à 1) pour un individu i donné.

$$W = \sum_{i=1}^N U(C_i, T_i, D_i, l_i) \quad (10)$$

où $\frac{\partial W}{\partial U} \frac{\partial U}{\partial C_i} = \lambda_i$

Le planificateur dispose de deux types d'instruments pour y arriver : la capacité du réseau routier I et différentes taxes (taxe ou subvention uniforme G, taxe sur le travail τ_L , taxe sur le transport de passagers τ_T et une taxe sur le transport de marchandises τ_F)⁹. Le budget du gouvernement planificateur se présente de la façon suivante (où le coût de la capacité¹⁰ équivaut à N fois l'unité) :

$$\left(\tau_T \sum_{i=1}^N T_i + \tau_L \sum_{i=1}^N e_i L_i + [\tau_D + \tau_F F_{ND}] \sum_{i=1}^N D_i \right) - I N = NG \quad (11)$$

où F_{ND} représente la quantité de transport de marchandises nécessaire par unité produite du bien D et $F_{ND} \sum_{i=1}^N D_i = F$ représente la demande totale de transport de marchandises dans l'économie.

9. Toutes les taxes sont du type « accise » et s'ajoutent donc au prix à la production pour obtenir le prix à la consommation.

10. On présente cette contrainte sans option d'emprunt. Cette approche est justifiée dans le cadre d'une économie stationnaire. Dans une économie stationnaire, le coût de capacité a le caractère d'un coût de location comprenant l'annuité, l'entretien et les intérêts intercalaires de l'infrastructure.

Nous allons utiliser ce modèle pour dériver le bénéfice net de bien-être d'une petite augmentation de la capacité de l'infrastructure de transport dI financée par les différents types de taxes prévus par notre modèle de base. On utilise une approche de réforme marginale de la taxation¹¹ partant de niveaux des taxes qui ne sont pas nécessairement optimaux. On commence avec la taxe uniforme G, ensuite la taxe sur le travail et on finira par étudier le financement par une taxe sur le transport.

3.1. Extension de l'infrastructure financée par une taxe uniforme G sur tous les individus

Utilisant les fonctions d'utilité indirectes dans la fonction de bien-être social (10) nous donne :

$$W = \sum_{i=1}^N V^i \left[(1 - \tau_L) e_i, (1 + \tau_D) c_D(\tau_F, \phi), \phi \left(\sum_{i=1}^N T_i + F, I \right), G \right] \quad (12)$$

où c_D représente le coût à la production du bien D, le bien qui requiert beaucoup de transport de marchandises ; ce coût est lui-même une fonction de la taxe sur le transport de marchandises et du niveau de congestion

L'effet sur le bien-être d'un investissement en capacité de transport dI devient (différentiant (12) et en utilisant $\frac{\gamma_i}{\lambda_i}$ qui est le rapport entre l'utilité marginale du temps et l'utilité marginale d'une unité de revenu et représente donc la valeur subjective du temps pour l'individu i) :

$$\frac{1}{\lambda^*} \frac{dW}{dI} = - \sum_{i=1}^N \frac{\lambda_i}{\lambda^*} \left(D_i \frac{\delta c_D}{\delta \phi} + \frac{\gamma_i}{\lambda_i} T_i \right) \frac{d\phi}{dI} + N \frac{dG}{dI} - \sum_{i=1}^N \frac{\lambda_i}{\lambda^*} \left(D_i \frac{\delta c_D}{\delta \phi} + \frac{\gamma_i}{\lambda_i} T_i \right) \frac{d\phi}{dG} \frac{dG}{dI} \quad (13)$$

où $\lambda^* = \sum_{i=1}^N \frac{\lambda_i}{N}$

Dans (13), le paramètre $\frac{\lambda_i}{\lambda^*}$ est connu comme l'utilité marginale sociale d'une unité de revenu pour le ménage i. Plus grande est l'aversion à l'inégalité, plus grande sera la concavité de la fonction d'utilité et plus grande sera ce paramètre pour un individu pauvre. Afin

11. L'approche réforme de la taxation ne tient que pour des petits changements dans les taxes et la capacité (voire Guésnerie, 1977). L'avantage d'une telle réforme est que beaucoup moins d'information est nécessaire pour juger une réforme. Parce que le secteur de transport est petit par rapport à l'ensemble de l'économie, les changements de taxation générale seront petits.

d'arriver à une expression interprétable il nous faut encore calculer $\frac{dG}{dI}$ en utilisant la contrainte budgétaire du gouvernement :

$$R(\tau_T, \tau_L, \tau_D, \tau_F, \tau_X, G, \phi) = G \quad (14)$$

$$\frac{dG}{dI} = - \frac{1 - \frac{\partial R}{\partial \phi} \frac{d\phi}{dI}}{1 - \frac{\partial R}{\partial G} \Big|_{\phi} - \frac{\partial R}{\partial \phi} \frac{d\phi}{dG}} \quad (15)$$

Après substitution dans (13) nous obtenons une expression qui devient facile à interpréter quand on émet l'hypothèse $\frac{d\phi}{dG} = 0$, ceci nous donne :

$$\begin{aligned} \frac{1}{\lambda^*} \frac{dW}{dI} = & \\ & - \sum_{i=1}^N \frac{\lambda_i}{\lambda^*} (D_i \frac{\delta c_D}{\delta \phi} + \frac{\gamma_i}{\lambda_i} T_i) \frac{d\phi}{dI} \\ & - N \Gamma_G \\ & + \tau_L \frac{\partial L}{\partial \phi} \frac{d\phi}{dI} \Gamma_G \\ & + \tau_T \frac{\partial T}{\partial \phi} \frac{d\phi}{dI} \Gamma_G + \tau_F \frac{\partial F}{\partial \phi} \frac{d\phi}{dI} \Gamma_G \end{aligned} \quad (16)$$

$$\text{où } \Gamma_G = \frac{1}{1 - \frac{\partial R}{\partial G} \Big|_{\phi}}$$

Le premier terme du membre de droite représente le bénéfice marginal direct pour le transport de marchandises et pour le transport de personnes dû à la réduction du temps de transport. Les gains au niveau des coûts de transport des marchandises vont affecter les individus via le prix à payer pour les biens D.

Le terme $\gamma_i/\lambda_i T_i$ représente, pour les ménages, les économies en coût de temps associées à l'investissement. Les bénéfices directs pour le transport de marchandises et de personnes ont été pondérés par un poids distributif égal à l'utilité marginale sociale relative d'une unité de revenu.

Le deuxième terme représente le coût marginal des fonds publics collectés via une augmentation de la taxe uniforme G et utilisés pour

un projet public qui n'a, par définition, aucun effet sur la consommation de biens taxés.

Le troisième terme représente l'effet de retour d'une diminution de la congestion sur les recettes fiscales de la taxe sur le travail. Parce que cette augmentation des recettes fiscales permet de réduire la taxe uniforme, on retrouve comme facteur de pondération le coût marginal des fonds publics collectés par une taxe uniforme Γ_G .

Le quatrième et cinquième terme représentent l'effet de retour de la diminution de la congestion sur les recettes fiscales des taxes sur le transport.

Nous pouvons résumer les résultats obtenus dans (16) en soulignant tous les termes qui peuvent être calculés à l'aide d'un modèle de transport « partiel » :

$$dW = \sum_i \text{poids.distrib.} [\text{bénéfice.transport.de.personnes}_i + \text{bénéfice.fret}_i] [NRC_G] - \Gamma_G [\text{coût.unit.capacité} - \text{recettes.fisc.induites.transport}_{NRC_G} - \text{recettes.fisc.induites.travail}_{NRC_G}] \quad (17)$$

Où le coût marginal des fonds publics pour une taxe uniforme est approximé par¹² (où R représente les recettes fiscales totales et où le niveau de congestion est constant au niveau $\bar{\phi}$) :

$$\Gamma_G = \frac{1}{1 - \frac{\partial R}{\partial G} \Big|_{\bar{\phi}}} \quad (18)$$

NRC_G représente la « réduction nette de la congestion » qui résulte de l'augmentation de la capacité d'une unité. Cette réduction nette est calculée en tenant compte de trois effets : l'effet direct d'une augmentation de la capacité pour un volume de trafic constant, l'augmentation de la congestion qui résulte du trafic additionnel attiré par la diminution des coûts en temps et troisièmement, le changement de la congestion qui résulte des changements de la consommation de transport résultants des effets de revenus associés à l'augmentation de la taxe uniforme. C'est à cause de cette augmentation de la taxe uniforme que l'effet NRC devra être indexé par le type de taxe utilisé pour financer l'extension de l'infrastructure.

12. Ceci est une approximation parce qu'on néglige l'effet direct, à infrastructure donné du changement de la taxe forfaitaire sur le volume total de transport, sur la congestion et sur les recettes fiscales. Cet effet est probablement petit et nous permet d'utiliser des estimations du coût marginal des fonds publics publiés dans la littérature. Nous utilisons la même simplification pour le financement par une taxe sur le travail. Nous n'utilisons pas cette hypothèse quand l'investissement est financé par des taxes transport.

Pour des degrés élevés de l'aversion à l'inégalité (concavité forte des fonctions d'utilité dans la fonction de bien être W (10)) la pondération sera importante pour les individus pauvres et ceci peut fortement augmenter les bénéfices sociaux d'un projet de transport.

La seconde ligne de l'expression des coûts et bénéfices est égale au coût des fonds publics (pour une taxe uniforme) multiplié par les besoins net de financement. Dans cette seconde ligne, il n'y a pas de poids distributifs parce que, dans ce cas, tous les individus paient la même taxe.

Si on peut croire la plupart des études¹³ le coût des fonds publics est plus petit pour une taxe uniforme que pour une taxe sur le travail et pourrait même être inférieur à l'unité.

Dans l'expression (17) nous avons souligné tous les termes qui peuvent être calculés à l'aide d'un modèle de transport. Une approche du type équilibre général nous a amené à ajouter les éléments suivants :

Les poids distributifs qu'on peut estimer à partir d'autres choix de politique de taxation

Les recettes fiscales de l'impôt sur le revenu induites par le projet – ce terme peut être d'autant plus important pour les projets qui affectent les conditions de transport des navetteurs.

Le coût marginal des fonds publics d'une taxe uniforme.

3.2. *Investissement financé par une taxe proportionnelle sur le travail*

L'effet sur le bien-être d'une petite augmentation de la capacité de l'infrastructure financé par une augmentation de la taxe sur le travail devient maintenant (où les termes soulignés sont calculables à l'aide de modèles de transport du type équilibre partiel) :

$$dW = \sum_i \text{poids.distrib}_i \left[\underline{\text{bénéfice.passagers}_i} + \underline{\text{bénéfice.fret}_i} \right] \left[\underline{NRC_{\tau_t}} \right] \quad (19)$$

$$- \Gamma_{\tau_t} \left(\frac{\sum_i \frac{\lambda_i}{\lambda^*} e_i L_i}{\sum_i e_i L_i} \right) \left[\underline{\text{coût.unit.cap}} - \underline{\text{rec.fisc.ind.transport}_{NRC_{\tau_t}}} - \underline{\text{rec.fisc.ind.travail}_{NRC_{\tau_t}}} \right]$$

13. On peut consulter Snow & Warren (1996) qui font la synthèse de différentes études. Dans une étude pour les différents pays de l'OCDE, Kleven & Kreiner (2003) obtiennent pour la France un coût des fonds publics de l'ordre de 1.08 pour une taxe du type uniforme et de 1.57 pour une taxe sur les revenus du travail.

Le coût marginal des fonds publics collectés par une augmentation de la taxe proportionnelle sur le travail est égal à :

$$\Gamma_{\tau_L} = \frac{1}{1 + \frac{\tau_L}{\sum_i e_i L_i} \frac{\partial(\sum_i e_i L_i)}{\partial \tau_L}} \Bigg|_{\bar{\phi}} \quad (20)$$

Les bénéfices de l'extension de la capacité par unité de congestion réduite sont identiques aux bénéfices calculés pour le cas où une taxe uniforme est utilisée. La Réduction Nette de la Congestion (NRC) peut cependant être différente à cause d'autres effets de substitution et de revenu.

La partie relative au financement de l'extension de capacité est ici pondérée à l'aide de poids distributifs parce qu'il est maintenant important de savoir qui paie les taxes additionnelles sur le travail. Le coût des fonds publics dépend maintenant de la somme des élasticité de l'offre de travail. On s'attend en principe à un coût des fonds publics Γ_{τ_L} plus grand que un parce que l'augmentation de la taxe sur le travail va en général diminuer l'offre de travail¹⁴.

3.3. Investissement financé par une augmentation des taxes sur le transport de passagers ou de fret

L'effet sur le bien-être d'une petite augmentation de la capacité de transport devient maintenant :

$$dW = \sum_i \text{poids.distrib}_i \left[\text{bénéfice.passagers}_i + \text{bénéfice.fret}_i \right] \left[\text{NRC}_{\tau_T} \right] \quad (21)$$

$$- \Gamma_{\tau_L} \left(\frac{\sum_i \lambda_i T_i}{\sum_i \lambda^* T_i} \right) \left[\text{coût.unit.cap} - \text{rec.fisc.ind.transport}_{\text{NRC}_{\tau_T}} - \text{rec.fisc.ind.travail}_{\text{NRC}_{\tau_T}} \right]$$

14. Nous supposons ici que l'élasticité prix de l'offre de travail est positive. Il s'agit de l'élasticité agrégée au niveau des ménages qui tient compte des heures prestées mais aussi des décisions de participation. Il y a un certain consensus à ce sujet dans la littérature. Pour une revue de la littérature de l'élasticité prix de l'offre de travail voir Blundell et McCurdy (1999).

Où le coût marginal des fonds publics collectés à l'aide d'une augmentation de la taxe sur le transport de passagers est approximé par :

$$\Gamma_{\tau_r} = \frac{1}{1 + \frac{\tau_r}{\sum_i T_i} \frac{d(\sum_i T_i)}{d\tau_r}} \quad (22)$$

Les bénéfices d'une extension de la capacité de transport sont identiques aux cas précédents par unité de congestion réduite mais la réduction de la congestion sera beaucoup plus grande pour la même augmentation de la capacité parce qu'une taxe plus élevée sur le transport de personnes réduit aussi la congestion d'une façon directe.

La partie de l'expression relative au financement de la capacité est de nouveau pondérée par des poids distributifs parce que les paiements par les individus sont proportionnels à l'utilisation et celle ci peut être très différente pour les ménages.. Le coût marginal des fonds publics est maintenant une fonction de l'effet direct de la taxe sur le transport de personnes plus l'effet indirect de la taxe sur le niveau de la congestion et du nombre de voyages¹⁵. Ce coût marginal peut être calculé à l'aide d'un modèle de transport.

On peut refaire le même exercice pour la taxe sur le transport de marchandises. La différence majeure sera que les bénéfices proviennent d'un changement des prix des biens intensifs en transport de marchandises. Le coût des fonds publics sera lui aussi différent parce que les élasticités aux prix du transport de marchandises seront différentes.

3.4. *Le coût marginal des fonds publics « pur ».*

Il existe beaucoup de définitions et d'estimations pour le coût marginal des fonds publics collecté par une certaine taxe. La définition la plus courante nous dit que c'est le coût en termes d'efficacité de collecter un euro de recettes fiscales supplémentaires en augmentant la taxe pourvu que l'utilisation de la recette supplémentaire (pour des biens publics) n'affecte pas la consommation de biens taxés.

Si nous pouvons négliger les effets secondaires d'une taxe uniforme ou d'une taxe proportionnelle du travail sur le niveau de congestion

15. Ceci est écrit comme une dérivée totale mais la capacité est gardée inchangée.

nous pouvons utiliser un fonds public « pur » pour (18) et (20). Kleven & Kreiner (2003) donnent un ordre de grandeur pour le coût des fonds publics « pur » qui tient compte des effets sur le nombre d'heures de travail et sur le degré d'activité de la population. Ils tiennent aussi compte des effets du degré d'activité sur les dépenses de la sécurité sociale et considèrent la taxe indirecte moyenne aussi comme une taxe proportionnelle sur le travail.

Tableau 1.1

**Exemples de coûts des fonds publics « purs » utilisés pour l'analyse
Coût- Bénéfice**

	France	Royaume Uni
Taxe uniforme	1,08	1,09
Taxe proportionnelle sur l'emploi	1,57	1,37

Source : Kleven & Kreiner (2003)

La valeur 1,57 signifie que chaque euro qui doit être financé par des taxes sur le travail a un coût économique de 1,57 euro. Les valeurs dans ce tableau illustrent clairement l'importance du moyen de financement pour le résultat final de l'analyse coût-bénéfice.

4. RÈGLES D'INVESTISSEMENT DANS UNE ÉCONOMIE CROISSANTE

Jusqu'ici nous avons raisonné dans le cadre d'une économie statique qui se répète indéfiniment, qui n'a pas de taxes sur le capital et qui n'a pas d'incertitude. Le taux d'intérêt était implicite dans le coût de capacité.

Aussi longtemps qu'on néglige le risque et l'incertitude, nous pouvons étendre les règles dérivées pour une économie statique à une économie croissante. Nous suivons ici Liu (2003) ce qui nous permet de définir une règle d'investissement pour une économie avec une taxe sur le capital et une taxe proportionnelle sur le travail et des individus identiques qui ont une durée de vie infinie.

Un projet d'infrastructure peut être défini comme une série de coûts d'investissement et d'entretien $\Delta I_t ; t=1, \dots, \infty$, et une série de bénéfices pour les ménages $\Delta B_t ; t=1, \dots, \infty$ ainsi qu'une série de recettes

fiscales induites ΔR_t ; $t=1, \dots, \infty$. La règle d'investissement devient maintenant :

$$dW = \sum_t \frac{\Delta B}{(1+r_n)^t} - MCF_{\tau_L} \sum_t \frac{\Delta I_t - \Delta R_t}{(1+r_g)^t}$$

ou $r_n = r_g(1 - \tau_K)$

$$MCF_{\tau_L} = \frac{\sum_t \frac{w_t L_t}{(1+r_n)^t}}{\frac{\partial \left[\sum_t \frac{R_t}{(1+r_g)^t} \right]}{\partial \tau_L}} \quad (23)$$

ou τ_K est la taxe sur le revenu de capital.

Dans cette règle, les bénéfices sont actualisés au taux d'intérêt des consommateurs (après taxe) mais le gouvernement utilise un taux d'intérêt avant taxes qui est égal au taux de rentabilité du capital dans l'économie. Cette procédure est plus rigoureuse que l'utilisation d'une moyenne pondérée du taux avant et après taxes comme taux d'actualisation social.

L'extension de ce cadre à une économie avec risque et incertitude n'est pas facile parce que ces deux facteurs sont difficiles à intégrer dans une analyse d'équilibre général.

5. AUTRES EXTENSIONS DU MODÈLE DE BASE

On peut considérer d'autres extensions du modèle de base. On se contente de les commenter brièvement.

Premièrement, un traitement explicite du trafic des navetteurs étant donné que ceci a un effet direct sur le marché de travail. On peut consulter Parry et Bento (2002) pour un modèle intéressant qui montre qu'il reste intéressant de tarifier le trafic navetteurs au coût marginal social si le produit des tarifs est utilisé pour réduire la taxe sur le travail.

Deuxièmement, plusieurs modes de transport où certains modes ne sont pas tarifés correctement. Ceci est généralement traité à l'aide d'un modèle d'équilibre partiel qui intègre les différents marchés de transport. On peut étendre notre modèle de base à plusieurs marchés de transport.

Troisièmement, considérer d'autres imperfections de certains marchés de transport (concurrence monopolistique dans les opérateurs ferrés ou aériens). On peut utiliser notre cadre d'analyse pour ce

problème si les marges monopolistiques sont constantes et si la distribution des rentes monopolistiques est donnée (cf. note de bas de page 5). Si la rente monopolistique revient à l'État, on peut l'assimiler à une taxe. Une analyse d'équilibre général avec de la congestion où le comportement monopolistique (la marge monopolistique) est endogène est moins facile (voir de Palma et Proost, 2006).

Quatrièmement, on peut ajouter d'autres types de taxes et de subventions. Notre cadre peut les traiter avec cependant une perte de transparence dans les résultats.

Finalement, la modélisation explicite des effets externes autres que la congestion (pollution de l'air, etc.). La façon facile de le faire est de supposer que ces externalités n'affectent pas la consommation de biens taxés. Dans ce cas il suffit de soustraire les coûts externes du bien être social. Une façon plus correcte mais plus complexe est de, comme on l'a fait pour la congestion, intégrer dans le modèle les réactions des consommateurs aux dommages externes. Par exemple un accroissement de la pollution atmosphérique affecte la santé, et donc réduit l'offre de travail et les recettes fiscales de la même façon que la congestion affectait le volume de transport et l'offre de travail dans notre modèle (voir Williams, 2003, pour un exemple).

6. REVUE DE LA LITTÉRATURE

C'est en France que s'est développée une tradition du calcul économique, à partir des travaux pionniers de Jules Dupuit (1844). Lesourne, Boiteux et bien d'autres grands économistes français ont été les pionniers, pendant les années 50 et 60, de l'analyse coût-bénéfice avec une dimension d'équilibre général. Il est à ce sujet utile de consulter l'ouvrage de Lesourne (1964) qui contient beaucoup d'études de cas. Dans la tradition anglo-saxonne, il y a eu Harberger (1964) qui est bien connu pour le triangle de distorsion sur un marché mais qui a, dans ses travaux, aussi proposé une formule pour tenir compte des effets sur les distorsions sur les autres marchés.

La problématique a été reposée en termes théoriques pour les investissements publics (Atkinson et Stern, 1974 ; King, 1986 ; Sandmo, 1998) et pour le cas des effets externes (Ballard et Medema, 1993 et Bovenberg et Van der Ploeg, 1994) mais son application au secteur des transports est encore en développement (Parry et Bento, 2002 ; Mayeres et Proost, 2001).

Il reste une grande différence entre la littérature théorique et les manuels d'analyse coûts-bénéfices utilisés par les différentes adminis-

trations. La Banque européenne d'investissement recommande d'étudier plus en détail la problématique du coût des fonds publics mais ne fait pas de proposition concrète (voir Railpag, 2005). En Angleterre il y a eu récemment une étude approfondie (« Sactra » report) de la méthodologie de l'analyse Coûts Bénéfices (Myles, 1999) mais la question du coût des fonds publics et des interactions avec les autres distorsions dans l'économie n'a pas été résolue. Aux Pays-Bas il y a aussi eu une analyse approfondie de la méthodologie de l'analyse coûts bénéfices (Eijgenraam *et al*, 2000), cette analyse ne donne pas des recommandations concrètes sur les procédures à adopter pour le coût des fonds publics. Néanmoins, le Bureau du Plan néerlandais a développé une tradition d'analyse de projets d'infrastructure publique dans laquelle elle inclut une analyse d'équilibre générale.

Références

- Atkinson A.B. et Stern N.H., « Pigou, taxation and public goods », *Review of Economic Studies*, 41, 1974, p. 119-128.
- Ballard C.L. et Medema S.G., « The marginal efficiency effects of taxes and subsidies in the presence of externalities : a computational general equilibrium approach », *Journal of Public Economics*, 52, 1993, p. 199-216.
- Blundell R. et MacCurdy T., « Labour Supply : A review of alternative approaches », in Ashenfelter O. et Card D. (eds), *Handbook of Labor Economics*, North Holland, New York, 1999, p. 1559-1695.
- Bovenberg A.L. et van der Ploeg F., « Environmental policy, public finance and the labour market in a second-best world », *Journal of Public Economics*, 55, 1994, p. 349-390.
- Bovenberg, A.L. et Goulder L., « Optimal Environmental Taxation in the Presence of Other Taxes : General-Equilibrium Analysis », *American Economic Review*, 86(4), 1996, p. 985-1000.
- Calthrop E., De Borger B. et Proost S., « Tax Reform for dirty intermediate inputs », 2004, <http://www.econ.kuleuven.be/ete/downloads/ETE-WP-2003-02.PDF>
- De Palma A. et Proost S. (2006), « Imperfect competition and the city », *Journal of Urban Economics*, 60, 185-209.
- Dreze J. et Stern N.H., « The theory of cost benefit analysis », *Handbook of public Economics* (Auerbach and Feldstein ed), North Holland, 1987.
- Dupuit J., « De la mesure de l'utilité des travaux publics », *Annales des Ponts et Chaussées*, vol. 8, 1844.

- Eijgenraam C., Koopmans C., Tang P. et Venster A., « Evaluatie van infrastructuurprojecten », Leidraad voor kosten-baten analyse, Ministerie van Verkeer en Waterstaat en Ministerie van Economische Zaken, Den Haag, 2000.
- Fujita M. et Thisse J.-F., *Economics of Agglomeration*, Cambridge University Press, 2002.
- Guesnerie R., « On the direction of tax reform », *Journal of Public Economics*, 7, 1977, p. 179-202.
- Harberger A., « The measurement of waste », *American Economic Review*, papers and proceedings, 54, 1964, p. 58-76.
- King M., « A Pigouvian rule for the optimal provision of public goods », *Journal of Public Economics* 30, 1986, p. 273-291.
- Kleven H.J. et Kreiner C.T., « The marginal cost of public funds in OECD countries : hours of work versus Labor force participation », CESifo Working paper n° 935, 2003.
- Lesourne J., *Le calcul économique*, Dunod, Paris, 1964.
- Liu L., « A marginal cost of funds approach to multi-period public project evaluation : implications for the social discount rate », *Journal of Public Economics*, 87, 2003, p. 1707-1718.
- Mayeres I. et Proost S., « Optimal Tax and Investment Rules for Congestion Type of Externalities », *Scandinavian Journal of Economics*, 99 (2), 1997, p. 261-279.
- Mayeres I. et Proost S., « Marginal tax reform, externalities and income distribution », *Journal of Public Economics*, 79, 2001, p. 343-363.
- Mirlees J.A., « An exploration in the theory of optimum income taxation », *Review of Economic Studies* 38(114), 1971, p.175-208.
- Myles D., « Taxation, Economic Growth and the double dividend », report to the SACTRA group, Department of the Environment, Transport and the Regions, 1999.
- Parry I. et Bento A., « Revenue recycling and the welfare effects of road pricing », *Scandinavian Journal of Economics*, 103(4), 2002, p. 645-671.
- Pigou A.C., *The Economics of Welfare*, London, Macmillan and Co., Ltd, 1920.
- Proost S., De Berper B., Koskenoja P. (2007), « Public Finance Aspects of transport charging and investment », Chap. 3 dans A. de Palma, Lindsey R., Proost S., « Investment and the use of tax and toll revenues in the Transport Sector », Elsevier.
- Railpag, « Orientations pour l'évaluation des projets ferroviaires », brochure publiée par la Commission Européenne et la Banque Européenne d'investissement, 2005.

Sandmo A., « Redistribution and the marginal cost of public funds », *Journal of Public Economics*, 70, 1998, p. 365-382.

Snow A. et Warren R.S., « The marginal welfare cost of public funds : theory and estimates », *Journal of Public Economics*, 61, 1996, p. 289-305.

Williams R.C., « Health effects and optimal environmental taxes », *Journal of Public Economics*, 87, 2003, p. 323-335.

CHAPITRE 12

ÉQUITÉ, EFFICACITÉ ET ACCEPTABILITÉ DANS LA LOCALISATION DES ÉQUIPEMENTS COLLECTIFS

Jacques-François Thisse¹

Je souhaiterais commencer ce chapitre² par trois remarques qui me semblent importantes pour la suite de l'exposé et même, peut-être, pour l'ensemble de la réflexion relative au thème général de cet ouvrage.

Premièrement, dans toutes les questions portant sur la localisation d'équipements collectifs ou d'infrastructures, on retrouve la même tension entre, d'une part, les coûts d'installation et de fonctionnement des équipements et, d'autre part, les coûts d'accès des usagers à ces mêmes équipements³. La prise en compte du seul facteur financier conduit, en effet, à construire le plus petit nombre d'installations compatible avec un niveau d'encombrement – plus ou moins – admissible pour les usagers ; au contraire, le souci de minimiser les seuls coûts d'accès incite à accroître le nombre d'installations jusqu'à ce que le service considéré soit mis à la disposition des usagers à proximité de leurs domiciles. Dans le premier cas, la solution impose des frais de déplacement considérables aux utilisateurs, réduisant ainsi de manière substantielle l'utilité générale du service du fait d'une accessibilité médiocre. Dans le second, les dépenses d'investissement et de fonctionnement deviennent prohibitives, imposant à la collectivité un fardeau considérable sous la forme de prélèvements fiscaux additionnels et susceptibles de devenir abusifs. En conséquence, dans tous les cas de

1. Professeur à l'Université catholique de Louvain et à l'École nationale des ponts et chaussées

2. Je remercie Marc Fleurbaey, Joël Maurice et David Meunier pour leurs nombreux commentaires.

3. Les coûts d'accès incluent un grand nombre d'éléments puisqu'on y retrouve à la fois les dépenses monétaires proprement dites (amortissement du véhicule et carburant dans le cas de la voiture), la valeur du temps pour les individus concernés, l'inconfort provoqué par un moyen de transport surchargé, le risque d'accident, etc.

figure, *le décideur public est confronté à un arbitrage-clé et doit trouver un point d'équilibre entre ces deux forces opposées*, point qui doit en même temps apparaître comme une solution de compromis acceptable pour l'ensemble de la collectivité formée par les usagers, mais aussi par les contribuables⁴.

Mais de quel décideur s'agit-il ici ? Ce fut pendant longtemps le ministère concerné (éducation, équipement, santé), mais la décentralisation, qu'elle soit passée ou future, ne peut laisser les choses en l'état. Il est, en effet, raisonnable de penser que les compétences des instances locales de décision vont croître, compliquant ainsi la recherche de ce point d'équilibre du fait d'une focalisation de ces dernières sur les besoins des seuls habitants dont elles sont les représentantes, et ce souvent au détriment des autres utilisateurs – et contribuables – plus éloignés et, par conséquent, situés en dehors de leurs juridictions. Un nouveau conflit risque ainsi de naître entre instances nationales et locales, les premières étant *a priori* plus soucieuses de l'intérêt général tandis que les secondes sont naturellement et spontanément tournées vers les préoccupations locales. Si les États-Unis et la Suisse ont une longue tradition en la matière, la France manque cruellement d'expérience dans ce domaine, ce qui devrait compliquer le problème pour un temps qui ne sera pas nécessairement court. L'étude des pratiques suivies en la matière par les pays ayant une longue pratique en matière de décentralisation des décisions publiques aurait dû être entreprise depuis longtemps afin d'éviter de gros écueils et de grandes désillusions.

Deuxièmement, le redéploiement spatial des services publics est à l'ordre du jour dans de nombreux pays, de sorte que le cas français n'en est qu'un parmi d'autres, même s'il présente certaines spécificités historiques et politiques qui doivent être soulignées et étudiées. Dans presque tous les pays industrialisés qui ont été marqués par un fort centralisme des pouvoirs de décision, les problèmes de congestion, le gigantisme des administrations, l'inefficacité et la lourdeur des mécanismes de décision et, finalement, les exigences en matière de développement régional se combinent plus ou moins pour conduire à une nouvelle répartition spatiale de nombreux services publics. Les bouleversements politiques que connaît actuellement l'Europe avec, d'un côté, l'apparition de nouveaux États et, de l'autre, l'effort général d'unification réclament une nouvelle définition des compétences terri-

4. Nous retrouverons tout au long de ce chapitre le concept d'arbitrage, lequel joue un rôle crucial dans les toutes questions qui nous occupent ici. Il est donc fondamental d'en comprendre, chaque fois, les éléments ainsi que la nature.

toriales en matière de gestion des services publics. Quels que soient les choix effectués, on peut raisonnablement s'attendre à un accroissement des conflits entre instances représentatives des utilisateurs, et ce jusqu'au niveau de décision devant procéder aux grands arbitrages macro-économiques et politiques.

Avant de passer à la prochaine remarque, il convient encore de souligner que les politiques budgétaires restrictives actuelles, et probablement futures, provoquent une augmentation du coût d'opportunité des fonds publics, ce qui impose un effort plus grand de rationalisation dans les choix d'équipements et d'infrastructures. La raréfaction croissante des fonds publics qui vient, en même temps, compliquer l'arbitrage dont j'ai parlé précédemment, devrait militer pour le recours à des techniques d'évaluation, empruntées à la recherche opérationnelle, plus adaptées et plus rigoureuses que celles qui furent utilisées dans le passé.

Troisièmement, et ce sera ma dernière remarque introductive, la compréhension de la dimension spatiale des équipements collectifs renvoie à une question peu discutée et, surtout, souvent mal comprise : à quelle échelle spatiale doit-on évaluer la performance d'un système d'équipements collectifs ? Outre les relations entre instances de niveaux différents impliquées dans les décisions d'implantation, il existe une difficulté conceptuelle fondamentale d'une autre nature, due au fait qu'une grandeur peut être pertinente à une échelle, disons au niveau local, mais ne plus l'être à une autre, par exemple au niveau national. S'il est vrai que certains principes gouvernant l'organisation des territoires restent valables à toutes les échelles, cela ne veut pas dire qu'il en aille de même pour toutes les questions. La raison en est que *les espaces économiques ne s'emboîtent pas comme les petites poupées russes* ; ils ont leurs spécificités et celles-ci doivent être intégrées d'entrée de jeu.

Qui plus est, les instruments à mettre en œuvre pour combattre les inégalités spatiales ne sont pas les mêmes selon l'échelle retenue. En première approximation, il est donc commode de distinguer entre le *local* (ou le micro-spatial) et le *national* (ou le macro-spatial). On vient de le dire, les forces économiques dominantes agissant à ces deux échelles ne sont pas nécessairement les mêmes. Par exemple, au niveau local, le marché foncier et son fonctionnement constituent un déterminant crucial des choix résidentiels et des localisations des entreprises, puisque le marché fonctionne comme une trieuse entre les utilisateurs potentiels du sol. En revanche, au niveau national, il est raisonnable de penser que l'interaction entre travail et capital, tout comme leur capacité à se déplacer entre régions, devient fondamentale. De surcroît, s'il

est vrai, dans les deux cas, que la mobilité des facteurs de production est un élément critique, il ne s'agit pas du même type de mobilité de sorte que les politiques requises pour corriger les inégalités spatiales ne sont pas les mêmes.

En résumé, le choix d'un *espace de référence* se révèle être un choix critique pour les résultats attendus de toute politique de construction d'équipements ou d'infrastructures (Gérard-Varet et Thisse, 1997). Reste que ce choix devrait, à son tour, être accompagné d'une intégration des préoccupations des autres niveaux concernés. C'est particulièrement important en matière de transport – mais pas seulement – où les enquêtes d'utilité publique se limitent uniquement aux communes traversées par les infrastructures, alors que les usagers couvrent en général une zone beaucoup plus large. L'évaluation doit être capable d'intégrer un maximum de dimensions. Faire l'impasse sur ces diverses questions pour des raisons d'opportunisme politique ou pire, en les reléguant sans discussions sérieuses dans des mesures de décentralisation mal ficelées ne peut conduire qu'à des échecs et des désillusions.

Il n'est pas inutile de préciser d'emblée quelques-unes des conséquences qu'impose la distinction entre échelles. Au niveau micro-spatial, on assiste depuis plusieurs décennies à un étalement croissant des villes. La déconcentration des emplois et la péri-urbanisation des logements en sont deux manifestations importantes que l'on rencontre en France, mais également dans plusieurs autres pays européens. La tendance est donc bien générale. À ce niveau d'analyse, ces phénomènes font penser à une certaine forme de *décentralisation « locale » des activités*. Simultanément, au niveau macro-spatial, une forme particulière de polarisation de l'espace se manifeste sous la forme d'une métropolisation croissante des économies. Le fait que les régions les plus prospères de l'Union européenne soient presque toutes organisées autour d'une grande métropole urbaine ne doit rien au hasard. Une telle polarisation des espaces nationaux renvoie ainsi à une *centralisation « multirégionale » des activités*, tendance qui semble s'opposer à la première⁵. Il s'agit en fait d'une question d'échelle spatiale, laquelle est conceptuellement similaire à celle que pose

5. Dans ce dernier cas, toutefois, il ne faut pas aller trop vite en besogne. En effet, il est tout à fait possible que, durant le processus de baisse des coûts de transport et l'effacement progressif des autres obstacles aux échanges commerciaux, les activités économiques commencent, dans une première phase du processus d'intégration, par se concentrer dans un nombre restreint de grandes régions urbaines pour, dans une seconde phase, se re-disperser vers un plus grand nombre de régions afin de bénéficier des avantages comparés des unes et des autres selon les firmes (Combes *et al.*, 2006). Qui plus est, la contradiction peut n'être qu'apparente puisque la taille des régions à dominance urbaine peut augmenter pendant que, simultanément, les centres-villes perdent des habitants et des activités.

l'agrégation des grandeurs économiques que l'on rencontre en théorie économique.

Quoi qu'il en soit, il faut accepter d'utiliser des outils et des concepts différents en fonction du territoire de référence. Cette question sera abordée tout au long du présent chapitre. En outre, ce territoire, s'il est purement administratif, a de plus beaucoup de chances de ne pas être adéquat au niveau économique dans la mesure où les espaces publics ont été conçus, pensés et dessinés à une époque où les technologies de transport et de communication n'avaient rien à voir avec celles que nous connaissons aujourd'hui (Ozouf-Marignier, 1986). Les choses se compliquent davantage quand on constate que la France est en retard par rapport à ces voisins dans la redéfinition des nouveaux espaces de l'intervention publique : le nombre de communes y est largement plus élevé que dans les autres grands pays européens, alors que l'articulation entre régions et départements est loin d'être claire. En revanche, le découpage administratif existant présente l'avantage d'être cohérent avec la collecte de données statistiques. Reste que des changements d'ensemble s'imposent lorsque le désaccord entre la réalité économique et le fait administratif devient trop grand.

À ce stade de mon argumentaire, je me dois d'avertir le lecteur d'une limite importante de l'analyse qui va suivre. Dans la suite de ce chapitre, j'ai choisi de me concentrer sur la localisation d'équipements pouvant être représenté par un *point* – une école, un aéroport ou une gare – afin de faciliter au maximum la présentation. La localisation d'équipements de type *réseau* est beaucoup plus compliquée à formaliser car elle pose, non seulement, la question de la distance à parcourir pour y accéder, comme pour un équipement classique, mais également celle de savoir comment le nouvel équipement de type réseau va modifier, pour les usagers, les distances aux autres équipements existants, incitant par-là certains agents à modifier leur localisation. Si, dans ce cas, les idées et principes qui vont être présentés ci-dessous conservent leur pertinence, les résultats ne peuvent pas être interprétés tels quels, bien au contraire. Je dois reconnaître que, sauf cas particuliers⁶, ce problème n'a pas encore été abordé de manière satisfaisante par les théoriciens de la localisation et de l'économie géographique. Reste, toutefois, que les méthodes de localisation d'un très grand équipement de nature « ponctiforme », par exemple un grand échangeur ou une plate-forme multi-modale, peut aider à déterminer

6. Comme, par exemple, l'impact que peut avoir une nouvelle infrastructure de transport sur la répartition des activités entre deux régions ainsi reliées (Combes *et al.*, 2006).

le réseau de transport permettant de le desservir. Il s'agit là, d'ailleurs, d'une des interprétations les plus communes du modèle minisum qui sera discuté ici. Ce problème admet d'ailleurs une généralisation, connue sous le nom de problème de l'*arbre minimum de Steiner*, qui cherche à déterminer le réseau de longueur minimale connectant un certain nombre de points tout en permettant d'ajouter des sommets intérieurs. Ce problème compte parmi les plus difficiles à résoudre en optimisation⁷.

1. COMMENT CHOISIR LA LOCALISATION D'UN ÉQUIPEMENT COLLECTIF ?

Partons de l'exemple du conseil municipal d'une commune rurale qui doit décider du lieu de construction d'une nouvelle école, étant entendu que l'ancienne sera détruite pour cause de vétusté. Une solution semble s'imposer d'elle-même : reconstruire la nouvelle école là où se trouvait l'ancienne. Cette solution est moins évidente qu'il n'y paraît à première vue car la commune a connu une forte croissance démographique du fait de l'arrivée de nouveaux habitants qui ont contribué au développement de plusieurs hameaux, auparavant peu peuplés. En outre, il existe encore de nombreux terrains vacants, dispersés au sein de la commune, ce qui donne au conseil une assez grande liberté dans le choix d'un emplacement nouveau⁸. Enfin, les élus sont conscients que chaque famille désire que l'école soit construite à proximité de son domicile afin d'éviter à ses enfants des déplacements longs et fatigants. Dans un tel contexte, il n'est pas déraisonnable de retenir la *distance* parcourue comme une assez bonne approximation des inconvénients suscités par ces déplacements. Cette distance peut être mesurée en kilomètres, en minutes, voire en nombre de rues à traverser, selon les cas. Dans la suite de ce chapitre, je supposerai que la distance est mesurée en euros afin de simplifier l'argumentaire.

En dépit de sa simplicité apparente, cet exemple inclut pratiquement tous les ingrédients nécessaires à la discussion du problème-type de la *localisation* d'un équipement. En effet, pour une localisation donnée de l'école, on obtient une distribution de distances dont la

7. Voir, par exemple, Hwang *et al* (1992).

8. La rareté des espaces disponibles peut faire que le choix d'une localisation est, en pratique, très contraint. Cela ne signifie pas que la localisation des équipements répond moins à une logique d'optimisation que de gestion des opportunités. Bien au contraire, cela veut dire que l'optimisation doit intégrer davantage de contraintes.

moyenne et la variance dépendent de la localisation retenue. En règle générale, cette distribution est inégalitaire, ce qui veut dire que les enfants ont à parcourir des distances différentes pour se rendre à l'école. Pour des raisons morales et politiques, on peut être opposé à de telles disparités et proposer de les réduire autant que faire se peut. Comme point de départ, on suppose l'existence d'une localisation telle que tous les enfants soient situés à la même distance de celle-ci. Il s'agit là, sans conteste, d'une localisation que l'on peut qualifier de *juste* puisque aucune discrimination entre enfants ne subsiste. Pourtant, il n'est pas évident que cette localisation soit souhaitable. Un choix aussi égalitaire peut, en effet, conduire à une solution globalement très inefficace. Plus précisément, on dit qu'une localisation est *efficace* s'il n'existe pas de localisation alternative pour laquelle tous les enfants seraient plus proches de leur école. Autrement dit, on ne peut pas se rapprocher de certains enfants sans nécessairement s'éloigner des autres. On retrouve ici l'optimum de Pareto bien connu en théorie économique du bien-être.

Pour illustrer ces différents concepts, je vais considérer le cas de trois hameaux de taille identique formant un triangle ABC , qui est par ailleurs identique à l'ensemble des localisations efficaces. Si ce triangle est équilatéral (voir Figure 1), le centre du cercle passant par les trois sommets semble s'imposer. Il minimise, en effet, la somme des distances parcourues par tous les enfants en leur imposant des trajets identiques. Cette localisation semble donc posséder toutes les propriétés socialement désirables.

Figure 1

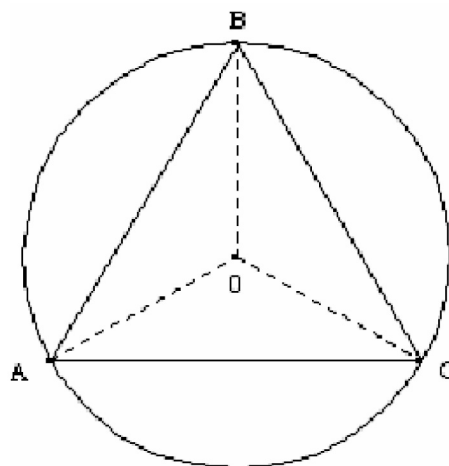
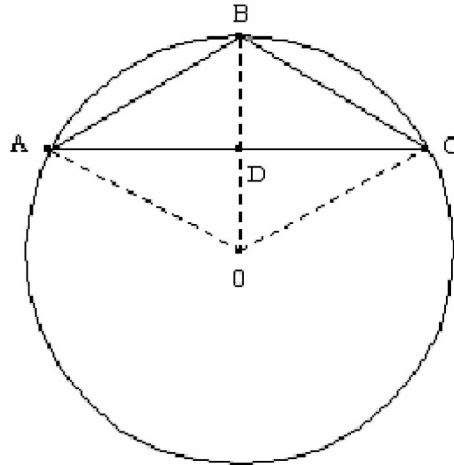
Triangle équilatéral

Figure 2
Triangle obtus



On suppose maintenant que le sommet supérieur B du triangle se rapproche de son côté inférieur, de sorte que le triangle devienne obtus (voir Figure 2). Dans ce cas, la localisation juste est toujours donnée par le centre O du cercle passant par les trois sommets. Toutefois, il est aisé de voir que ce centre est maintenant situé à l'extérieur du triangle. On vérifie alors facilement que le point du triangle qui en est le plus proche, à savoir le pied D de la perpendiculaire partant du centre O , est préféré à ce dernier par la totalité des personnes concernées. Autrement dit, si on demande aux résidents d'exprimer leur préférence entre les deux localisations retenues, on peut s'attendre à ce qu'ils choisissent à l'unanimité une localisation injuste – puisqu'engendrant des disparités –, mais efficace puisque aucune réduction de toutes les distances n'est plus possible, à une localisation juste imposant des déplacements inutilement longs à l'ensemble des enfants.

Cet exemple est beaucoup moins extrême qu'il ne semble à première vue. Pour l'illustrer, je vais considérer un problème totalement différent. Le Tableau 1 donne les évaluations du PIB par habitant des pays européens du début de la Révolution industrielle à la Première Guerre mondiale, telles qu'elles ont été calculées par Bairoch (1997).

Que constate-t-on à la lecture de ce tableau ? En début de période, tous les pays européens avaient des revenus moyens similaires, à l'exception peut-être des Pays-Bas mais la différence avec les autres n'était guère considérable. Au fur et à mesure que la Révolution

industrielle s'étend, on remarque que la croissance économique affecte tous les pays européens qui connaissent un développement économique généralisé, la moyenne augmentant régulièrement de 199 dollars en 1800 pour atteindre 550 dollars en 1913.

Toutefois, les pays sont concernés de manière très variable par ce mouvement car les gains de revenus engendrés varient considérablement d'un pays à l'autre. Plus précisément, l'écart-type de la distribution part de la valeur 24 en 1800 pour atteindre celle de 229 en 1913, soit une augmentation qui équivaut à presque quatre fois celle de la moyenne ; en particulier, les écarts internationaux se creusent progressivement pour atteindre un rapport allant de 1 à 4 entre les pays les plus riches et les plus pauvres. Mais quel état de l'économie européenne préférons-nous, celui de 1800 ou de 1913 ? Je reconnais bien volontiers que la réponse à cette question n'est pas évidente dans la mesure où de fortes disparités sont apparues entre les deux périodes. Toutefois, je vois mal comment argumenter de manière convaincante en faveur de la première situation sans faire appel à des modalités extrêmes, pour ne pas dire pathologiques, d'égalitarisme.

Tableau 1

PIB par tête des pays européens, exprimés en dollars et prix de 1960

Pays	1800	1830	1850	1870	1890	1900	1913
Allemagne	200	240	305	425	540	645	790
Autriche-Hongrie	200	240	275	310	370	425	510
Belgique	200	240	335	450	555	650	815
Bulgarie	175	185	205	225	260	275	285
Danemark	205	225	280	365	525	655	885
Espagne	210	250	295	315	325	365	400
Finlande	180	190	230	300	370	430	525
France	205	275	345	450	525	610	670
Grèce	190	195	220	255	300	310	335
Italie	220	240	260	300	315	345	455
Norvège	185	225	285	340	430	475	615
Pays-Bas	270	320	385	470	570	610	740
Portugal	230	250	275	290	295	320	335
Roumanie	190	195	205	225	265	300	370
Royaume-Uni	240	355	470	650	815	915	1035
Russie	170	180	190	220	210	260	340
Serbie	185	200	215	235	260	270	300
Suède	195	235	270	315	405	495	705
Suisse	190	240	340	485	645	730	895
Moyenne	199	240	285	350	400	465	550
Ecart-type	24	43	68	110	155	182	229

Ce bref détour par l'histoire économique n'est pas inutile car il devrait suffire à montrer que les problèmes évoqués précédemment ont une signification très générale, dans la mesure où ils portent sur des questions fondamentales concernant l'ensemble de la société. Cela étant dit, revenons au cas de l'école. Il est rare que l'on puisse trouver une localisation équidistante pour tous, sauf, ce qui serait évidemment absurde, à retenir une localisation qui soit infiniment éloignée de tous les hameaux ou, ce qui est tout aussi insatisfaisant, à ne rien faire afin de ne pas susciter de mécontentement. À titre d'exemple de solution possible, on pourrait considérer comme juste la localisation minimisant la variance de la distribution des distances. Toutefois, pour éviter le paradoxe mentionné ci-dessus, il faut minimiser cette variance sous la contrainte que la localisation de l'école soit efficace. Une telle approche peut, dans certains cas, améliorer les choses, mais la question de fond reste posée : *quel sens y a-t-il à réduire les disparités entre les enfants si, en même temps, la situation de certains d'entre eux – qui pourraient d'ailleurs être les plus mal lotis – se détériore dans l'absolu ?* On connaît la réponse : c'est au niveau des fondements théoriques du critère de justice que réside la difficulté majeure. L'emploi d'autres indicateurs de disparité pourrait peut-être éliminer certains cas gênants, mais la même difficulté de principe subsiste quoi que l'on fasse (Sen, 1973).

La seule façon connue à ce jour de résoudre la contradiction possible entre efficacité et justice est de choisir une localisation qui minimise la distance maximale parcourue par un enfant. On reconnaît là ce que les économistes appellent le « critère maximin », lequel correspond au principe d'équité proposé par le philosophe américain John Rawls : choisir comme socialement désirable toute action ou politique visant à maximiser le bien-être des plus défavorisés. Ce critère a suscité de nombreuses controverses qui ne peuvent être abordées ici. On le retiendra tel quel dans la suite de ce chapitre, car sa pertinence empirique est incontestable⁹. Dans le cas qui nous occupe, les réductions d'inégalités considérées comme souhaitables sont donc celles dont bénéficient les enfants les plus éloignés de l'école. On dit qu'une localisation satisfaisant ce critère est *équitable*.

Avant de discuter plus en détails les autres critères possibles, je souhaite revenir à la question évoquée dans ma première remarque préliminaire. Cette dimension du problème est, en effet, trop importante pour être évasive. Quels sont les ingrédients de base du problème

9. Le lecteur intéressé est renvoyé à van Parijs (1991) pour un aperçu détaillé de ces débats.

de la localisation de *plusieurs* installations, et ce surtout lorsque leur nombre est *a priori* variable ? Le choix à faire obéit à un arbitrage simple et connu depuis longtemps car on le retrouve dans toutes les questions relatives à l'organisation des espaces économiques ; il est, cependant, souvent mal compris des décideurs ou des usagers. Le premier terme en est donné par le coût de la construction, mais aussi de fonctionnement, des équipements. De très nombreux équipements sont *indivisibles*. Par-là, on entend qu'il n'est pas possible d'offrir une partie d'une infrastructure sans en affecter fondamentalement la qualité ou la fonction. Par exemple, on ne peut pas découper une piscine en petits morceaux pour la distribuer sur l'ensemble du territoire, ni réduire la largeur d'une route afin de lui permettre de relier deux lieux très éloignés. L'une et l'autre doivent avoir une taille minimale, faute de quoi leur raison d'être disparaît. Quand bien même ces exemples sont assez extrêmes, il n'en reste pas moins vrai que de telles indivisibilités sont courantes dans la réalité et ont une implication fondamentale pour l'organisation du territoire. Du fait de l'existence d'un coût fixe $G > 0$ imposé par leur construction, ***les équipements ne peuvent pas être disponibles partout*** à cause de la rareté générale des ressources. Autrement dit, on ne peut y avoir accès qu'en un nombre forcément limité de lieux, ce qui implique nécessairement des disparités dans l'accessibilité.

Pour comparer celles-ci, il faut avoir recours à une *distance* permettant d'évaluer l'effort demandé à, voire encore la perte de satisfaction ressentie par, un usager potentiel pour rejoindre l'équipement, faute de quoi on passe à côté de la dimension la plus importante du problème. Le second terme de l'arbitrage renvoie alors à la ***dispersion géographique*** des usagers établis au sein du territoire en question et, par conséquent, à la distance les séparant des équipements disponibles. Dès lors, une population plus dispersée incite à construire un plus grand nombre d'équipements afin de garantir aux usagers une accessibilité minimale. Partant, on devine aisément comment l'arbitrage entre ces deux forces va jouer : la première pousse à concentrer l'offre de service en un petit nombre de grands établissements, alors que la seconde incite à la construction d'un grand nombre d'établissements pouvant être de plus petite taille. Le nombre et la localisation des installations construites dépendent *in fine* de l'intensité relative des deux forces – qui changent elles-mêmes avec la nature du service et l'état de la technologie – mais aussi, on le verra, avec le mécanisme de décision mobilisé pour résoudre cet arbitrage. Nier cet arbitrage ne peut mener qu'à des déconvenues considérables, aussi bien pour le

décideur que pour la population. Un travail approfondi d'information s'impose donc de manière urgente.

2. ÉQUITÉ SPATIALE OU UTILITÉ GÉNÉRALE : L'ÉTERNEL DILEMME ?

Décrivons maintenant le problème de la localisation d'un équipement de manière plus formelle. Pour cela, on considère un espace métrique (X, d) . Si X représente l'ensemble de tous lieux pertinents pour le problème considéré, d est la distance mesurant l'éloignement entre toute paire de lieux de X . d peut être donné par la distance euclidienne, la longueur du plus court chemin reliant deux points quelconques d'un réseau de transport, voire toute autre métrique que l'on désirerait retenir dans un contexte particulier. On se donne également un nombre fini de points x_1, \dots, x_m de X , représentant les localisations des usagers de l'équipement considéré – les maisons des enfants dans le cas de l'école discuté précédemment. Une remarque importante s'impose à ce stade : le choix de ces points est susceptible de varier sensiblement avec l'échelle spatiale retenue¹⁰. En outre, il dépend avant tout de l'information dont le décideur dispose. Par exemple, ces points pourraient correspondre aux capitales régionales, aux communes françaises, voire si l'analyse est très fine, aux localisations individuelles des usagers. On rencontre de nouveau le problème souligné précédemment : le choix d'un niveau d'agrégation spatiale n'est pas neutre pour les recommandations à faire.

Dans le contexte présent, il faut insister sur le fait que les localisations des usagers sont considérées comme des données, tandis que *seule* la localisation de l'équipement est variable. Plus précisément, cette dernière est choisie en fonction de la position des usagers dans l'espace de référence X , tandis que les localisations des usagers ne sont pas susceptibles d'être modifiées. On verra plus tard que cette hypothèse, qui décrit assez bien la faible mobilité géographique des populations en France, est loin d'être sans conséquences pour la politique à suivre.

À chaque point x_i est ensuite associé un nombre N_i représentant, par exemple, le nombre d'usagers vivant en ce lieu. Ce nombre pourrait lui-même être pondéré par un coefficient exprimant le « souci » que le décideur se fait vis-à-vis du groupe installé en cet endroit ; il

10. Sauf s'il est possible d'utiliser un modèle entièrement désagrégé, ce qui devient possible du fait des nouvelles techniques de traitement de l'information, tels que les SIG.

peut également correspondre au nombre espéré d'usagers dans un espace où la mobilité serait élevée. Rappelons qu'en théorie économique, on appelle *utilité* toute indicatrice du degré de satisfaction ressentie par la personne considérée. Dans ce qui suit, il s'agit plus précisément ici de « l'utilité indirecte », qui est exprimée en fonction de contraintes auxquelles cette personne est confrontée. Il est alors courant – mais aussi commode – de retenir comme première approximation de la véritable utilité d'un usager résidant en x_i une fonction linéaire de son revenu et de la distance le séparant de l'équipement dont la localisation est notée s :

$$U_i(s) = Y - d(x_i, s) \quad i = 1, \dots, N \quad (1)$$

Y désignant le revenu d'un usager, N le nombre total d'usagers et $d(x_i, s)$ la distance entre sa localisation et celle du service public. En retenant l'expression (1), on fait l'hypothèse hautement irréaliste que tous les individus possèdent le même revenu. Il faut insister sur le fait que cette hypothèse est faite pour une seule raison : on cherche à se concentrer *exclusivement* sur la dimension spatiale du problème de la localisation des équipements. Les critères proposés dans les prochaines sections conservent toute leur pertinence lorsque les revenus diffèrent entre individus, à condition qu'ils soient connus du décideur.

En outre, dans la mesure où la satisfaction des usagers décroît avec la distance qu'ils doivent parcourir pour se rendre à l'équipement, l'expression (1) peut être considérée comme une approximation du premier ordre de la véritable fonction d'utilité. On pourrait, sans rien changer à ce qui suit, pondérer la distance par un coefficient spécifique à chaque usager et qui mesure la perte de satisfaction – son coût d'accès – que lui impose un allongement de la distance à parcourir. La difficulté est alors d'un autre ordre : comment le décideur peut-il déterminer la valeur de ces coefficients pour chaque usager ? On retrouve ici le problème classique de la révélation des préférences individuelles, bien connu par ailleurs en économie publique. Qui plus est, cette formulation des préférences individuelles néglige le fait que, du point de vue d'un usager, la désirabilité d'un équipement varie souvent de manière non monotone avec la distance qui les sépare. Si celle-ci est trop grande, il va se plaindre d'être pénalisé par une mauvaise accessibilité au service public, mais une trop grande proximité l'amène souvent à contester le choix ainsi fait à cause des nuisances diverses que l'équipement engendre dans son voisinage immédiat (encombrement de la voirie et difficulté de stationnement, différentes

formes de petites pollutions auxquelles le reste de la population est souvent insensible)¹¹.

L'application du critère maximin conduit au choix d'une localisation qui minimise la fonction suivante :

$$M(s) = \underset{i=1,\dots,N}{\text{Max}} d(x_i, s) \quad (2)$$

Cette fonction satisfait le critère d'anonymat, à savoir que la dénomination des usagers n'a pas d'influence sur la décision finale. C'est, en effet, la minimisation de la fonction $M(s)$ qui détermine qui sera l'usager le plus mal loti, dans la mesure où le bien-être de celui-ci dépend de la localisation retenue pour l'équipement. En théorie de la localisation, on a pris l'habitude d'appeler ce critère *minimax* du fait du renversement des signes des variables dans les fonctions d'utilité. Sa diffusion en théorie de la localisation est contemporaine, bien qu'indépendante, des travaux de Rawls.

Le problème minimax fut introduit en géométrie il y a plus de cent trente-cinq ans sous la forme suivante : déterminer le centre d'une circonférence de rayon minimal incluant un ensemble donné et fini de points du plan. Une méthode simple de résolution consiste à partir de circonférences centrées en x_1, \dots, x_m et de rayon identique, mais suffisamment grand, pour que leur intersection commune soit non vide. On réduit alors la valeur du rayon de manière progressive jusqu'à ce que la taille de l'intersection commune soit inférieure à une tolérance donnée *a priori*. Tout point de cette intersection est une solution du problème maximin. Depuis, des algorithmes considérablement plus performants ont été élaborés, permettant de trouver très rapidement une solution au problème (2) pour des distances pouvant différer de la distance euclidienne, mais aussi pour un très grand nombre m de points décrivant les localisations des usagers. On est, par conséquent, capable de travailler à une échelle spatiale extrêmement fine, retenant les quartiers, les rues et même les habitations.

Si le problème (2), sous sa forme mathématique, est assez ancien, ce n'est que récemment qu'il fut interprété en termes d'application du critère maximin en théorie de la localisation. En fait, la grande majorité des modèles de localisation ont retenu plutôt une approche de type *utilitariste* définie comme la somme du bien-être des individus appartenant à la collectivité considérée. Dans le cas présent, cela revient à

11. Notons que cette attitude n'a rien à voir avec les comportements de type « Not in my backyard », plus connu par son acronyme NIMBY.

sélectionner une localisation qui minimise la somme des distances aux usagers :

$$D(s) = \sum_{i=1}^m N_i d(x_i, s) \quad (3)$$

ou encore, puisque le nombre total d'usagers

$$N = \sum_{i=1}^m N_i$$

est donné, la distance moyenne définie par :

$$\bar{D}(s) = \sum_{i=1}^m (N_i/N) d(x_i, s) \quad (4)$$

Une localisation optimale minimise donc la somme des distances parcourues par les enfants. Autrement dit, on agrège les pertes de satisfaction individuelles en les additionnant. C'est la raison pour laquelle, en théorie de la localisation, on lui a donné le nom de critère *minisum*¹². Il est clair qu'une telle localisation satisfaisant ce critère est toujours efficace ; en revanche, elle est rarement équitable car les lieux occupés par un nombre élevé d'usagers ont un impact fort comparé à ceux possédant peu d'habitants. L'avantage de l'approche utilitariste est de réduire au maximum le coût supporté par l'ensemble de la population, en accordant à chacun le même poids. En outre, en accordant plus de poids aux grandes masses d'usagers, on pourrait penser qu'une localisation utilitariste est susceptible de recevoir l'appui de nombreux usagers, ce qui la rendrait « politiquement » souvent plus acceptable qu'une localisation équitable. J'y reviendrai dans la section suivante.

En conséquence, on peut affirmer qu'il y a généralement *conflit entre équité spatiale et utilité générale*. Une manière possible de le « résoudre », proposée par Morrill et Symons (1977), est d'affecter un exposant α , supérieur ou égal à 1, à la distance dans l'expression (3). Lorsque celui-ci prend une valeur unitaire, on retrouve le critère minisum. À l'autre extrême, lorsque α tend vers l'infini, on obtient le critère minimax. En conséquence, l'idée de retenir des valeurs intermédiaires de ce paramètre, pouvant exprimer des attitudes différentes du décideur vis-à-vis des inégalités spatiales, semble assez naturelle. Dès lors, la fonction

$$D_\alpha(s) = \sum_{i=1}^m N_i d^\alpha(x_i, s) \quad (5)$$

12. Il existe un large accord au sein de la communauté scientifique pour penser que la première formulation de ce problème remonte à Fermat quand il posa la question : comment trouver le point qui minimise la somme des distances aux sommets d'un triangle ?

peut être considérée comme l'expression d'un *compromis* entre les critères d'utilité générale et d'équité. On peut alors interpréter l'exposant α comme étant le degré d'*aversion pour l'inégalité* qui caractérise les préférences du décideur¹³.

Toutefois, une certaine prudence s'impose lorsque l'on emploie ce type d'approche. Ainsi, en étudiant la localisation de casernes de pompiers dans la province du Luxembourg en Belgique – le découpage provincial belge correspond au découpage départemental français –, Richard *et al* (1990) observent que le temps moyen d'intervention des pompiers augmente rapidement dès que le paramètre α prend des valeurs légèrement supérieures à l'unité. Cela suggère que le conflit entre utilitarisme et équité spatiale puisse, en pratique, être sévère – et que dire alors de celui entre utilitarisme et justice. Je me répète donc : un recours trop systématique au seul critère d'équité spatiale, voire même à des objectifs moins contraignants de ce point de vue, est de nature à affecter considérablement le bien-être d'une grande partie de la population. Cela se fera soit au travers d'une accessibilité assez faible aux équipements pour une part importante de la population, soit en augmentant le nombre d'équipements mais en pesant lourdement sur les budgets des collectivités concernées.

À ce stade, un certain nombre de points cruciaux pour l'ensemble du chapitre doivent être soulevés. Premièrement, il faut préciser que tout ce que l'on vient de voir ne se limite en rien au cas d'une école : toute la discussion qui précède garde sa pertinence pour la très grande majorité, sinon la totalité des équipements collectifs habituels. Deuxièmement, l'espace de référence retenu étant quelconque, *l'approche proposée* n'est nullement confinée au cas d'un équipement à construire dans une commune, mais *peut être étendue au cas de plusieurs équipements susceptibles de couvrir un territoire de dimension a priori quelconque*, par exemple toute une région, voire même la totalité du territoire national. À ce niveau macro-spatial, les arbitrages précédents subsistent dans leur intégralité, même si la formalisation est, elle, un peu plus complexe. C'est davantage au niveau des modalités de financement que se pose une question nouvelle, à savoir celle du *nombre* d'équipements désirés.

Cette dernière observation est importante car le problème de l'équité spatiale se matérialise, principalement parce qu'on ne peut

13. Si l'on insère l'expression (1) en lieu et place de la distance dans une fonction de type CES telle que celle donnée par (5), l'aversion pour l'inégalité diminue au fur et à mesure que le revenu des usagers augmente. L'arbitrage entre utilité générale et équité spatiale est donc moins contraignant dans une société riche que dans une société pauvre. Je remercie Marc Fleurbaey de m'avoir fait cette remarque.

souvent construire qu'un nombre très restreint d'équipements par rapport au nombre de sites occupés par les usagers. Précisons de nouveau que si l'on pouvait installer un équipement en chacun de ces sites, c'est-à-dire à domicile comme le Minitel ou l'Internet, les disparités spatiales n'auraient plus guère de raison d'être. C'est la présence d'économies d'échelle dans la production de nombreux services collectifs qui empêche cette multiplication à l'infini des équipements. En d'autres termes, l'origine de la question de l'équité spatiale réside dans la combinaison de la *dispersion de la demande* et de la *concentration spatiale de l'offre* causée par l'existence d'une taille minimum des équipements. Si de plus, on rappelle la liberté dont on dispose dans la spécification concrète des lieux de demande x_i , il devrait être clair que l'existence des outils proposés ici est de nature à aider considérablement le décideur dans ses choix en matière d'aménagement d'espaces publics de différentes natures. Dans ce contexte élargi, le fait de disposer d'outils d'évaluation performants est fondamental, car ils permettent l'étude et la comparaison d'un très grand nombre de scénarios possibles (Labbé *et al.*, 1995 ; Peeters *et al.*, 2002), là où les méthodes habituelles restreignent considérablement l'ensemble des opportunités considérées du fait de la lourdeur des procédures d'évaluation.

Remarquons enfin que la répartition spatiale des usagers comporte deux facettes, à savoir la distribution de la population sur l'ensemble X – dont le rôle est largement reconnu –, mais aussi la *position relative* des points de X dans l'espace – un aspect du problème qui est trop souvent ignoré, en théorie comme en pratique. Une répartition plus ou moins symétrique de ces points a plus de chances de favoriser une convergence entre les critères utilitariste et d'équité, ce qui simplifie considérablement le processus de décision. En revanche, une dispersion très asymétrique va probablement accentuer la divergence entre ces deux critères, compliquant par la même la prise de décision. On n'en sait malheureusement pas beaucoup plus, l'impact de la dispersion spatiale des usagers sur la localisation désirable d'un équipement étant, à la fois, complexe et assez peu étudié¹⁴. C'est d'autant plus regrettable que le rôle joué par cet aspect du problème varie largement avec la configuration et la capacité des différents réseaux de transport que les usagers peuvent emprunter. L'absence de résultats significatifs dans ce domaine montre donc, s'il était nécessaire de s'en convaincre, la nécessité de disposer de *méthodes numériques* de résolution perfor-

14. Voir Thomas (2002) pour un panorama des travaux des géographes sur la question.

mantes pouvant aider significativement la planification des infrastructures de transport.

Avant d'aborder une autre question essentielle, celle de l'acceptabilité par la population du choix d'une localisation, il est important de conclure cette section en se rappelant que la solution pratique à certaines difficultés que le décideur peut rencontrer réside dans le recours à des installations légères et potentiellement mobiles, susceptibles de suivre l'évolution de la distribution spatiale des usagers, en lieu et place d'équipements lourds et immobiles, qui sont trop souvent privilégiés par les élus locaux pour des raisons de prestige : par exemple, le maintien d'une ligne ferroviaire plutôt que son remplacement par une ligne de bus. Bien que connue depuis longtemps, cette manière de faire mérite d'être rappelée car elle est souvent « oubliée » pour de mauvaises raisons telles que, par exemple, des formes extrêmes de justice distributive ou des questions de susceptibilité entre différentes juridictions ou niveaux de décision.

3. DE L'ACCEPTABILITÉ DES POPULATIONS

Que ce soit en tant qu'usagers ou en tant que riverains, il est clair qu'un nombre grandissant de Français et d'Européens ont abandonné toute attitude de passivité ou de résignation face à la construction d'équipements collectifs. De nouvelles installations modifient, souvent de manière sensible, l'environnement et/ou les valeurs foncières, provoquant ainsi des réactions qui peuvent être brutales. À l'inverse, l'absence ou la fermeture de certains équipements (scolaires ou sanitaires) au sein de territoires à faible densité fait que la population concernée a souvent le sentiment d'être délaissée du fait de la fermeture ou du non-remplacement d'équipements, ce qui engendre un mécontentement pouvant avoir des répercussions politiques importantes. C'est donc sans surprise que les décideurs publics s'interrogent afin de savoir comment la population concernée va réagir à ces nouvelles décisions. L'acceptation d'un système d'équipements collectifs par la population devient, par conséquent, un critère additionnel à prendre en compte. À cette occasion, soulignons qu'une question-clé est de savoir de quelle population nous parlons et donc de se rappeler que les jugements peuvent différer selon les échelles spatiales. En effet, un équipement considéré comme hautement nécessaire pour l'ensemble de la collectivité (un aéroport ou un incinérateur de déchets) est souvent tout aussi hautement indésirable au niveau local.

À première vue, il semblerait qu'une telle question ne soit pas du ressort de la théorie économique. Sans prétendre y répondre de

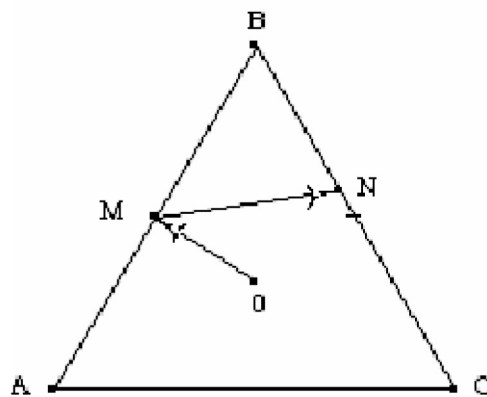
manière exhaustive, il est pourtant légitime de mentionner quelques résultats qui viennent compléter utilement ce que l'on a vu dans les sections précédentes. Dans la suite de cette section, le principal mécanisme de décision collective utilisé dans nos sociétés, à savoir le **vote à la majorité**, va être mobilisé pour appréhender l'idée d'acceptabilité par la population. Cette approche est évidemment réductrice car le concept d'acceptabilité est plus large¹⁵.

On commencera par rappeler que, dans un jeu de vote, on appelle « gagnant de Condorcet » toute action pour laquelle il n'existe pas d'autre action admissible qui lui soit strictement préférée à la majorité absolue des votants. Par conséquent, on parle d'une **localisation de Condorcet** si l'on ne peut en trouver une autre au sein d'un ensemble préétabli qui la défasse lors d'un vote où l'une est opposée directement à l'autre. Une des difficultés principales que pose une localisation de Condorcet est qu'elle peut tout simplement ne pas exister. Pour s'en convaincre, revenons à l'exemple des trois hameaux formant un triangle équilatéral discuté dans la première section et partons de la localisation située au centre O du cercle passant par les trois hameaux. On a vu que cette localisation satisfaisait un grand nombre de critères susceptibles d'être retenus, de sorte que le décideur peut assez légitimement penser qu'elle est *a priori* politiquement acceptable. Pourtant, si on demande à la population des trois hameaux de choisir entre la localisation O et une autre, M , située au milieu du côté AB du triangle de la figure 3, la totalité des usagers vivant dans les deux hameaux reliés par ce côté va choisir M à la majorité absolue car M est plus proche des habitants de ces hameaux que la localisation O (voir Figure 3).

Cela signifie-t-il que cette nouvelle localisation est politiquement acceptable ? La réponse est de nouveau négative. Certains habitants vont, en effet, proposer une troisième localisation N située sur le côté BC du triangle, mais qui soit légèrement plus proche du point B que du point C (voir figure 3). Dans ce cas, tous les habitants de ces deux hameaux vont préférer M à N . On devine et on peut effectivement démontrer que ce processus va se poursuivre indéfiniment, sans jamais converger puisque aucun hameau ne peut obtenir une majorité de suffrages. Par conséquent, la population est incapable de prendre une décision finale concernant son école, attestant ainsi de la difficulté de trouver un compromis démocratique portant sur la localisation d'un

15. Voir Raux et Souche (2004). Je reviendrai en outre sur une autre approche de ce problème dans la sous-section 4.2.

Figure 3
Instabilité du vote



équipement. En fait, on retrouve ici le fameux *paradoxe de Condorcet*, mais sous une forme quelque peu différente.

Rappelons que cet exemple suppose que les trois hameaux ont la même taille. Si, au contraire, un des hameaux, disons *A*, possède une taille qui excède la moitié de la population totale, alors le lieu correspondant est évidemment une localisation de Condorcet. Qui plus est, cette localisation est également celle qui minimise la somme des distances, un résultat connu comme étant le « théorème de la majorité ». En d'autres termes, vote et utilitarisme conduisent à la même solution, un résultat qui reste vrai dans le cas d'un nombre quelconque de points et pour n'importe quelle métrique¹⁶. En revanche, la solution démocratique diverge d'avec la solution équitable qui est toujours donnée par le centre du cercle passant par les trois points.

Malheureusement, les résultats généraux permettant une comparaison systématique du vote et de critères de bien-être sont assez peu nombreux. Toutefois, ceux qui sont disponibles jettent un éclairage original et pertinent sur la question de l'acceptabilité d'une localisation (Hansen et Thisse, 1981 ; Hansen *et al*, 1986). Considérons, en premier lieu, le cas d'un réseau de transport au sein duquel il est impossible de tourner en rond¹⁷. À titre d'exemple, mentionnons la très grande majorité des réseaux fluviaux, mais aussi le système autoroutier de la France qui fut pendant longtemps étoilé et centré sur Paris. Dans un

16. L'existence d'une très grande agglomération suffit donc pour réconcilier les deux procédures.

17. En théorie des graphes, une telle structure est appelée un arbre.

tel cas de figure, il existe toujours une localisation de Condorcet, quelle que soit la forme du réseau et la répartition spatiale des usagers. De surcroît, ce point se confond avec celui qui minimise la somme des distances pondérées par le nombre d'usagers en chaque lieu. Il y a donc, comme dans le cas du théorème de la majorité, concordance entre vote et utilitarisme. Toutefois, comme on peut s'y attendre, cette équivalence n'est plus valable dans le cas d'un réseau contenant des circuits. D'une part, un gagnant de Condorcet n'existe plus nécessairement et d'autre part, quand il existe, il peut diverger considérablement de la solution utilitariste. On comprend donc mieux pourquoi une décision publique, prise dans l'intérêt général, peut engendrer un large mécontentement au sein de la population concernée.

La comparaison des deux procédures dépend évidemment de la répartition des usagers, de sorte qu'il est pratiquement impossible d'énoncer des résultats généraux. Dans de tels cas, on a recours au critère dit du « plus mauvais cas » : *quel est l'écart maximum entre les deux solutions envisagées* ? Si c désigne un gagnant de Condorcet et u la solution du problème minisum, l'inégalité suivante est toujours satisfaite :

$$D(c) \leq 3D(u) \quad (6)$$

Autrement dit, la perte d'utilité générale subie en retenant une localisation de Condorcet ne peut jamais excéder 200 % de l'utilité générale. Il est important de comprendre que ce résultat correspond à l'écart maximum qui puisse être atteint. En pratique, il sera souvent plus faible. Quoi qu'il en soit, il convient d'être conscient que l'emploi d'une procédure de vote pour choisir la localisation d'un équipement collectif est susceptible de provoquer des pertes significatives de bien-être pour l'ensemble de la population ; inversement, l'utilisation du critère utilitariste peut mécontenter de larges segments de la population.

En revanche, si l'on se restreint à des solutions locales, c'est-à-dire à des comparaisons ou à des évaluations en des points situés dans un voisinage de la localisation candidate – voisinage qui peut être grand ou petit, selon les cas –, *l'ensemble des gagnants locaux de Condorcet coïncide avec l'ensemble des maxima locaux du critère minisum*. Autrement dit, en se limitant à ne considérer que des localisations alternatives pas trop éloignées d'un lieu correspondant à un minimum local de la somme pondérée des distances, ce lieu est également accepté par la majorité des usagers à condition de se limiter à des comparaisons par paires dans lesquelles on ne retient que des localisations alternatives qui ne soient pas non plus trop éloignées. On peut

ainsi réconcilier, du moins partiellement, vote et utilitarisme. Le bilan est donc loin d'être aussi négatif qu'on aurait pu le craindre à première vue.

Les résultats portant sur la comparaison entre vote et équité sont moins nombreux. Toutefois, comme ci-dessus, on peut borner supérieurement l'écart entre les deux localisations. Si r désigne la localisation équitable, l'inégalité (6) doit être remplacée par la suivante :

$$M(c) \leq 2M(r).$$

Au lieu d'une perte de 200 %, le vote entraîne maintenant une perte maximum de 100 % cette fois-ci en termes d'équité. Ce moindre écart révèle, une fois de plus, que le choix d'un critère est loin d'être neutre en termes d'acceptabilité.

En conclusion, demander aux usagers de choisir la localisation d'un équipement au moyen d'un vote est susceptible de conduire à une indécision permanente, quand ce n'est pas à des pertes pouvant être importantes en termes d'utilité générale et/ou d'équité. Malgré la simplicité du contexte retenu, qui n'exclut cependant pas une certaine généralité puisque le réseau et l'équipement sont de nature quelconque, ces résultats suffisent à révéler la difficulté qu'il peut y avoir pour un décideur public à retenir une localisation particulière sans provoquer une opposition pouvant être vaste ; opposition qui, par ailleurs, est loin d'être démunie de légitimité puisqu'elle trouve sa source dans le mécanisme de base des démocraties.

Dans quelle mesure ces résultats et préconisations dépendent-ils de l'hypothèse d'un seul équipement à localiser ? Le problème prend ici une nouvelle dimension, à savoir que les usagers sont aussi des contribuables. Dans un tel contexte, il est raisonnable de penser que ceux-ci vont comprendre, du moins à partir d'un certain niveau de dépenses, qu'un accroissement du nombre d'équipements doit inmanquablement conduire à une augmentation du même ordre de grandeur des impôts à supporter. Chaque usager est ainsi confronté à un arbitrage nouveau. Soit l'usager se déclare en faveur d'un plus grand nombre d'installations au prix d'un revenu réduit et donc d'une consommation plus faible des autres biens, soit il souhaite préserver son revenu et renonce par-là même à un parc plus large d'équipements.

Afin de mettre en lumière ce nouvel arbitrage et ses implications politiques éventuelles, on suppose que l'on demande aux usagers de voter sur le nombre d'équipements à installer ainsi que sur leur localisation. Par souci de simplicité, on considère un espace donné par un segment de droite de longueur L . On suppose en outre que la distribution des individus le long de ce segment est uniforme, de sorte que la

densité de population est égale à N/L . Le vote est alors modélisé à l'aide d'une procédure en deux étapes. Dans la première, les individus choisissent à la majorité absolue le nombre d'équipements qu'ils souhaitent ; dans la seconde, ils choisissent leur localisation. On a pris l'habitude de résoudre ce genre de problème en suivant la méthode de rétroduction, ce qui veut dire que le nombre d'équipements est d'abord supposé donné et que les usagers doivent choisir où les installer. Ensuite, connaissant le résultat de ce scrutin, ils déterminent collectivement le nombre total d'équipements souhaité, chacun anticipant correctement où sera situé l'équipement le plus proche pour chaque nombre susceptible d'être choisi. En d'autres termes, la procédure de décision est séquentielle et se déroule en présence d'anticipations des usagers. Malgré un certain manque de réalisme, cette approche nous permet d'appréhender l'idée que se font les usagers de l'arbitrage fondamental discuté précédemment.

La solution de la seconde étape est immédiate et intuitive : tout usager se déplaçant vers l'équipement le plus proche, chaque équipement est installé au centre de son aire de service puisque la distribution des usagers est uniforme. Cela revient à dire que les équipements sont répartis de manière équidistante le long de l'intervalle. Venons-en maintenant au choix du nombre d'équipements. Cette question est plus délicate à traiter que la précédente, car elle réclame de préciser comment les équipements sont financés. On suppose ici que tous les usagers ont le même revenu Y et que les coûts sont couverts par un impôt sur le revenu dont le taux est donné par θ . Si le nombre d'équipements à construire est égal à n , les dépenses sont égales à nG , alors que les recettes sont égales à θYN . Du fait de la contrainte budgétaire, c'est donc sans surprise que l'on constate que les prélèvements, et, par conséquent, le taux d'imposition θ , augmentent avec le nombre d'équipements. En revanche, ce montant diminue avec la densité et/ou le revenu des usagers. Autrement dit, si la multiplication des équipements provoque une augmentation des dépenses publiques, celle-ci a plus de chance d'être acceptée lorsque la population est dense et riche.

Le choix du nombre d'équipements peut alors être formulé de la manière suivante. On consulte la population en lui demandant si elle préfère accroître le nombre proposé d'une unité, procédure qui commence en comparant les nombres 1 et 2. Le nombre d'équipements finalement choisi est celui qui est préféré à la majorité absolue à n'importe quel autre. Le mécanisme ainsi retenu est plus subtil qu'il ne semble *a priori*. En effet, lorsque le nombre n d'équipements est augmenté d'une unité, cela affecte la localisation des autres et, par conséquent, l'accès de tous les usagers à l'ensemble du parc. Certains d'entre

eux peuvent ainsi se trouver dans la situation où, à la fois, ils payent plus d'impôts et souffrent d'une accessibilité moins bonne aux équipements. D'autres bénéficient d'une meilleure accessibilité qui est, à leurs yeux, insuffisante pour compenser l'augmentation des impôts qui en résulte. Restent ceux pour qui le gain en accessibilité l'emporte sur la réduction de leur consommation qu'impose un élargissement de la taille du parc.

Dans l'exemple considéré, on peut démontrer qu'il existe un seul nombre qui puisse être choisi à la majorité. Il est donné par le nombre entier le plus élevé n^* pour lequel l'inégalité suivante est vérifiée¹⁸ :

$$n(n-1) \leq tLN / 2G \quad (7)$$

où t désigne le coût unitaire de déplacement d'un usager.

Venons-en maintenant au critère utilitariste, dans la mesure où il est raisonnable de supposer que le planificateur cherche à minimiser la somme des coûts totaux qu'entraîne le système dans son ensemble. Dans le cas présent, cela équivaut à sommer les coûts de construction payés par les pouvoirs publics au travers de l'impôt et les coûts d'accès supportés directement par les usagers. Comme précédemment, les équipements sont répartis de manière équidistante le long du segment. La détermination du nombre d'équipements est, en revanche, plus simple que dans celui de la procédure de vote. On vérifie aisément que le nombre optimal pour le critère utilitariste est donné par le plus grand des nombres entiers n^o satisfaisant la nouvelle inégalité :

$$n(n-1) \leq tLN / 4G \quad (8)$$

En comparant les membres de droite des inégalités (7) et (8), on constate sans peine que **le nombre d'équipements minimisant les coûts totaux est inférieur au nombre choisi par la population à l'occasion d'un scrutin majoritaire** ($n^o < n^*$). Le corollaire immédiat de cette proposition est que les équipements ont de bonnes chances d'être trop nombreux au regard du critère d'utilité générale. En d'autres termes, le recours à une procédure de vote est susceptible de déboucher sur des dépenses publiques élevées. Plus inattendu, **la solution adoptée par les électeurs correspond très exactement à la solution obtenue si le planificateur avait appliqué le critère d'équité plutôt que celui de l'utilité générale**. Même si ce résultat est établi dans un cadre d'hypo-

18. Pour une démonstration de ce résultat, voir Cremer *et al* (1983) ou Fujita et Thisse (2003, chapitre 5).

thèses très particulier, il n'en reste pas moins très suggestif d'une tendance lourde, à savoir que les critères d'acceptabilité et d'équité poussent vers un plus grand nombre d'équipements.

4. TRANSFERTS ENTRE USAGERS

4.1. *La réconciliation des critères au moyen de transferts entre usagers*

Le problème de l'équité et de l'acceptabilité d'une localisation vient d'être discuté dans un cadre somme tout restreint puisque le décideur ne dispose que d'un seul instrument, à savoir la localisation des équipements. On peut se placer dans une perspective plus large en permettant des transferts croisés entre usagers, par exemple sous la forme de taxes ou de subventions. L'idée est simple et repose sur un principe qui guide de nombreuses décisions publiques : en ayant recours à des transferts entre les différents usagers, les « gagnants » ont-ils la possibilité de *compenser* les « perdants » ? La réponse est ici positive puisque le recours à des transferts rend l'ensemble des utilités réalisables plus vaste et lui donne la forme d'un simplexe. Dans ce cas, l'ensemble des allocations efficaces est identique à l'ensemble des allocations utilitaristes et contient comme toujours l'allocation maximum, provoquant ainsi une réconciliation partielle. Comme on va le voir, il est même possible ici de réconcilier tous les points de vue. Cette observation est d'autant plus pertinente que ce que l'on va voir dans cette sous-section peut être étendu sans difficulté au cas de consommateurs possédant des revenus différents.

Dans le cadre présent, il est raisonnable de laisser les usagers recevant des subventions décider comment ils souhaitent concrétiser cette compensation, par exemple en investissant dans des services publics différents ou en payant moins d'impôts. De cette façon, *si certains usagers semblent favorisés du fait d'un meilleur accès aux équipements, les autres bénéficient en contre-partie d'une consommation plus élevée d'autres biens ou services*. Il n'y a aucune raison pour que des individus dispersés dans l'espace aient la même structure de consommation (même si ceux-ci sont identiques en termes de préférences et de revenu). En effet, la seule dispersion spatiale des résidences suffit pour imposer des solutions différentes dans les arbitrages individuels entre déplacement et consommation. Cela étant admis, si les montants sont évalués correctement, une politique de transferts compensatoires n'a donc rien d'injuste, bien au contraire.

Reste à déterminer le montant de telles compensations – ou du moins une approximation praticable et satisfaisante – qui sont ici calculées en unités monétaires (uniquement pour des raisons de commodité dans l'exposé). Si T_i désigne la taxe payée ou la subvention reçue par un usager localisé en x_i , l'utilité de cet usager devient :

$$U_i(s) = Y + T_i - d(x_i, s) \quad i = 1, \dots, N \quad (9)$$

ce qui implique que son revenu net soit donné par $Y + T_i$. Ces taxes sont forfaitaires et indépendantes du revenu des usagers. En effet, dans la mesure où elles trouvent leur origine dans les différences supportées par les usagers dans leur accès au service, on voit mal pourquoi elles devraient être influencées par le niveau de revenu de ceux-ci. En outre, le budget de cette collectivité doit être en équilibre, ce qui revient à

$$\text{dire } \sum_{i=1}^N T_i = 0.$$

Cette dernière hypothèse implique que les transferts sont effectués à l'intérieur du groupe des seuls usagers et par conséquent qu'il n'y a aucun prélèvement ou financement extérieur. De telles interventions sont pourtant courantes dans la réalité, dès que le territoire considéré fait lui-même partie d'une entité politique plus large. Par contre, elle appréhende très précisément l'idée de *solidarité* au sein du groupe territorial considéré.

Le but poursuivi est alors de déterminer un système de transferts T_1, \dots, T_N qui permette de réconcilier utilitarisme, vote et équité. En partant de (9), il est relativement aisé de vérifier que, pour toute localisation s , de tels transferts existent. Ils ont la forme suivante :

$$T_i^*(s) = d(x_i, s) - \bar{D}(s) \quad i = 1, \dots, N \quad (10)$$

Ces transferts peuvent être positifs ou négatifs. Plus précisément, dans le cas où la distance qu'il parcourt est supérieure à la distance moyenne, l'usager reçoit une subvention égale à la différence entre la première et la seconde ; l'objectif est de le compenser pour la plus grande distance qu'il doit parcourir pour se rendre à l'équipement, du moins *relativement* aux autres usagers. En revanche, le transfert a la nature d'un prélèvement dès que l'usager est situé à une distance inférieure à la distance moyenne ; ce prélèvement correspond à la rente de localisation que lui confère sa plus grande proximité de l'équipement installé en s . La subvention (ou le prélèvement) augmente au fur et à mesure que l'usager est plus éloigné (ou plus proche) de l'équipement.

En sommant les montants payés et reçus, on vérifie immédiatement que la contrainte budgétaire est toujours satisfaite. Les transferts (10) sont donc applicables quelle que soit la localisation retenue pour l'équipement. De par leur nature, ils ne permettent donc pas de déterminer une localisation particulière. Si on additionne les utilités (9), on obtient

$$\sum_{i=1}^N U_i(s) = NY - D(s)$$

de sorte que l'utilité générale est maximisée si et seulement si la localisation retenue minimise la somme des distances pondérées. Autrement dit, les transferts proposés permettent de maximiser la somme des utilités tout en assurant une péréquation des niveaux de bien-être. De fait, le principe qui les soutient est celui de la *compensation*, bien connu en économie du bien-être¹⁹.

Qui plus est, si l'on introduit (10) dans (9), il s'ensuit que tous les usagers atteignent le même niveau de satisfaction lorsque les transferts (10) sont mis en oeuvre. Finalement, remarquons que si les usagers anticipent que les transferts donnés par (10) vont être effectivement mis en oeuvre, *la localisation ainsi retenue est acceptée par l'unanimité des usagers*. En effet, toute autre localisation qui serait, disons, plus favorable à un usager est accompagnée automatiquement d'une augmentation de la taxe qu'il doit payer ou d'une diminution de la subvention qu'il perçoit.

Dans la mesure où (9) constitue une bonne approximation des préférences des usagers, les transferts (10) présentent ainsi un double avantage. D'une part, ils permettent une *réconciliation* des différents critères retenus jusqu'à présent. D'autre part, ils reposent sur des grandeurs qui sont, à la fois, observables et mesurables. Bien entendu, un mécanisme direct de transferts entre usagers ne sera accepté par la population qu'après un travail approfondi d'information. Il n'est politiquement acceptable qu'à condition d'être appliqué dans la transparence la plus grande pour que ses mérites soient reconnus par le plus grand nombre. Entre autres choses, ce résultat met en lumière une idée importante, à savoir que *le décideur a intérêt à augmenter son nombre d'instruments* car cela lui permet toujours d'élargir l'ensemble des utilités réalisables et par conséquent de réduire le nombre de conflits potentiels.

19. Une difficulté bien connue dans la mise en oeuvre de telles compensations réside dans l'ignorance que le décideur a souvent des véritables préférences des agents. Dans le cas qui nous occupe, la distance constitue une approximation raisonnable de celles-ci.

L'analyse précédente peut être étendue afin d'intégrer le financement de l'équipement. Sachant que G désigne le coût d'installation de celui-ci, la contrainte budgétaire devient

$$\sum_{i=1}^N T_i + G = 0.$$

On vérifie alors aisément que, dans ce cas, les transferts (10) doivent prendre la forme suivante :

$$T_i^*(s) = d(x_i, s) - \bar{D}(s) - G/N \quad i = 1, \dots, N \quad (11)$$

Les deux premiers termes, on l'a vu, correspondent à un prélèvement (ou à une subvention) qui dépend de l'écart entre distance moyenne et distance individuelle à l'équipement, tandis que le troisième assure le financement de l'équipement à l'aide d'une taxe identique auprès des usagers.

Au vu de l'ensemble de ces résultats, il est tentant de conclure que le problème de l'équité spatiale est un faux problème puisque l'emploi de transferts appropriés suffirait pour rendre équitable la localisation utilitariste, pour assurer la péréquation spatiale des niveaux de satisfaction retirés des équipements collectifs et par conséquent la cohésion nationale. Une telle conclusion serait – on s'en doute un peu – exagérément optimiste. Tout d'abord, il se peut que l'on ne puisse mettre en pratique les transferts (10) ou (11) du fait de contraintes légales prohibant des prélèvements et subventions discriminatoires basés sur la localisation des usagers. Plus important, les travaux précités retiennent des fonctions d'utilité quasi-linéaires, et ce parce qu'elles permettent la mise en pratique de transferts sans induire de redistribution implicite pouvant être indésirable. Or, Mirrlees (1972) et Wildasin (1986) ont démontré que des utilités non quasi-linéaires impliquent généralement que des individus, par ailleurs identiques mais dispersés dans l'espace, atteignent des niveaux de satisfaction différents dès que l'on retient le critère utilitariste. Par conséquent, la réconciliation de l'utilité générale et de l'équité spatiale a toutes les chances de ne plus être vraie dans un contexte plus large. Comme la section 5 le montrera, on peut cependant identifier un autre mécanisme, transitant par les marchés fonciers, qui conduit à des résultats similaires, mais qui ne réclame ni l'utilisation de transferts, ni l'hypothèse de fonctions d'utilités quasi-linéaires.

4.2. La stabilité des collectivités territoriales

Un territoire sous-équipé est susceptible d'amener sa population à se sentir exclue de la communauté nationale ou régionale, parce

qu'estimant que ses intérêts légitimes sont trop lésés. Une telle attitude est susceptible d'engendrer une demande pour une plus grande autonomie politique, qui peut aller, dans les cas les plus extrêmes, jusqu'à la sécession pure et simple. Dans la mesure où la montée du régionalisme en Europe rend de telles revendications de plus en plus plausibles – formellement, cela signifie ici que la distance voit sa valeur augmenter du fait d'une plus grande valeur idéologique et symbolique –, il est raisonnable de penser que le problème ainsi soulevé est réel. Insistons sur le fait que le risque de sécession évoqué ici ne dépend pas des différences de revenu entre territoires mais de la seule distribution géographique des équipements.

Le nombre et la taille des équipements maximisant l'utilité générale dépendent évidemment de la distribution spatiale des usagers. Toutefois, ce n'est pas, comme on pourrait s'y attendre, la densité de population qui compte dans le calcul de ce nombre, mais *la racine carrée de cette densité*. Si les usagers mécontents des choix ainsi effectués sont dispersés sur l'ensemble du territoire, le risque de conflit est réel, mais celui de sécession est quasi-inexistant. En revanche, si les mécontents sont regroupés dans un petit nombre de sous-régions aisément identifiables, le dernier risque devient réel. Formellement, cela revient à poser la question de savoir si, dans le jeu coopératif où les usagers peuvent former des coalitions – ici des groupes territoriaux – et financer leurs équipements uniquement à partir de leurs propres ressources, il existe des allocations, spécifiant la localisation des équipements et leurs modalités de financement, satisfaisant l'ensemble des usagers et évitant toute sécession. Dans un tel contexte, Haimanko *et al* (2005) ont démontré qu'*il existe des transferts entre usagers tels que la totalité de la population concernée souhaite rester au sein du groupe national*. En revanche, le dédommagement total (on soustrait de la taxe le coût de transport) ne permet pas toujours de prévenir un risque de sécession.

À quoi de tels transferts ressemblent-ils ? Dans le cas particulier où il est optimal de construire un seul équipement pour le groupe territorial considéré – soit du fait que celui-ci est de petite taille, soit du fait d'un coût de construction de l'équipement très élevé –, Le Breton et Weber (2003) établissent que *les usagers éloignés de l'équipement doivent être, du moins partiellement, compensés*. À l'autre extrême, lorsque le nombre d'équipements est très élevé, du fait d'un espace occupé par les usagers de très grande taille ou, ce qui est formellement équivalent, du fait d'un coût d'installation très bas d'un équipement, Le Breton *et al* (2004) démontrent que seuls les schémas où *les usagers supportent le même coût total par tête*, en termes de financement et

d'accessibilité, assurent la cohésion et la stabilité du groupe national sur une base qui est à la fois sociale et territoriale. On retrouve donc, mais dans un contexte différent, le principal résultat obtenu dans la sous-section précédente.

On devine sans peine les implications que ces travaux peuvent avoir pour la politique de localisation des équipements, puisque les résultats varient significativement avec la taille des équipements et celle du territoire concerné, rendant ainsi toute réponse générale quasi-impossible. En outre, il faut aussi que la mise en œuvre des transferts requis soit, comme précédemment, politiquement acceptée par l'ensemble de la population ce qui, justement, peut poser problème dans une société géographiquement ou politiquement polarisée.

5. MOBILITÉ DES USAGERS ET MARCHÉS FONCIERS

Jusqu'à présent, l'hypothèse a été posée que les usagers ne changeraient pas de résidence, de sorte que les équipements collectifs doivent être implantés en considérant ces localisations comme des données qui s'imposent au décideur public. Ce point de vue est souvent défendu par les aménageurs pour qui familles et individus ont le droit d'habiter où ils veulent, les activités, surtout les services publics, devant les accompagner là où ils sont. Cette opinion ignore tout simplement le fait, banal mais prégnant, qui veut que la grande majorité des activités humaines soit soumise à des rendements d'échelle croissants. Dans ce cas, on voit mal comment satisfaire l'objectif précédent sans provoquer une explosion des dépenses publiques et la mise en place de transferts inter-régionaux considérables.

Quelle que soit l'opinion que l'on puisse avoir en cette matière, il ne faut jamais négliger le fait que ceux qui vivent dans des régions peu peuplées font souvent un choix qui traduit leurs préférences pour un certain type d'environnement, incluant un grand nombre d'aménités naturelles. Ce choix peut également être motivé par le désir de consommer de grandes surfaces. Dans ce dernier cas, l'attitude des ménages doit immanquablement provoquer, par agrégation des décisions individuelles, un éloignement des grands équipements. Cela étant admis, il est important de remarquer que la consommation d'espace tend à augmenter sensiblement avec le niveau des revenus. C'est ce qu'atteste, par exemple, le développement de la périurbanisation dans de nombreux pays riches (Cavailhès, 2004). Partant, on voit mal pourquoi des agents économiques ayant effectués de tels choix résidentiels devraient, automatiquement, avoir le « droit » de bénéficier de toutes les facilités que l'on trouve dans les grandes aggloméra-

tions urbaines. Bien sûr, ces dernières procurent à leurs habitants de multiples avantages de nature très différente, mais ils subissent également les nombreux inconvénients causés par l'encombrement et les fortes densités de peuplement. On a tendance à l'oublier un peu vite : en termes de bien-être, aménités naturelles et consommation d'espace peuvent compenser, du moins dans une large mesure, l'absence de certains services publics résultant d'un peuplement peu dense.

De surcroît, on néglige souvent un autre fait, pourtant d'importance majeure : les coûts fonciers ainsi que les temps et conditions de déplacement varient considérablement entre zones urbaines et rurales, et même entre villes de taille différente. Les coûts fonciers augmentés des coûts de déplacement entre domicile et lieu de travail définissent ce que l'on appelle les *coûts urbains*. Dans l'ensemble des pays développés, ceux-ci représentent une part importante et croissante du budget des ménages. Ainsi, dans le cas de la France, entre 1960 et 2000, la somme des postes logement et transport est passée de 22,8 à 39,8 % des dépenses des ménages, ce qui représente une hausse de près de 75 %, et ce malgré un quasi-quadruplement du revenu réel par tête durant la même période²⁰. En outre, ces coûts augmentent rapidement avec la taille de la ville. Les données françaises montrent, qu'en 2000, la part du logement et du transport représente approximativement 45 % des revenus individuels dans les très grandes villes, mais seulement 34 % dans les petites. Ignorer ces différences de coûts dans les évaluations des niveaux de bien-être équivaut donc à biaiser considérablement celles-ci en faveur des grandes villes où les rémunérations nominales sont souvent plus élevées. Dès lors, en tenant compte explicitement des coûts supplémentaires que supportent les habitants des grandes agglomérations, le besoin d'une plus grande péréquation spatiale des revenus en est réduit d'autant.

Afin de mieux comprendre comment joue ce mécanisme de compensation, voyons au moyen d'un exemple simple comment se forme le prix du sol – la *rente foncière* – et comment les usagers s'organisent au

20. Ces pourcentages sont obtenus à partir des données présentées par Rignols (2002). L'indice des dépenses de consommation des ménages était égal à 87,6 en 1960 et à 78,7 en 2000. Les dépenses de logement correspondent à 10,7 % en 1960 et à 19,1 % en 2000, celles de transport à 9,3 % en 1960 et à 12,2 % en 2000. En conséquence, on obtient $(10,7+9,3)/0,876 = 22,8$ en 1960 et $(19,1+12,2)/0,787 = 39,8$ en 2000. Remarquons que les dépenses de transport reprises dans la comptabilité nationale n'intègrent pas la valeur du temps passé dans les déplacements journaliers entre domicile et lieu de travail, mais incluent d'autres postes qui n'ont rien à voir avec ces déplacements. Même si les valeurs calculées ici surestiment les coûts urbains réels, elles sont suffisamment élevées pour que l'on puisse affirmer sans grands risques que ces derniers représentent une fraction importante des dépenses des ménages.

sein de l'espace en fonction des variations de ce prix, au sein d'une zone dotée d'un grand équipement. Pour ce faire, on va se placer dans une perspective micro-spatiale, mais celle-ci a évidemment des implications comparables au niveau macro-spatial comme le montrent les écarts de coûts urbains entre petites et grandes villes. Il importe de distinguer plusieurs types de rentes foncières. Ce que l'on appelle familièrement « rente foncière » recouvre en effet trois concepts totalement différents (Duranton et Thisse, 1996). Tout d'abord, la *rente différentielle* – c'est la seule qui sera retenue ici – qui correspond à la valorisation du sol causée par la proximité d'un lieu particulier, soit parce qu'on y travaille, soit parce qu'on y consomme certains biens ou services. On trouve ensuite la « rente absolue » dont l'origine réside dans le fait que le sol est un facteur de production dont l'offre est limitée par les obstacles naturels (contraintes topographiques) et les plans d'occupation du sol (contraintes institutionnelles). Le sol est ici rémunéré du fait de sa rareté absolue. Dans la suite, cette composante n'est pas retenue afin d'alléger l'exposé au maximum. Enfin, on appelle aussi rente foncière la rémunération qui provient du capital physique affecté au secteur foncier. Dans la réalité, le logement est produit par une combinaison de sol et de capital physique. Ce capital doit être rémunéré partout où il est investi. Cette dernière composante de la rente, appelée « rente immobilière », ne sera pas abordée ici dans la mesure où nous identifions, par souci de simplicité dans l'exposé, logement et quantité de sol utilisée par un consommateur.

La théorie économique de la rente différentielle est fondée sur un principe simple que l'on trouve chez von Thünen : tout se passe comme si chaque ménage comparait toutes les localisations possibles et évaluait, pour chacune d'elle, le montant maximum qu'il serait disposé à payer pour y résider ; chaque lot est alors attribué au ménage proposant l'enchère la plus élevée pour le lot en question²¹. Dans le cas présent, les usagers étant mobiles, la concurrence pour la proximité de l'équipement les amène à payer une rente foncière qui varie avec la distance les séparant du dit équipement. Voyons comment ce principe fonctionne en pratique.

Pour des raisons de simplicité dans l'exposé, supposons que l'espace X soit donné par l'intervalle $[0, L]$. À l'équilibre résidentiel, tous les usagers – supposés identiques par hypothèse – atteignent le même niveau de satisfaction faute de quoi certains seraient incités à s'installer là où leur satisfaction serait la plus élevée, ce qui provoque-

21. Cf. Huriot (1994).

rait une hausse du prix des terrains. Si chacun consomme une unité de sol à des fins de logement (on reviendra plus loin sur cette hypothèse), les usagers sont alors distribués uniformément sur X , impliquant que l'équipement soit installé en $x = L/2$. Dans ce cas, la rente différentielle d'équilibre est donnée par l'expression suivante :

$$R(y) = t |L/2 - y| \quad (12)$$

où $R(y)$ désigne la valeur de la rente qui prévaut à une distance y de l'équipement et t le coût unitaire de déplacement d'un usager. Cette expression est obtenue en partant du point $x = 0$ où la rente différentielle est égale à zéro ; elle atteint sa valeur maximale au lieu où l'équipement est implanté, pour décroître ensuite jusqu'en $x = L$ où sa valeur est nulle. La comparaison de (11) et (12) révèle que la rente foncière différentielle et les transferts compensatoires jouent le même rôle, du moins à une constante près dont la moyenne est égale à $tL/4$. Si les usagers situés à proximité de l'équipement payent une rente foncière plus élevée, le montant diminue avec la distance de sorte que les usagers les plus éloignés ne payent rien. En d'autres termes, **la rente différentielle mime les transferts de compensation**. Un marché foncier concurrentiel, combiné à la mobilité spatiale des usagers, permet donc de résoudre le conflit entre utilité générale et équité spatiale. Autrement dit, la concurrence pour l'occupation du sol entre usagers est la source d'une compensation spontanée entre coûts de déplacement et rente foncière différentielle qui, à son tour, favorise la péréquation des niveaux de bien-être.

On pourrait objecter à cet argument que les usagers atteignent un niveau d'utilité inférieur à celui qui serait le leur si les transferts compensatoires étaient mis en pratique, et ce parce qu'ils doivent payer la rente foncière au propriétaire du logement occupé. Si la rente différentielle agrégée obtenue en additionnant les rentes individuelles, dont la valeur totale est égale à $tL/4$, est taxée intégralement et redistribuée uniformément entre les usagers et habitants, ceux-ci retrouvent le même niveau d'utilité qu'en appliquant les transferts (10). La différence entre les deux mécanismes tient donc à la seule structure des droits de propriété. La rente foncière différentielle trouvant ici son origine dans la présence de l'équipement public, dont la raison d'être résulte elle-même d'un rassemblement humain, on voit mal pourquoi elle ne reviendrait pas à l'agent qui en est le maître d'œuvre puisque celui-ci est **le seul responsable de l'augmentation de la rente différentielle**. En effet, dans le cas qui nous occupe, les propriétaires fonciers n'ont rien entrepris qui puisse justifier le versement de la rente fon-

cière différentielle dont l'origine est à trouver dont l'activité des pouvoirs publics²².

En dépit des apparences, une telle proposition n'a rien de révolutionnaire. Par exemple, elle est en accord avec le contenu d'une lettre envoyée en 1990 à Michael Gorbatchev, signée par de nombreux économistes parmi lesquels on compte plusieurs prix Nobel, et reprise en encadré. Cette proposition n'est, en fait, que la suite d'une autre, plus ancienne et formulée par Henry George dans *Progress and Poverty* en 1879. Elle est souvent mise en pratique au travers de mécanismes, moins extrêmes que la taxation intégrale, dans des pays comme les Etats-Unis ou le Japon (Duranton et Thisse, 1996). Les pouvoirs publics français ont d'ailleurs mené une telle politique à Paris au XIX^e siècle. L'idée de financer les aménagements de la ville au moyen des gains fonciers qui leur sont associés semble, en effet, remonter à Hausmann. Marchand (1993, p. 80-81) résume fort bien ce que fut sa politique :

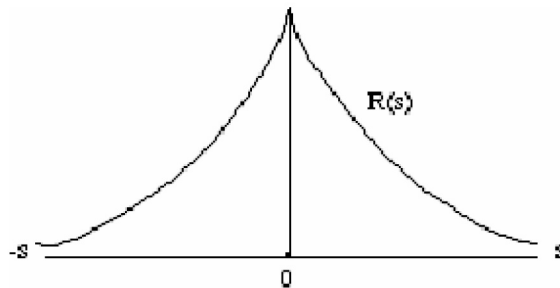
« Ceux qui profitent d'un nouveau parc, d'une nouvelle avenue ou d'un système de transport plus moderne sont non pas les locataires occupants, car leurs loyers augmentent en même temps que la qualité des services dont ils disposent, mais les propriétaires fonciers. Un programme de grands travaux aboutit ainsi à diriger un flux considérable d'argent public vers des poches privées. Comment la collectivité peut-elle recouvrer le fruit de ses investissements ? Hausmann avait trouvé une solution élégante à ce problème classique : la loi de 1852 lui permettait d'exproprier des terres privées, d'y ouvrir une nouvelle avenue et de revendre sur le marché les terres qu'il n'avait pas utilisées, à un prix très supérieur à celui de l'expropriation puisque les travaux avaient considérablement augmenté leur attrait. Le coût des travaux était en partie remboursé par les plus-values produites par ces reventes. »

Bref, rien de bien nouveau finalement, si ce n'est le rappel d'une idée que l'on a oublié un peu vite.

Il est important de remarquer que les résultats présentés ci-dessus ne dépendent pas de l'hypothèse d'une consommation du sol indépendante de la distance à l'équipement. Quand celle-ci est endogène, la rente foncière reste décroissante mais n'est évidemment plus linéaire ; elle devient convexe comme le montre la figure 4, dont l'exponentielle négative constitue souvent une bonne approximation. La densité de population admet donc un profil comparable à celui de la rente fon-

22. La spéculation foncière n'est pas davantage, on s'en doute, une pratique justifiant le versement de cette rente.

Figure 4

Rente différentielle

cière : les usagers implantés à proximité de l'équipement choisissent des logements plus petits, tandis que les usagers éloignés consomment de plus grands logements (Fujita, 1989). Quand bien même les usagers sont a priori identiques, la structure de leur consommation change donc avec la distance à l'équipement. De ce fait, la compensation entre rente différentielle et coût de déplacement est moins directe que dans le cas précédent, car elle doit tenir compte du changement qui s'opère dans la structure de la consommation individuelle. Reste que *la rente différentielle* décroît dans la distance et *capitalise l'avantage procuré par une habitation située à proximité de l'équipement*, de sorte que les usagers éloignés, payant des montants plus faibles par unité de surface, consomment davantage d'espace, ce qui compense leur moindre accessibilité (Fujita, 1989 ; Duranton et Thisse, 1996)²³.

Que de telles forces, transitant par les coûts fonciers et poussant à la péréquation des niveaux de bien-être, du moins pour des classes de revenus similaires, soient effectivement à l'œuvre semble confirmé dans le cas de la Grande-Bretagne par une étude encore préliminaire de Rice et Venables (2003). Des recherches plus poussées et couvrant un grand nombre de cas sont néanmoins nécessaires avant de pouvoir conclure.

6. CONCLUSIONS

En résumé, on peut dire sans grands risques que, s'il est important de ne pas confondre égalitarisme et équité, il est tout aussi crucial de comprendre qu'équité spatiale et utilité générale ne forment pas nécessairement un « couple infernal ». Pour cela, il faut accepter de

23. Le lecteur septique pourra s'en convaincre en pensant à l'évolution des prix du logement au voisinage des « bonnes écoles ».

compenser les usagers défavorisés, soit sous forme monétaire, soit sous la forme d'un menu d'opportunités de services différents ou substituts. Dans ce dernier cas, il est urgent de remplacer le « dur » par l'« adaptable » partout où cela est possible, car le second convient souvent mieux aux régions peu peuplées et rurales, étant susceptible de s'adapter plus rapidement à l'évolution des besoins alors que le premier les fige.

Autrement dit, si l'on souhaite pouvoir réconcilier des desiderata, par ailleurs souvent légitimes même s'ils sont contradictoires, il est fondamental d'élargir l'ensemble des opportunités proposées à la population, tout se réservant une flexibilité suffisante afin de préserver l'avenir. C'est probablement la seule manière d'éviter le contresens de vouloir « tout partout » au travers de transferts interrégionaux qui deviendraient vite excessifs ou de provoquer, à rebours, un sentiment d'injustice susceptible de menacer la cohésion nationale et sociale du pays et des régions. Dans cette perspective, la solution rawlsienne apparaît souvent comme la seule acceptable, même s'il se peut qu'elle ne soit pas soutenue par une majorité d'usagers.

De plus, le recours à des techniques efficaces de recherche opérationnelle est de nature à permettre aux décideurs publics de mieux percevoir les grands enjeux de leur politique et d'illustrer, en particulier auprès des populations concernées, les arbitrages nécessaires mais souvent incompris à la prise de décision. Pour ce faire, il faut aussi mobiliser les outils que proposent les nouvelles technologies de l'information afin de rendre ces arbitrages plus clairs et transparents. Dans ce cas, les pouvoirs publics gagneraient beaucoup à mieux saisir les concepts et fondements de l'économie publique spatiale, ainsi qu'à expliquer aux usagers la nature des problèmes multiples que soulève l'installation d'un système d'équipements.

Enfin, l'économie urbaine nous a permis de mettre en lumière un fait trop négligé, à savoir que les marchés fonciers jouent un rôle crucial en capitalisant les avantages que la proximité des équipements offre aux usagers. Plus précisément, il conviendrait de développer des concepts « équipements + services associés » valorisant mieux le foncier au voisinage du dit équipement, en générant une rente foncière supplémentaire et en la captant pour aider à son financement. Pourquoi se priver de tels instruments, mobilisés pourtant dans des pays peu suspects de faire preuve d'un interventionnisme forcené comme, par exemple, les États-Unis ou le Japon.

Reste un problème d'importance non abordé ici car les résultats sont, à notre connaissance, quasi-inexistants : *quelles sont les interactions possibles entre les politiques de localisation des équipements*

collectifs et la redistribution des revenus entre individus – rappelons que le revenu Y est le même pour tous les consommateurs dans (1) –, dès que l'on reconnaît que la distribution observée n'est pas « optimale » ? Les politiques en matière d'équipement ont pourtant un impact – positif ou négatif, selon les cas – sur les disparités territoriales. Celles-ci, à leur tour, sont très souvent le résultat de la concentration spatiale de travailleurs qualifiés disposant de revenus élevés ou, à l'opposé, de travailleurs peu qualifiés admettant des rémunérations plus faibles (Combes *et al.*, 2003 ; Combes *et al.*, 2006). En conséquence, bien que la redistribution individuelle n'entre pas dans les préoccupations premières du MET, il est clair que son action a des implications redistributives de nature territoriale, qui sont elles-mêmes largement de nature individuelle, phénomène que ce ministère ne peut plus ignorer. Au final, la relation entre les différentes formes d'inégalité doit être reconnue et, partant, il convient de ne plus en séparer les différentes facettes pour appréhender les politiques de redistribution dans leur globalité.

Références

- Bairoch P., *Victoires et déboires, Histoire économique et sociale du monde du XVI^e siècle à nos jours*, Volume II, Paris, Editions Gallimard, 1997.
- Cavailhès J., « L'extension de villes et la périurbanisation », dans *Institut des villes, Villes et économie*, Paris, La Documentation Française, 2004.
- Combes P.-P., Duranton G. et Gobillon L., « Externalités d'agglomération et inégalités salariales entre zones d'emploi », complément C du rapport Compétitivité, Conseil d'Analyse Economique, 2003, p. 163-183.
- Combes P.-P., Mayer Th. et Thisse J.-F., *Économie géographique. L'intégration économique des régions et des nations*, Paris, Economica, 2006.
- Cremer H., de Kerchove A.-M. et Thisse J.-F., « An economic theory of public facilities in space », *Mathematical Social Sciences*, 9, 1985, p. 249-262.
- Duranton G. et Thisse J.-F., « La politique foncière dans une économie spatiale », *Revue Économique*, 47, 1996, p. 197-231.
- Fujita M., *Urban Economic Theory, Land Use and City Size*, Cambridge, Cambridge University Press, 1989.
- Fujita M. et Thisse J.-F., *Économie des villes et la localisation*, Bruxelles, De Boeck, 2003.

- Gérard-Varet L.-A. et Thisse J.-F., « Économie publique locale et économie géographique », *Annales d'Économie et Statistique*, 45, 1997, p. 1-18.
- Haimanko O., Le Breton M. et Weber S., « Transfers in a polarized country : bridging the gap between efficiency and stability », *Journal of Public Economics*, 89, 2005, p. 1277-1303.
- Hansen P. et Thisse J.-F., « Outcomes of voting and planning : Condorcet, Weber and Rawls Locations », *Journal of Public Economics*, 16, 1981, p. 1-15.
- Hansen P., Thisse J.-F. et Wendell R., « Equivalence of solutions to network location problems », *Mathematics of Operations Research*, 11, 1986, p. 672-678.
- Huriot J.-M., *Von Thünen : Economie et espace*. Paris, Economica, 1994.
- Hwang F., Richards D. et Winter P., *The Steiner Tree Problem*, Amsterdam, North-Holland, 1992.
- Labbé M., Peeters D. et Thisse J.-F., « Location on networks », in Ball M., Magnanti T., Monma C. et Nemhauser G. (eds), *Handbook of Operations Research and Management Science : Networks*, Amsterdam, North-Holland, 1995, p. 551-624.
- Le Breton M. et Weber S., « The art of making people happy : how to prevent a secession », *IMF Staff Papers* 50, 2003, p. 403-435.
- Le Breton M., Weber S. et Drèze J., « The Rawlsian principle and secession-proofness in large heterogeneous societies », 2004. <http://www.core.ucl.ac.be>
- Marchand B., *Paris, histoire d'une ville*, Paris, Editions du Seuil, 1993.
- Mirrlees J., « The optimum town », *Swedish Journal of Economics*, 74, 1972, p. 114-135.
- Morrill R.L. et Symons J., « Efficiency and equity aspects of optimum location », *Geographical Analysis*, 9, 1977, p.215-225.
- Ozouf-Marignier M.-C., « De l'universalisme constituant aux intérêts locaux : le débat sur la formation des départements en France (1789-1790) », *Annales ESC*, t. 41, 1986, p. 1193-1213.
- Peeters D., Thisse J.-F. et Thomas I., « Modèles opérationnels de localisation des équipements collectifs », in Cliquet G. et Josselin J.-M. (éd.), *Stratégies de localisation des entreprises commerciales et industrielles*, Bruxelles, De Boek, 2002, p. 59-91.
- Raux C. et Souche S., « The acceptability of urban road pricing. A theoretical analysis applied to experience in Lyon », *Journal of Transport Economics and Policy*, 38, 2004, p. 191-216.
- Rice P. et Venables A., « Equilibrium regional disparities : theory and British evidence », *Regional Studies*, 37, 2003, p. 675-686.

- Richard D., Beguin H. et Peeters D., « Fire stations location in a rural environment », *Environment and Planning A*, 22, 1990, p. 39-52.
- Rignols E., « La consommation des ménages depuis quarante ans », INSEE, Première n° 832, 2002.
- Sen A.K., *On Economic Inequality*, Oxford, Clarendon Press, 1973.
- Thomas I., *Transportation Networks and the Optimal Location of Human Activities. A Numerical Geography Approach*, Cheltenham, UK, Edward Elgar, 2002.
- Van Parijs P., *Qu'est ce qu'une société juste ? Introduction à la pratique de la philosophie politique*, Paris, Seuil, 1991.
- Wildasin D., « Spatial variation of marginal utility of income and unequal treatment of equals », *Journal of Urban Economics*, 19, 1986, p. 125-129.

Dear Mr. Gorbachev :

The movement of the Soviet Union to a market economy will greatly enhance the prosperity of your citizens. Your economists have learned much from the experience of nations with economies based in varying degrees on free markets. Your plans for freely convertible currency, free trade, and enterprises undertaken and managed by individuals who receive the profit or bear the losses that result from their decisions are all highly commendable. But there is a danger that you will adopt features of our economies that keep us from being as prosperous as we might be. In particular, there is a danger that you may follow us in allowing most of the rent of land to be collected privately.

It is important that the rent of land be retained as a source of government revenue. While the governments of developed nations with market economies collect some of the rent of land in taxes, they do not collect nearly as much as they could, and they therefore make unnecessarily great use of taxes that impede their economies--taxes on such things as incomes, sales and the value of capital.

Social collection of the rent of land and natural resources serves three purposes. First, it guarantees that no one dispossesses fellow citizens by obtaining a disproportionate share of what nature provides for humanity. Second, it provides revenue with which governments can pay for socially valuable activities without discouraging capital formation or work effort, or interfering in other ways with the efficient allocation of resources. Third, the resulting revenue permits utility and other services that have marked economies of scale or density to be priced at levels conducive to their efficient use.

The rental value of land arises from three sources. The first is the inherent natural productivity of land, combined with the fact that land is limited. The second source of land value is the growth of communities ; the third is the provision of public services. All citizens have equal claims on the component of land value that arises from nature. The component of land value that arises from community growth and provision of services is the most sensible source of revenue for financing public services that raise the rental value of surrounding land. These services include roads, urban transit networks, parks, and public utility networks for such services as electricity, telephones, water and sewers. A public revenue system should strive to collect as much of the rent of land as possible, allocating the part of rent derived from nature to all citizens equally, and the part derived from public services to the governmental units that provide those services. When governments collect the increase in land value that results from the provision of services, they are able to offer services at prices that represent the marginal social cost of these services, promoting efficient use of the services and enhancing the rental value of the land where the services are available. Government agencies that use land should be charged the same rentals as others for the land they use, or services will not be adequately financed and agencies will not have adequate incentive or guidance for economizing on their use of land.

Some economists might be tempted to suggest that the rent can be collected publicly simply by selling land outright at auction. There are a number of reasons why this is not a good idea. First, there is so much land to be turned over to private management that any effort to dispose of all of it in a short period would result in an extreme depression in prices offered. Second, some persons who could make excellent use of land would be unable to raise money for the purchase price. Collecting rent annually provides access to land for persons with limited access to credit. Third, subsequent resale of land would enable speculators to make large profits unrelated to any productive services they offer, resulting in needless inequity and dissatisfaction. Fourth, concern about future political conditions would tend to depress offers. Collecting rent annually permits the citizens of future years to capture the benefits of good future public policies. Fifth, because investors tend to be averse to risk, general uncertainty about the future will tend to depress offers. This risk aversion is sidestepped by allowing future rental payments to be determined by future conditions. Finally, the future rent of land can more justly be claimed by future generations than by today's citizens. Requiring annual payments from the users of land allows each year's population to claim that year's rent. While the proceeds of sales could be invested for the benefit of future generations, not collecting the money in advance guarantees the heritage of the future against political excesses.

The attached Appendix provides a brief technical discussion of issues of the duration of rights to use land, the transfer of land, the assessment of land, social protection against the abuse and subsequent abandonment of property, and redistribution among localities to adjust for differences in natural per capita endowments. While these issues need to be addressed, none of them present insoluble problems.

A balance should be kept between allowing the managers of property to retain value derived from their own efforts to maintain and improve property, and securing for public use the naturally inherent and socially created value of land. Users of land should not be allowed to acquire rights of indefinite duration for single payments. For efficiency, for adequate revenue and for justice, every user of land should be required to make an annual payment to the local government, equal to the current rental value of the land that he or she prevents others from using.

Sincerely,

Nicolaus Tideman, William Vickrey, Mason Gaffney, Lowell Harriss, Jacques Thisse, Charles Goetz, Gene Wunderlich, Daniel R. Fusfeld, Elizabeth Clayton, Robert Dorfman, Carl Kaysen, Tibor Scitovsky, Richard Goode, Susan Rose-Ackerman, James Tobin, Richard Musgrave, Franco Modigliani, Warren J. Samuels, Guy Orcutt, Eugene Smolensky, Ted Gwartney, Oliver Oldman, Zvi Griliches, William Baumol, Gustav Ranis, John Helliwell, Giulio Pontecorvo, Robert Solow, Alfred Kahn, Harvey Levin.

Appendix on Technical Issues

All individuals and enterprises should have the right to continue using the land they have been using, for as long as they are prepared to pay the rent of that land. The amount of rent to be paid will vary as the economy evolves. As is traditional in countries with market economies, if land is needed for some public purpose such as a highway, the judicial process should guarantee the user fair compensation for any improvements that have been made in good faith. Every user of land should also have the right to transfer ownership of the improvements on the land, together with the right to continue using the land upon payment of rent, to any buyer on any terms upon which they mutually agree. For the rent of land to be collected publicly, land must be assessed, and then reassessed regularly. The assessment process is simplified by the fact that land rental values tend to change smoothly with location. Initially, a map of the value of land can be made by auctioning scattered sites on a rental basis, and then interpolating for the value of other sites, based on the experience of Western appraisers and assessors regarding the manner in which the value of land varies systematically. To update assessments in future years, the assessment office would auction sites that had been relinquished by their users, or sites with improvements that were almost fully depreciated, that had been acquired in voluntary transactions. Interpolation would again be used to estimate the rent of sites that had not been transferred.

With all or nearly all of the rent of land collected publicly, it would be necessary to guard against the possibility that users of land with fully depreciated improvements would abandon their property, leaving the State to demolish the improvements in preparation for the next use of the site. This potential problem can be avoided by requiring every user of land to post a government bond as a "security deposit" that the land will not be abandoned in a run-down condition. Interest on the bond could be applied to the annual rent.

Collection of the rent of land is best managed by local governments, but justice, as well as efficiency in migration incentives, requires that the part of rent that is attributable to nature rather than community development be shared on an equal per capita basis. Thus there is need of clearinghouse mechanism, into which all localities would deposit collections of rent from nature in excess of the average per capita amount, and from which other localities would receive compensation for their deficiencies of rent from nature, relative to the average per capita amount.

ÉQUITÉ TERRITORIALE, ACCEPTABILITÉ ET GRANDES INFRASTRUCTURES DE TRANSPORT¹

Alain Trannoy²

1. LE CALCUL ÉCONOMIQUE : OTAGE DU NOUVEAU RAPPORT DE FORCES ÉTAT-RÉGIONS

La France peut-être considérée comme la patrie du calcul économique, avec la longue tradition des ingénieurs économistes qui démarre au XIX^e siècle avec Jules Dupuit, créateur de la notion de surplus du consommateur, et qui trouve son apogée dans les travaux d'Allais, Boîteux et Lesourne au XX^e siècle. Ce calcul économique, et plus précisément l'analyse coûts-bénéfices appliquée aux choix d'infrastructure de transport telle que consignée dans les deux rapports Boîteux I et II du Plan, n'a pas été contestée dans leur rationalité tant que la France était un État authentiquement centralisé et que la principale valeur était la notion d'efficacité. Le concept de centralisation renvoie à deux phénomènes concrets. D'une part, le peuple ne délègue sa souveraineté qu'à une seule instance, l'instance centrale. D'autre part, l'origine des financements ne provient que de celle-ci. Force est de constater qu'à l'aune de ces deux critères, la France ne possède plus les caractéristiques « pures » d'un pays centralisé. En effet, avec la loi de décentralisation de 1982 et l'élection des conseils régionaux au suffrage universel, les régions se sentent aussi dépositaires d'une partie de

1. Je remercie, Pierre-Philippe Combes, Michel Le Breton, Joël Maurice, Charles Raux, Jacques Thisse, Tanguy Van Ypersele pour d'utiles conversations et les participants à la réunion de travail à Lyon. Je remercie tout particulièrement Marie-Vic Ozouf-Marignier pour plusieurs discussions et pour m'avoir communiqué des sources historiques. Je remercie également Pierre Henri Bono et Laurence Bouvard pour leur aide à la manipulation de données. La remarque d'usage s'applique.

2. ÉHESS, IDEP-GREQAM.

la souveraineté populaire des électeurs domiciliés dans leur périmètre. D'autant que l'article 59 de la loi de 1982 définit les champs d'intervention des conseils régionaux d'une manière plutôt large : « *l'institution régionale a compétence pour promouvoir le développement économique et social, sanitaire, culturel et scientifique de son territoire et pour assurer la préservation de son identité* ». Si le développement du transport régional entre de plein pied dans les compétences régionales, il est non moins clair que les élus régionaux apparaissent comme des interlocuteurs légitimes disposant d'un mandat de leurs électeurs pour discuter avec les représentants de l'État du développement des infrastructures nationales traversant leur région. En outre, les régions peuvent participer au financement de celles-ci au travers des contrats de plan État-région, à telle enseigne que le TGV Est est la première ligne à grande vitesse cofinancée par les collectivités locales traversées. Les régions, les départements et même les villes ont pris à leur charge 25 % de la facture, soit 736 millions d'euros sur 3,125 milliards. La ligne Paris-Marseille avait elle été financée à 90 % par RFF et 10 % par l'État. Et « *cette nouvelle règle du jeu est la seule méthode adaptée au financement des infrastructures ferroviaires pour le présent et le futur* », a prévenu le patron de RFF, Claude Martinand³ ? Qui peut croire que le rapport de force n'a pas changé à la lecture de la délibération adoptée le 6 mai 2001 par le Conseil Régional Midi-Pyrénées « *Le Conseil Régional Midi-Pyrénées exige la réalisation d'une ligne à grande vitesse intégrale Paris-Bordeaux-Toulouse en 3 heures à l'horizon 2016. Pour réaliser cet objectif, il demande que la ligne nouvelle Paris-Tours-Bordeaux soit prolongée jusqu'à Toulouse et que les études préliminaires et les procédures associées sur le tronçon Toulouse-Bordeaux soient engagées dès 2004, afin que les travaux soient réalisés au même horizon que Paris-Bordeaux, soit 2016. Il s'engage sur le principe à participer au financement du projet* ».

Les régions depuis la loi de décentralisation sont apparues pour un temps comme des collectivités peu endettées, laissant aux départements et à l'État le soin de porter le fardeau de la solidarité sociale. D'une certaine manière, les régions ont récolté une partie des missions de feu le Commissariat Général du Plan (CGP) dans leur périmètre d'action mais il n'est pas certain qu'elles aient hérité de sa culture économique. Dans les trente glorieuses, le Plan apparaissait comme une institution légitime qui mettait l'accent en particulier sur la promotion de l'efficacité économique. Le calcul économique y avait une

3. *La Tribune*, article du 29/01/2002.

place d'honneur parce qu'il apparaissait comme un instrument de sélection des investissements publics basé sur la recherche de la maximisation de l'efficacité sociale des deniers publics. La recherche de l'efficacité parétienne, semblait plus urgente dans une France qui cherchait d'abord à combler ses retards accumulés par une première moitié de XX^e siècle marquée par une stagnation relative. Le Plan était une institution technicienne originale, où d'une certaine manière le politique acceptait de s'effacer en partie pour laisser parler l'intérêt général non-partisan. Sa progressive perte d'influence, correspond aussi à la montée des préoccupations d'équité dans la société française. L'arrivée de la gauche au pouvoir en 1981 traduit d'une certaine manière cette inflexion des priorités. Le problème de l'équité est éminemment politique et les instances d'arbitrage politiques, le conseil des ministres, le parlement apparaissent comme les lieux légitimes pour trancher un différent de cette nature. Le débat sur l'équité ne peut pas être tranché par une instance technicienne.

Ce double changement, de priorité relative des valeurs, efficacité – équité et dans le rapport de force État-régions importe pour mieux comprendre les difficultés rencontrées pour faire accepter les conclusions de l'analyse coûts-bénéfices. Elles tiennent d'abord en ce que les élus des conseils régionaux n'ont pas naturellement adopté son langage, dans la mesure où elle est attachée à une valeur, l'efficacité économique, dont la charge politique est assez faible. Elles tiennent ensuite à ce que l'homme politique local prend en compte l'intérêt de ses administrés, ce que l'on ne peut en bonne logique lui reprocher. Si un calcul ou arbitrage est de nature à défavoriser sa circonscription, il prendra partie contre celui-ci et cherchera à le disqualifier, à le délégitimer. Un élu se doit d'être éloquent, il se comporte comme un avocat de sa circonscription. Dans ces conditions, il ne faut pas s'étonner que la faiblesse de l'État se traduise pas une remise en cause de son diagnostic, voire de ses méthodes de diagnostic, en particulier dans la phase de démocratie participative. Les phases d'enquête publique et de débat public obligent en effet les représentants de l'État à des exercices difficiles au cours desquels ils doivent justifier des conclusions sur la base d'une rationalité économique qui n'est pas toujours partagée par le public qui leur fait face. L'une des difficultés, mais pas la seule tourne autour du thème de l'équité territoriale et de sa prise en compte ou non par l'analyse coûts-bénéfices. C'est ce thème qui nous occupe ici exclusivement. Une grande infrastructure de transport entraîne deux effets. D'une part une meilleure accessibilité pour tous les individus ou entreprises desservis, à répartition donnée de la population des individus et des entreprises sur le territoire national. D'autre

part, une modification des différentiels de conditions de production et de consommation entre territoires qui peut entraîner des phénomènes de modification de trajectoire de croissance économique et entraîner par là-même une modification des flux migratoires entre régions desservies et non desservies. Jacques Thisse aborde dans son chapitre, tout comme Bonnafous et Masson (2003) dans leur étude sur l'équité spatiale et les infrastructures de transport, les problèmes d'équité en termes d'accessibilité, et raisonnent donc en grande partie à population donnée. J'aborde ici les problèmes d'équité posés par le second type d'impact, qui est un impact à plus long terme puisqu'il met en jeu la croissance économique et les phénomènes de mobilité inter-régionale.

À cet égard, la critique la plus courante envers l'analyse coûts avantages (ACA) de la part des élus réside dans ce qu'elle avantage mécaniquement les infrastructures de transport entre les régions les plus densément peuplées. En reprenant l'heureuse expression de Marc Gaudry selon laquelle les infrastructures de transport redistribuent du temps, l'ACA favorise une redistribution du temps au profit des zones peuplées et au détriment relatif et parfois absolu de ceux qui habitent les zones peu peuplées. La première ligne de TGV a bien relié Paris à Lyon, c'est-à-dire les deux plus grandes agglomérations françaises, alors que, par exemple, le projet de train pendulaire sur la ligne Paris-Orléans-Limoges-Toulouse (POLT) a été abandonné en raison d'une faible rentabilité socio-économique induite par la faible densité de population des régions traversées.

Or, au moins depuis le livre de Gravier (1947) sur Paris et le désert français, ce thème de l'abandon des zones les moins denses par l'État touche une corde sensible et tous les élus locaux des départements « ruraux », quel que soit leur appartenance politique, entonnent le même cantique comme en témoigne la spectaculaire réunion de Guéret en mars 2005 qui portait sur la disparition des services publics accompagnant la diminution de population dans un certain nombre de départements. À lire ou à écouter les débats parlementaires, cette revendication d'« équité territoriale » semble aller de soi dans le débat politique. À savoir que les hommes politiques des départements peuplés, comme saisis par le remords d'être les élus de zones bien desservies en services publics, ne semblent pas contester le bien-fondé de l'appel à cette notion d'équité. Ainsi, dans le champ politique, enfourcher cette notion d'équité territoriale permet soit d'obtenir des compensations, soit l'annulation de décisions basées sur la rationalité économique.

La question examinée ici est de savoir si cet argument d'équité territoriale est recevable sur le plan d'une théorie « normative » de la justice. Je me fonde sur une discipline relativement récente née des réflexions d'Harsanyi (1953, 1955) sur l'utilitarisme, de Rawls (1971) sur le critère du maximin, l'économie normative, dont deux manuels Fleurbaey (1996) et Roemer (1996) rendent compte du développement. Alors que la théorie du bien-être traditionnelle s'interdisait d'aller au-delà du critère d'efficacité parétienne, la théorie normative étudie les fondements des critères de justice. Que l'on ne se méprenne pas. L'idée ici n'est pas que l'économiste prenne parti ou émette un jugement de valeur sur ce qui est désirable ou pas. Il se borne à inspecter avec ses outils habituels la construction intellectuelle qui conduit à tel ou tel jugement de valeur, il cherche à en tirer un critère opérationnel et il examine quelles sont les politiques qui permettent d'atteindre tel ou tel critère d'équité. Comme exemple célèbre, citons la construction de Rawls qui a montré comment le critère du maximin pouvait être défendu sur la base d'une rationalité individuelle sous la voile d'ignorance. Ce critère a été axiomatisé par les économistes et il est désormais intégré comme l'un des objectifs que la société puisse retenir parmi d'autres.

Ici, nous allons prendre au mot ceux qui parlent d'équité territoriale et examiner dans quelle mesure ce type de discours peut être pris au sérieux sur un plan éthique en s'appuyant sur le corpus existant d'économie normative qui est non négligeable. En tant que chercheur en sciences sociales, nous sommes ainsi dans notre rôle d'interprète de la demande sociale en matière d'équité. Cette interprétation ne peut être servile, elle doit obéir à la rationalité du discours sur la justice et la question posée est de savoir si les critères d'équité territoriale formulés d'une manière plus ou moins latente par le corps social répondent à des critères de rationalité. L'intérêt essentiel de cette démarche est d'offrir un cadre dans lequel une discussion peut s'engager avec les différents acteurs, cadre dans lequel certains arguments sont recevables et d'autres *non*. Au-delà des points sur lesquels il pourrait y avoir désaccord, je propose un cadre de référence pour les différents acteurs. Ce qui est important, c'est que dans ce cadre on pourra défendre certaines propositions, mais pas d'autres : cette théorie élimine un certain nombre d'arguments non recevables et c'est là sa principale utilité. Nous mettons en place un principe d'argumentation, pouvant être reconnu par les interlocuteurs qui participent au débat public. Nous espérons offrir tout ou partie d'un cadre de raisonnement commun partagé pour la formulation et la prise en compte des intérêts particuliers et leur hiérarchisation au sens d'Habermas (1978). L'ACA

offre bien un cadre de raisonnement pour discuter de l'efficacité économique d'un projet. Mais elle n'offre pas de cadre de raisonnement adéquat pour discuter d'équité territoriale.

Signalons d'emblée la difficulté de la tâche qui tient dans ce que ces critères d'équité heurtent notre individualisme méthodologique. L'équité semble l'apanage de l'individu. La notion de territoire ne semble pas revêtir le caractère d'une personne morale. Parler d'équité entre des entités mal définies que sont les territoires suscite des préventions bien compréhensibles et donc le premier test de rationalité est de vérifier qu'il est encore légitime de parler d'équité entre territoires. L'étape préliminaire du raisonnement est donc bien de définir : l'équité entre qui ? Ensuite, le cœur de l'analyse aboutit à deux critères d'équité territoriale, dont nous cherchons à déduire un critère opérationnel qui peut alors se combiner avec un critère d'efficacité économique.

2. CRITÈRES D'ÉQUITÉ ENTRE DES RÉGIONS

D'emblée, la bonne échelle spatiale pour examiner l'impact des grandes infrastructures de transport en termes macroéconomiques semble être la région. Les échelons du département, de la zone d'emploi, de l'agglomération ou de la commune sont trop restreints pour être en mesure d'enregistrer tous les effets d'entraînement des grandes infrastructures de transport. La pertinence d'une certaine délimitation territoriale provient de l'histoire et de la géographie qui conditionnent la vie économique. Certaines régions possèdent une identité, fruit d'une longue histoire partagée, comme la Bretagne et l'Alsace. D'autres régions comme Rhône-Alpes n'ont pas de réalité historique, mais la proximité géographique sous l'impulsion de la métropole Lyon-Grenoble-Saint-Etienne assure à tous les territoires qu'elle rassemble un même destin économique. En tout état de cause, nous considérons comme donné le découpage en régions. Nous supposons que les liens de toutes sortes à l'intérieur d'une région sont suffisamment forts pour que la définition d'un intérêt commun puisse se faire sans trop de difficultés à travers l'instance représentative qu'elles se sont données avec le conseil régional. La légitimité de parler au nom de la région offerte par le suffrage universel constitue à cet égard un atout majeur pour les édiles régionaux dans leur discussion avec le pouvoir central. Les raisonnements effectués ci-dessous exigent en effet cette qualité démocratique des exécutifs régionaux.

Pour pouvoir parler d'équité entre régions comme entre toutes personnes morales, il faut encore que cela ait un sens, qu'une région

soit considérée comme gagnante ou perdante, lorsqu'on passe du schéma de transport A au schéma de transport B. En d'autres termes, les régions ou plutôt les personnes qui parlent en leur nom, l'exécutif régional, doivent être capable de classer des options. Cet exécutif régional doit pouvoir se doter d'un ordre de préférence qui possède les attributs de la rationalité, transitivité et complétude.

L'objection majeure qui survient lorsqu'on parle de préférence d'une collectivité territoriale est de savoir si les préférences du décideur public, ici, l'exécutif régional, peuvent provenir d'une véritable agrégation des préférences soit des membres de l'assemblée, les conseillers élus, soit des électeurs votant dans la région en question. Le théorème d'Arrow a inscrit dans le marbre la difficulté d'agrégation des préférences individuelles, lorsqu'on désire agréger des préordres de préférences, c'est-à-dire des classements d'options. Dans le jargon de la théorie du choix social, les bases informationnelles du choix social sont alors dites ordinales et non-comparables. Ici, je vais retenir un cadre informationnel plus riche, dit cardinal et comparable, qui semble pertinent si la discussion porte, non sur l'agrégation d'opinions, mais l'agrégation d'intérêts. Cette distinction due à Sen (1977) (voir aussi Roberts 1995), a un certain domaine de pertinence pour le problème en question. Une opinion, c'est par exemple, donner son avis s'il fallait ou non envoyer des troupes en Irak. On ne peut pas facilement associer un niveau de bien-être individuel à l'une des deux options en présence, sauf peut-être pour les soldats qui sont susceptibles de faire la guerre. Un intérêt, c'est par exemple, donner son opinion sur l'augmentation éventuelle du RMI de 100 euros par mois par rapport à une situation de référence bien précise. Pour ceux qui vont bénéficier de cette mesure tout comme pour ceux qui vont financer cette mesure, il n'est pas très difficile de concevoir que chaque option correspond à un certain niveau de bien-être. Il n'est pas douteux que le débat concernant la réalisation ou non d'une infrastructure de transport qui affecte l'accessibilité d'un certain nombre de points du réseau et qui est susceptible d'engendrer des flux de revenu à plus ou moins long terme porte sur des questions d'intérêt. Par là même, il est légitime de considérer des fonctions de bien-être social du type de l'utilitarisme ou du maximin et le problème même de l'existence d'une préférence collective s'en trouve surmonté.

Il nous faut maintenant spécifier un peu plus cette fonction de bien-être régional qui agrège donc des fonctions d'utilités individuelles. Il nous faut répondre à la question de savoir si la fonction de bien-être est totalement indifférente au nombre d'individus dans la population. Cette question est importante car la mobilité des agents est

au cœur de notre raisonnement. Il importe donc d'adopter des fonctions de bien-être adaptées à des situations de *population variable*. Considérons d'abord le cas d'école où tous les individus ont exactement les mêmes préférences et bénéficient de la même allocation. La dispersion des utilités est nulle, tous les individus bénéficient du même niveau d'utilité qui est donc le niveau d'utilité moyen. Supposons que la population d'une certaine taille un moment donné diminue suite à un surcroît de mortalité exogène qui n'affecte cependant pas le niveau d'utilité. Il semble assez évident que ce changement sera considéré comme défavorable par la plupart des décideurs politiques, une diminution de la population est considérée comme un moins et une augmentation de la population considérée comme un plus, dans le cas où les individus sont placés dans une situation de parfaite égalité. Ce petit test invalide donc le critère de l'utilité moyenne comme critère de bien-être satisfaisant lorsque la taille de la population est susceptible de varier. Est-ce à dire qu'il faille retenir le critère de l'utilité totale comme critère de bien-être ? Blackorby et alii (2005) nous incitent à être prudent sur ce point et à retenir un critère un peu plus sophistiqué. Plaçons-nous au XIX^e siècle et prenons l'exemple d'une région pauvre qui fait face à une importante fécondité et qui n'arrive pas à nourrir tous ses enfants vu la pauvreté de son sol. Tout supplément de population ne permet pas d'accroître la production qui stagne et qui est à répartir entre un plus grand nombre de bouches. En retenant l'hypothèse classique de décroissance de l'utilité marginale, l'utilité totale ne sera jamais plus élevée que lorsque toute la population sera tellement grande que la production par tête diminue jusqu'à tendre vers zéro. L'utilitarisme total conduit à une conclusion que le philosophe Parfit (1984) a qualifiée de répugnante, le maximum d'utilité sociale est atteint quand toute la population est au seuil de mourir de faim. Pour éviter une telle conclusion, Blackorby *et alii* (2005) ont proposé d'introduire un seuil d'utilité dit critique, c , tel que la fonction de bien-être s'écrit comme la somme⁴ des différences entre les utilités individuelles U_i et ce niveau critique sur l'ensemble de la population. Ainsi le bien-être dans la région j à la date t s'écrit avec une population d'individus égale à n_{jt}

$$W_{jt} = \sum_{i=1}^{n_{jt}} (U_{i_{jt}} - c_t) \quad (1)$$

4. On pourrait introduire un concernement collectif pour l'égalité en transformant les fonctions d'utilité et le niveau critique par une fonction concave. Nous nous sommes dispensés de le faire car il ne semble pas que les problèmes d'équité verticale occupent une grande place dans la définition de la politique de transport inter-régionale. En matière d'interurbain, cette préoccupation devrait être réintroduite.

qui s'écrit encore

$$W_{jt} = n_{jt} \overline{U_{ijt}} - n_{jt} c_t \quad (1')$$

en définissant l'utilité moyenne comme

$$\overline{U_{ijt}} = \frac{1}{n_{jt}} \sum_{i=1}^{n_{jt}} U_{ijt}.$$

Considérons le cas d'un pays dans lequel un certain nombre d'individus ont un niveau d'utilité inférieur au niveau critique. Le niveau de bien-être dans ce pays va augmenter si les habitants dont le niveau d'utilité est plus faible que le niveau critique émigrent. Par contre, le bien-être va baisser si ce sont les individus dont le niveau d'utilité est plus élevé que le niveau critique qui émigrent. Si la valeur du niveau critique est zéro, on retrouve l'utilitarisme classique. Ce paramètre a une valeur strictement positive si les utilités sont contraintes à être positives ou nulles.

Il nous faut maintenant spécifier les utilités individuelles dans un contexte de mobilité spatiale entre deux régions, la région 1 et la région 2⁵. Dans chaque région, à chaque date, l'utilité du ménage est fonction du revenu y et de nombreux facteurs non monétaires a dont des facteurs hédoniques exogènes comme le climat et autres facteurs culturels. Ainsi le niveau d'utilité dans chaque région $j=1,2$ s'écrira

$$U_{ijt} = U_i(y_{ijt}, a_{jt}). \quad (2)$$

Le modèle traditionnel pour analyser des migrations (voir, par exemple, Gobillon et Le Blanc (2003) pour une tentative d'estimation et Gobillon (2001) et Jayet (1996) pour des revues de littérature en langue française) introduit des coûts de migrations monétaires et non monétaires. Ici, pour simplifier, nous considérons uniquement des coûts de migrations monétaires m_{12} qui viennent en déduction du revenu si l'individu migre de la région 1 à la région 2. Ainsi le ménage migre à la date t

$$U_i(y_{i2t} - m_{12}, a_{2t}) - U_i(y_{i1t}, a_{1t}) > 0 \quad (3)$$

c'est-à-dire, si le différentiel d'utilité compte tenu du coût de migration est à l'avantage de la région 2.

5. On pourrait évidemment considérer le cas de k régions. L'essentiel du raisonnement tient avec deux régions.

Le modèle est complètement spécifié quand est précisé comment les infrastructures de transport jouent sur les arguments de la fonction d'utilité et de la décision de migration. Dans chaque région, une variable T_{jt} , que l'on peut ou pourrait vouloir multidimensionnelle, enregistre la qualité et la quantité d'infrastructure de transport auquel ont accès les habitants de la région à la date t . Cette variable impacte les trois variables clés de la décision de migration. Elle influe sur le revenu dans la région car les infrastructures sont l'un des inputs de la fonction de production à côté du facteur travail et du capital privé. Par conséquent, la productivité marginale du travail dépend positivement de la qualité de l'infrastructure de transport et par là même le salaire, au moins à long terme (pour une discussion théorique et un survey des résultats empiriques on pourra se reporter au chapitre 1 de Lafourcade (1998) et à Charlot et Lafourcade (2000)). En deuxième lieu, l'accessibilité est sans nul doute l'un des facteurs de bien-être non monétaires à prendre en compte dans le vecteur a . En troisième lieu, le coût de migration dépend bien de la qualité des infrastructures des deux régions. Plus on peut revenir aisément dans la région de départ et plus le coût de migration est faible. On en déduit un phénomène bien connu des économistes et des géographes mais pas toujours intégré dans le discours politique, à savoir qu'à court terme au moins, l'amélioration des relations de transport entre une région riche et une région pauvre peut se traduire par une accélération des migrations de cette dernière vers la première. Il existe donc des relations de dépendance matérialisées par les fonctions suivantes

$$y_{ijt}(T_{jt}), a_{jt}(T_{jt}), m_{12}(T_{1t}, T_{2t}). \quad j = 1, 2 \quad (4)$$

Chaque région cherche donc à maximiser la fonction objectif (1) par rapport aux infrastructures de transport qui lui sont propres et compte tenu des relations structurelles, (2), (3), (4), la première indiquant les préférences des agents, la seconde les décisions de migration et la troisième comment le transport influe sur la production, le bien-être et la décision de migration. En dynamique, le bien-être total est la variable d'état, et les infrastructures jouent le rôle d'une variable de contrôle. Bien sûr le modèle n'est pas encore bouclé, car il faut indiquer comment l'infrastructure de transport est financée. Cela dépend essentiellement du modèle d'organisation politique du territoire comprenant les régions 1 et 2.

Nous en arrivons au cœur de l'argumentation qui consiste en la construction de critères d'équité entre régions. On imagine une situation hypothétique de contrat social entre régions avant leur union.

Dans la situation originelle, chaque région est autonome politiquement, décide de lever des impôts et de l'affectation du produit de ces impôts levés sur son territoire à une politique d'infrastructures sur ce territoire qui sont les seuls instruments de politique publique considérés ici. Elle peut même passer des accords avec l'autre région qui portent sur la réalisation d'infrastructures communes, une autoroute joignant, par exemple, leurs deux capitales. L'issue de telles négociations peut être décrite par un équilibre de Nash ou par une solution de *bargaining* comme la solution de Nash. Si elle demeure autonome, chaque région connaît une trajectoire de développement en termes de revenu, de facteurs non monétaires de bien-être et de croissance de population. Ces différentes trajectoires commandent alors celle de la variable d'état qui est l'utilité totale au sens de l'équation (1). L'évolution de cette utilité totale incorpore l'évolution de la démographie.

Ces régions sont placées devant un choix hypothétique. Il leur est offert de renoncer à leur souveraineté pour fusionner et fonder un espace unifié, centralisé. Plus précisément, elles renoncent à lever l'impôt sur leurs consommateurs régionaux et ainsi renoncent à façonner de manière indépendante leurs infrastructures de transport sur une longue période de temps ; les décisions concernant les infrastructures intéressant la région sont prises à l'échelon central et sont financées par l'impôt central. Les régions s'accordent également sur le fait de renoncer à lever tout obstacle juridique à la migration de leur population dans le cadre du territoire formé par la réunion des deux régions. Voilà les deux options entre lesquelles chaque région doit choisir dans cette situation originelle.

Etre rationnel, c'est évaluer les différents termes du choix : en quoi le choix va-t-il affecter l'utilité totale modélisée par les équations (1), (2), (3), (4) ? Sur la base du modèle, une grande infrastructure qui dessert la région présente deux types de bénéfices : un bénéfice d'accessibilité qui entre immédiatement comme argument de la fonction d'utilité des agents résidents dans la région, dès la construction de la ligne – c'est un bénéfice statique –, et un bénéfice dynamique du fait que cette infrastructure est susceptible d'attirer de nouvelles activités, de nouvelles populations et, par là même, le taux de croissance de la population régionale va s'en trouver modifié. Cette possibilité d'effets d'agglomération engendrée par la grande infrastructure de transport n'est pas univoque. L'économie géographique considère les choix de localisation des agents comme résultant d'une tension entre forces d'agglomération et forces de dispersion. Selon les modèles d'économie géographique, (Combes et Lafourcade, 2001, pour une revue de littérature), lorsque les coûts de transport sont élevés, une baisse de

ceux-ci engendre une augmentation de la polarisation. En revanche, lorsque les coûts de transport sont très faibles, les effets de dispersion peuvent l'emporter en raison d'effets de congestion dans les zones centrales induits par la diminution cumulée des coûts de transport. Toutefois, ce raisonnement ne vaut que pour une baisse généralisée du coût de transport qui touche toutes les zones. Que se passe-t-il si le coût de transport baisse partout sauf dans une zone donnée, c'est-à-dire en cas de développement asymétrique du réseau de transport ? Il est encore probable que le bénéfice d'accessibilité soit positif pour les consommateurs de cette zone, en raison d'un effet de réseau. Par contre, cette région peut encourir une perte quand la phase de polarisation se déclenche dans les autres régions reliées par l'infrastructure en raison de la mobilité des facteurs dans l'espace national. Le fait qu'une région ne soit pas traversée par une infrastructure de transport est de nature à modifier la trajectoire de croissance de la variable d'état en termes d'utilité totale ; il va y avoir création d'effet d'agglomération autour des têtes de réseaux. Dans un espace unifié, les régions entrent en concurrence au regard des décisions d'infrastructure, ce qui n'est pas le cas ou moins le cas lorsque les régions appartiennent à des espaces nationaux différents. Par exemple, toutes choses égales par ailleurs, une autoroute entre Paris et Lyon⁶ affecte plus sensiblement les perspectives de croissance en Auvergne qu'une autoroute entre Bruxelles et Paris évitant le Nord de la France n'affecte la croissance de cette région. Les habitants de Clermont tout comme ceux de Lille sont mobiles et peuvent déménager à Paris mais en plus les premiers peuvent déménager pour Lyon alors que le passage de frontière réduit la mobilité des Lillois vers Bruxelles.

Nous plaçons les élus locaux dans une situation hypothétique d'information parfaite, ils connaissent les effets dynamiques des infrastructures de transport, ce qui signifie en particulier que les relations de dépendance définies par (4) sont bien identifiées. Une région, avant d'accepter la proposition de fusion, s'interroge sur les garde-fous qu'elle peut exiger du fonctionnement concret de l'État central. Ces garde-fous font partie intégrante du « contrat social » librement consenti entre les régions. La première garantie provient d'une application immédiate d'une analyse coûts-bénéfices au choix entre autonomie et fusion. Aucune région n'acceptera de joindre son destin à celui d'un ensemble plus vaste si elle ne gagne pas en bien-être. Le critère

6. Nous ne prenons pas partie ici sur le point de savoir qui de Paris ou de Lyon va bénéficier le plus de l'autoroute, c'est-à-dire sur le point de savoir si nous sommes dans la partie décroissante ou croissante du U mis en lumière par les économistes géographes.

absolument minimal d'équité inspirée de la littérature sur le partage des coûts entre opérateurs (on pourra se reporter à Moulin (2003)), s'appelle le *stand alone test* (test d'isolement). Dans une question de partage de coût d'un équipement, il ne faut pas que le coût qu'on vous impose soit supérieur à celui que vous subiriez si vous aviez réalisé l'équipement seul. Si la région transpose la logique de ce test au problème qui l'occupe, elle demande à ce **que le pilotage de la variable de contrôle par l'État central, le réseau de transport, soit tel que la trajectoire de sa variable d'état, son bien-être total donné par la formule (1), soit au moins aussi favorable avec la fusion qu'en restant isolée**. Cela revient à demander qu'aucune région partie prenante à la fusion ne voie sa situation se détériorer après celle-ci. Toutes doivent bénéficier *a minima* des économies d'échelle de vivre ensemble. Ce premier critère est un critère minimal d'équité ; s'il n'est pas respecté, l'unité du grand territoire est menacée, la tentation de sécession peut devenir irrésistible (pour une analyse économique des problèmes de stabilité dans la formation de juridictions, on peut se reporter à Alesina et Spolaore (2003), Le Breton et Weber (2003) et Haimanko et alii. (2005) et à Rota Graziosi (2004), pour une revue de littérature). Il apparaît donc également comme un critère de stabilité de la coalition de régions. Il peut être invoqué à n'importe quelle étape de la trajectoire.

Le critère précédent autorise encore des répartitions très inégales du gain à la coopération. On peut également songer à des critères d'équité plus exigeants selon lesquels l'écart initial en valeurs relatives, par exemple entre l'Ile-de-France et la Bretagne, n'augmente pas suite à la fusion ; les trajectoires des deux variables d'état sont alors parallèles. En d'autres termes, dès que **la fusion des territoires s'opère, $t = 1$, le pilotage de la variable de contrôle par l'État central, le réseau de transport, doit être tel que les bien-être totaux régionaux croissent d'une manière homothétique :**

$$\frac{W_{1t}}{W_{2t}} = \frac{W_{10}}{W_{20}} \text{ pour tout } t \geq 1.$$

Ce critère dit d'**équité fort pour la suite** renvoie donc à une répartition proportionnelle des fruits de la fusion aux poids économiques originels donnés par W_{10} , W_{20} . L'écart initial en termes relatifs entre les deux régions n'augmente pas. Si cette clause n'est pas respectée, la région perdante est en droit de contester les décisions d'infrastructure et de demander des compensations. La menace de la sécession n'est cependant pas forcément crédible, la région la plus maltraitée a peut-être encore intérêt à ne pas se dissocier de l'ensemble, car elle connaît

trait peut-être une situation plus difficile en étant isolée. Ce critère ne peut donc prétendre au statut de critère de stabilité de l'union de deux régions.

Nous imaginons par la suite que le contrat social originel liant les régions comprend l'une ou l'autre des deux clauses. L'inscription de l'une ou l'autre clause dans le pacte originel encadre les politiques publiques. Chaque région a le droit de faire appel en leur nom pour contester le bien-fondé d'une politique publique et en particulier de décisions concernant les infrastructures de transport. Un débat peut alors s'engager entre les parties en présence pour examiner la valeur de la requête de la région en question.

Par cette construction, nous avons voulu montrer qu'il n'était pas irrationnel d'évoquer des critères d'équité entre territoires. Il n'est pas difficile d'imaginer qu'une telle situation hypothétique idéalise l'étape de constitution d'un pays fédéral. Nous n'avons pas, cependant, fait de recherches sur ce point. La valeur éthique d'une construction justifiant un critère d'équité ne repose pas sur le fait qu'elle puisse être située historiquement, même si elle renforce incontestablement la validité empirique de son application à un cas d'espèce⁷. Ainsi, il peut être intéressant de se demander si la France a connu un moment de son histoire où ce genre de discussions ou de préoccupations a pu se faire jour. *A priori*, la période d'effervescence de la Révolution française semble être une candidate tout trouvée pour une telle interrogation. La notion de contrat social vient d'être inventée et la Révolution voit disparaître les vieilles provinces au profit d'une construction neuve, les départements. Les pays d'États comme le Languedoc ou la Guyenne disposaient d'une autonomie quant à la construction de grandes infrastructures, que l'on pense par exemple à la construction du Canal du Midi. Ces provinces, dont certaines conservaient une forte autonomie avec leur parlement comme la Bretagne ou le Languedoc, ont-elles réclamées au moment d'être englouties des clauses résolutoires ? Il semble bien que la réponse est plutôt négative pour ce qui est de la Bretagne. Par exemple, les députés bretons, représentant une province rebelle s'il en est, après avoir formé le Club breton à l'origine du Club des Jacobins, ont été à la tête de l'abolition des privilèges de la nuit du 4 Août, qui impliquait également l'abandon des privilèges de la province. Selon Cornette (2005)⁸, cet abandon ne fut pas ressenti doulou-

7. Ce n'est pas, par exemple, parce qu'il n'y a pas eu de « contrat social » entre les colons et les indigènes au moment de la formation d'une colonie, qu'il fallait se dispenser d'appliquer des règles d'équité à l'égard des indigènes par la suite.

8. Tome II, p. 142.

reusement et n'entraîna pas de réactions d'hostilité⁹. Des députés, comme le marquis de Châteauneuf Randon, représentant le Gévaudan, interviennent même pour se féliciter de la disparition de la province du Languedoc qu'il faudrait rendre responsable selon lui du mauvais état des routes dans le Gévaudan, « *parce que depuis le XV^e siècle que le gouvernement crut devoir réunir l'administration particulière de ce pays à celle de la province du Languedoc, le Gévaudan a toujours contribué aux grandes routes de la province, sans avoir pu obtenir, dans son centre, de passage ni d'ouverture de ses montagnes, si ce n'est à frais particuliers.[...] Ainsi, ce malheureux pays, après avoir été oublié en chemin par l'administration vicieuse de cette grande province, venait-il enfin depuis deux ans d'obtenir quelque regard de justice, par le projet de deux grandes routes du royaume qui doivent le traverser et qui sont déjà commencées*¹⁰ ». Ainsi, la solution centralisatrice n'apparaissait pas forcément porteuse d'inéquité. L'état d'esprit de l'époque était aux antipodes d'une réflexion fédéraliste. Écoutons, Sieyès¹¹, un acteur majeur de la révolution, « *La France ne doit pas être un assemblage de petites nations qui se gouverneraient séparément en démocraties. elle n'est point une collection d'États : elle est un tout unique, composé de parties intégrantes ; ces parties ne doivent point avoir séparément une existence complète parce qu'elles ne sont point des tous simplement unis, mais des parties ne formant qu'un seul tout.* » Même les Girondins n'avaient pas de schéma fédéraliste en tête (Furet et Ozouf, 1991) et la création des départements¹², selon Duquesnoy, « *doit produire cet inappréciable avantage de fondre l'esprit local et particulier en un esprit national et public ; il doit faire de tous les habitants de cet empire, des Français, eux qui jusqu'aujourd'hui n'ont été que des Provençaux, des Normands, des Parisiens, des Lorrains* », propos que relaye puissamment Sieyès encore, « *je sens depuis longtemps la nécessité de soumettre la France à une nouvelle division. Si nous laissons passer cette occasion, elle ne reviendra plus, et les provinces garderont éternellement leur esprit de corps, leurs privilèges, leurs prétentions, leurs jalousies* ». Si l'histoire de France n'avait dévié de cet idéal révolutionnaire, nul doute qu'en matière spatiale, seuls seraient

9. Des députés bretons s'opposent toutefois à la disparition de la province comme M. Pellerin voir intervention du 5 novembre 1789 (cf. Mavidal et Laurent, 1867). Il semble que les réticences ont été beaucoup plus importantes dans l'ancienne province du Languedoc, voir par exemple l'intervention de M. Ramel-Nogaret, député de Carcassonne, le 11 novembre 1789, approuvée par les députés de cette province, Mavidal et Laurent, op. citée.

10. Intervention du 4 novembre 1789, Mavidal et Laurent op. citée.

11. Tous les extraits proviennent de Rosanvallon (2004) p. 34-35.

12. Pour plus de détails sur l'état d'esprit au moment de la création des départements, on se reportera à Ozouf-Marignier (1986).

légitimes des critères d'équité entre individus tels que ceux présentés par Jacques Thisse (2006) et que l'invocation de critères d'équité territoriaux serait incongrue et déplacée. L'appel à ceux-ci ne peut pas s'appuyer sur le « moment » constitutionnel révolutionnaire, au contraire, celui-ci pourrait servir à les disqualifier. Toutefois, l'histoire ne s'est pas arrêtée là et, par exemple, vers 1830-1840, les départements expriment l'idée que les réseaux d'infrastructure soient décidés à un échelon plus local. À la même époque, Tocqueville qui fait l'éloge du fédéralisme aux États-Unis s'abstient de demander la déclinaison d'un modèle fédéral pour la France (Benoît, 2005). La question de la décentralisation revient tel un serpent de mer dans le débat politique pendant tout le XIX^e siècle (Deyon, 1992 ; Flory, 1966), comme par exemple vers 1895 mais le sujet est de nouveau enterré (Rosanvallon, 2004). La France a mis deux siècles à amender son organisation territoriale issue de la Révolution, tant elle a été absorbée par d'autres conflits, avec la royauté d'abord, avec l'Église ensuite, avec l'Allemagne et 3 guerres, pour finir avec les colonies. La question territoriale n'émerge au sommet des préoccupations que dans le grand débat qui précède la loi Defferre de 1982. Le 17 mars 2003, le Parlement réuni en Congrès à Versailles adopte le projet de loi constitutionnelle relatif à l'organisation décentralisée de la République. Il modifie le fameux article 1 selon laquelle la France est une République indivisible, pour ajouter que *son organisation est décentralisée*. En gardant en mémoire leur domaine large de compétences sur le plan économique, les régions sont donc fondées, constitutionnellement, à exprimer selon le mot de Siéyès *leurs prétentions et leurs jalousies*. C'est donc plus en interrogeant le présent des lois et non en fouillant l'histoire qu'une résonance partielle peut être trouvée entre le raisonnement mené d'une manière abstraite et un vécu politique. Après avoir été anéanti par l'ouragan révolutionnaire, le fait régional vient de réapparaître, la constitution lui donne force de loi, et il faut compter avec lui.

Avant de clore cette parenthèse historique, il nous faut mentionner que dans l'ancienne province de Bretagne¹³, l'acte union de 1532 entre la France et celle-ci a toujours été utilisé par les régionalistes et les autonomistes pour rappeler que si la France manquait à ses droits en matière de développement de la Bretagne, celle-ci était fondée à déclarer que l'acte d'union était forclos. La création du CELIB (comité d'étude et de liaison des intérêts bretons) en 1950 (Cornette, 2005) qui réunissait en particulier tous les parlementaires bretons, quelle que

13. Les 5 départements et non les 4 départements de la région actuelle.

soit leur couleur politique, pour pousser leurs revendications et discuter avec les représentants de l'État du développement de la région, avec comme plat de résistance les infrastructures, peut, compte tenu du contexte de l'époque, à savoir un retard manifeste de la Bretagne, s'interpréter comme une tentative pour faire respecter par l'autre partie des obligations jugées contractuelles.

3. VERS UN CRITÈRE OPÉRATIONNEL

Nous proposons maintenant un critère opérationnel à partir des réflexions précédentes. Il semble difficile d'essayer de mettre en œuvre le premier critère d'équité proposé. En effet, deviner la trajectoire d'utilité totale si la région était demeurée isolée est un exercice particulièrement difficile s'agissant d'un exercice rétrospectif. Il n'en demeure pas moins que beaucoup de régions dans le monde ont fait sécession sur l'anticipation qu'elles feraient mieux seules qu'en restant sous le toit commun. Lorsqu'une région riche fait sécession, telle la Slovaquie qui se détache de l'ensemble yougoslave, cela semble un calcul moins hasardeux que lorsqu'une région pauvre comme la Slovaquie quitte la Tchécoslovaquie. Et pourtant, les performances économiques de la Slovaquie n'ont cessé d'être meilleures que celle de la république Tchèque depuis la scission.

Nous montrons maintenant que le second critère d'équité, l'équité forte, conduit à un critère opérationnel bien que l'utilité des agents ne soit pas observable. Les trajectoires d'utilité totale ne doivent pas diverger. Nous admettons qu'en France, en 2006, tous les habitants sont au-dessus du niveau critique, c'est-à-dire que *grosso modo* leur vie mérite d'être vécue économiquement parlant. Le modèle n'incorpore pas les différences de fécondité comme un élément pertinent pour l'analyse. L'évolution de la population, en dehors de toute différence de comportement de fécondité entre régions résulte donc de phénomènes migratoires. Là encore, mettons de côté les flux migratoires vers l'étranger qui sont hors-modèle. Nous supposons donc que la somme des populations des deux régions est constante et égale à n . Adoptons une hypothèse commode pour la suite du raisonnement¹⁴.

Hypothèse simplificatrice : Dans chaque région, il existe un agent représentatif pour ce qui concerne le comportement de migration.

14. Cette hypothèse n'est pas en soit nécessaire mais elle permet d'alléger les notations. Les données de migrations sont le plus souvent ventilées par groupe d'âge. On supposerait alors qu'il existe un agent représentatif par groupe d'âge.

Cette hypothèse n'a pas de raison d'être plus ou moins vérifiée pour les migrations que pour d'autres types de comportement. Étudier l'évolution de l'utilité du ménage représentatif dans chaque région est alors suffisant pour savoir dans quel sens évoluent les flux migratoires inter-régionaux. Nous supposons que cet agent représentatif est l'agent moyen.

Nous nous plaçons enfin dans le cas où les deux régions en période originelle (période 0) sont sur un pied d'égalité. Enfin à l'arrière plan, nous supposons encore, dans le cadre du modèle, que la seule cause profonde des migrations tient dans des différences d'accès au réseau qui agissent sur les utilités et sur les possibilités de gain de revenu. Alors nous pouvons énoncer la proposition suivante qui fonde notre critère opérationnel.

Proposition : Une migration de la région 1 vers la région 2 l'année 1 est une condition suffisante pour que le critère d'équité fort ne soit pas respecté au détriment de la région 1 cette année là.

Preuve L'existence d'une migration de 1 vers 2, implique que l'inégalité (3) est respectée pour l'agent représentatif. (3) implique également que l'utilité moyenne dans la région 2 est supérieure à l'utilité moyenne dans la région 1. Comme, de plus, la population de la région 2 augmente au détriment de la région (1) et que l'utilité moyenne est supérieure au seuil critique c , le bien-être total dans la région (2) est plus faible que le bien-être total dans la région (1) en vertu de l'égalité (1').

En prenant la contraposée de la proposition, on en déduit encore que l'absence de migration est une condition nécessaire du respect du critère d'équité territorial fort. Ce critère basé sur l'étude des migrations inter-régionales n'est, cependant, pas une condition suffisante de l'absence d'inéquité en raison de la présence des coûts de migration. Un exemple permettra d'en sentir la difficulté. Une utilité moyenne plus forte en PACA qu'en Île-de-France à un moment donné, engendre un flux migratoire de la région Ile-de-France vers la région PACA à condition que le différentiel d'utilité soit d'un ordre de grandeur suffisant pour que les migrants acceptent de payer le coût de migration. Si celui-ci était très élevé, on pourrait assister à des divergences persistantes d'utilité moyenne entre régions sans constater de migrations. L'absence de migration ne devient une condition suffisante de l'équité territoriale qu'en l'absence de coût de migration.

Risquons-nous à un premier essai d'application sur la période la plus récente depuis le dernier recensement de population. Il est toujours supposé « faussement » que les différences d'accès au réseau sont

à l'origine des migrations *via* leur impact sur les utilités. Nous relâchons par la suite cette hypothèse.

Bornons-nous d'abord à constater que, sur la période la plus récente 1999-2004 selon les chiffres communiqués par l'INSEE (cf. Tableau 1), les flux migratoires qui comprennent également les flux de l'étranger¹⁵ pour 0,23 % par an sont loin d'être négligeables. Six régions connaissent un flux migratoire négatif, le Nord-Pas-de-Calais, la Champagne-Ardenne (0,4 % par an, le record), la Haute-Normandie, la Picardie, la Lorraine et l'Île-de-France.

Tableau 1

Estimation de Population (source INSEE Première n° 1058, janvier 2006)

Région	Estimations de population au 1 ^{er} janvier 2004 révisées (en milliers)	Estimations de population au 1 ^{er} janvier 2005 (en milliers)	Variation relative annuelle 1999-2005 (en %)		
			Totale	Due au solde naturel	Due au solde apparent des entrées et des sorties ¹
Alsace	1 794	1 805	0,68	0,46	0,22
Aquitaine	3 045	3 072	0,93	0,07	0,86
Auvergne	1 326	1 390	0,26	-0,07	0,33
Basse-Normandie	1 442	1 445	0,26	0,27	0,00
Bourgogne	1 623	1 626	0,16	0,08	0,13
Bretagne	3 021	3 044	0,78	0,21	0,58
Centre	2 482	2 490	0,34	0,20	0,14
Champagne-Ardenne	1 336	1 334	-0,11	0,29	-0,40
Corse	273	275	0,95	0,00	0,95
Franche-Comté	1 139	1 143	0,38	0,37	0,01
Haute-Normandie	1 802	1 806	0,24	0,42	-0,19
Île-de-France	11 291	11 362	0,62	0,89	-0,26
Languedoc-Roussillon	2 462	2 497	1,43	0,12	1,32
Limousin	722	724	0,29	-0,32	0,61
Lorraine	2 331	2 334	0,16	0,27	-0,11
Midi-Pyrénées	2 701	2 731	1,15	0,11	1,04
Nord - Pas-de-Calais	4 028	4 032	0,15	0,49	-0,35
Pays de la Loire	3 358	3 385	0,84	0,43	0,41
Picardie	1 875	1 877	0,17	0,44	-0,27
Poitou-Charentes	1 691	1 701	0,62	0,02	0,59
Provence - Alpes - Côte d'Azur	4 703	4 743	0,87	0,22	0,65
Rhône-Alpes	5 896	5 947	0,89	0,52	0,37
Métropole	60 340	60 702	0,62	0,39	0,23
Guadeloupe	448	453	1,21	1,08	0,13
Guyane	184	191	3,44	2,74	0,71
Martinique	395	398	0,73	0,76	-0,03
La Réunion	763	775	1,61	1,44	0,17
Métropole et Dom	62 130	62 519	0,64	0,42	0,23

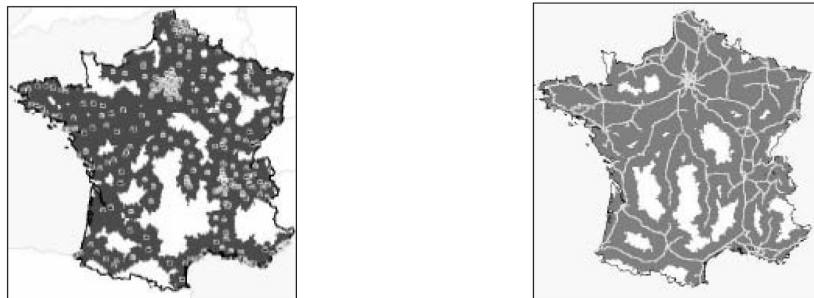
Pour toutes les autres régions, le flux migratoire net est positif ou nul. Il est positif en Auvergne et en Limousin et en particulier dans cette dernière région, le taux d'arrivée net est supérieur à celui de la

15. Il eut été mieux de disposer des flux migratoires internes pour le raisonnement mais ces chiffres ne sont pas encore disponibles.

région Rhône-Alpes. Le taux est supérieur à 1 % par an en Languedoc et en Midi-Pyrénées. Si l'on met en rapport, sans autre forme de procès, cette carte des flux migratoires et celle de l'accessibilité dressée par Sauvant et Salager (2002)¹⁶, soit au réseau ferroviaire soit au réseau routier, qui traduit comme on pouvait s'y attendre qu'une bonne partie des régions Limousin, Auvergne et Midi-Pyrénées figurent parmi les plus mal desservies par ces moyens de transport de terrestre, on constate qu'il existe *grosso modo* une corrélation négative entre le manque d'accessibilité au réseau de transport et le flux migratoire. Dans le cadre de notre modèle, cette corrélation serait une causalité et elle permettrait de rejeter l'hypothèse H0 selon laquelle le manque d'accessibilité de ces trois régions est à l'origine d'un décrochage de l'évolution de leur utilité totale sur la période 1999-2004 par rapport à des régions denses qui constituent des nœuds de transport comme la région Île-de-France. En effet, en appliquant la proposition ci-dessus, le déficit migratoire de la région Île-de-France combiné à l'excédent migratoire du Limousin implique que le niveau d'utilité total y a évolué moins favorablement que dans le Limousin.

Figure 1

Accessibilité par le train (à gauche) et par la route (à droite). Les zones blanches sont à + de 45' en voiture d'une gare importante ou d'une entrée d'autoroute



Source : Sauvant et Salager, 2002

Ce type de raisonnement n'est cependant pas totalement satisfaisant en raison du fait que nous savons que l'accès au réseau de transport et ses effets directs et indirects ne constituent qu'une des motivations des migrations. Par exemple, l'air pur du Limousin peut en constituer une autre, qui n'est que partiellement corrélée à

16. Le constat serait le même si on utilisait les travaux de Combes et Lafourcade sur le coût généralisé de transport (2002) et (2005).

l'absence de TGV ou d'aéroport important. Des causes permanentes qui n'ont rien à voir avec le transport sont peut-être déterminantes dans l'attrait du Limousin. Pour contrôler ces autres facteurs, il faut procéder à la technique de la double différence (voir Meyer, 1995, pour une présentation). Dans notre comparaison des flux migratoires entre Limousin et Île-de-France, nous avons mentalement opéré une différence entre les flux migratoires entre ces deux régions. Il eut fallu disposer d'une autre période plus ancienne, par exemple, la période 1990-1999, et calculer la variation de l'excédent en Limousin puis la comparer avec la variation du déficit en Île-de-France. Cette méthode permet de purger pour des effets fixes Limousin ou Île-de-France pouvant impacter l'évolution du solde migratoire dans ces deux régions d'une manière permanente. Nous nous sommes abstenus de le faire dans la mesure où il n'existe pas d'événement marquant dans le monde du réseau de transport qui soit survenu dans ces deux régions et puisse constituer une véritable expérience naturelle. À titre d'illustration, nous pratiquons cette méthodologie pour mesurer l'impact du TGV Méditerranée à qui l'on ne peut dénier de représenter un possible candidat à une expérience naturelle dans le domaine du transport.

Depuis l'inauguration, le 10 juin 2001, du TGV Méditerranée, Marseille n'est plus qu'à 3 heures de Paris au lieu de 5 heures auparavant. Les autres villes provençales, languedociennes ou de la vallée du Rhône ont vu également leurs temps de parcours ferroviaire considérablement baisser pour rejoindre la région parisienne. Ce TGV est incontestablement un grand succès qu'atteste un fort taux d'occupation. À en croire l'antienne, l'arrivée du TGV aurait provoqué un afflux de population dans la région. En pratiquant un exercice de double différence, le TGV n'a pas provoqué d'accélération de la croissance de population dans les départements qu'il traverse directement (Drôme, Vaucluse, Gard, Hérault, Bouches-du-Rhône, Var). En effet, si l'on compare le surcroît de population¹⁷ entre 1999 et 2004 dans ces départements à la hausse qu'ont connue ces mêmes départements entre 1999 et 1990, force est de constater... que l'on assiste à une baisse : la population n'y a augmenté que de 281 000 personnes alors qu'elle avait augmenté dans la dernière période intercensitaire de 356 000. Toutefois, cela ne prouve rien encore. Il se pourrait également que ce

17. Il eut été mieux de ne comptabiliser que les flux migratoires. Mais, ils ne sont pas encore disponibles sur la période 99-04 à l'échelon départemental. Le fait de raisonner en double différence permet d'éliminer les effets d'une différence de fécondité entre les départements traversés par le TGV et les autres, à condition que cette différence soit restée constante.

mouvement de repli soit général et concerne également toute la France hors agglomération parisienne non traversée par le TGV Méditerranée. C'est exact, la population a moins augmenté dans ces départements entre 1999 et 2004 (1 142 000 personnes) qu'entre 1990 et 1999 (1 262 000 personnes) mais le repli est moindre. La baisse n'est que de 10% pour cette partie de la France alors qu'elle est de 21 %¹⁸ pour les départements traversés par le TGV. Alors que ces départements ont bénéficié d'un atout supplémentaire du fait de cette innovation technologique, ils ont connu une décélération de leur croissance démographique plus forte que les départements n'ayant pas connu l'introduction d'un équipement analogue. Ce constat infirme l'idée naturelle d'un afflux particulier de population dans le midi induit par le TGV Méditerranée.

Après tout, il n'est pas difficile d'imaginer la raison pour laquelle le TGV Méditerranée n'a pas généré cet afflux de population. Grâce au TGV, il est plus facile de séjourner dans le midi pour de courtes périodes, un week-end par exemple. Le midi de la France étant plus accessible depuis Paris, cela peut atténuer l'envie de déménager de manière définitive puisque les escapades occasionnelles sont plus faciles. On pourrait donc s'attendre à ce que le TGV Méditerranée ait entraîné un afflux de population saisonnière, avec une augmentation des résidences secondaires, ce qui serait de nature à peser sur les prix. Le même type de raisonnement que celui effectué plus haut infirme cependant cette thèse, on constate, selon les premiers chiffres du recensement sur la période 1999-2004, une décélération de la croissance du parc de résidences secondaires dans toute la France et celle-ci est encore plus marquée dans les régions traversées par le TGV, la région PACA et la région Languedoc-Roussillon ! Ajoutons d'ailleurs que la hausse des prix de nature essentiellement spéculative a pu dissuader des personnes de s'installer dans le Sud-Est de manière définitive ou temporaire pour s'installer de préférence dans d'autres régions. C'est peut-être là l'une des raisons du redémarrage démographique du Massif-Central.

Ainsi, sous bénéfice d'inventaire¹⁹, un des exemples emblématiques d'une infrastructure lourde en matière de transport ne semble pas avoir engendré de déplacement de population supplémentaire du moins à court terme. Bien évidemment, il est également nécessaire de raisonner sur des horizons plus longs. Mais lorsque l'on

18. Il faut raisonner en taux pour tenir compte de la différence des effectifs de population.

19. La méthode de différences en différences est loin d'être exempte de critiques (cf. Trannoy, 2003).

étend la période d'étude des phénomènes migratoires, on ne peut plus considérer la fécondité comme exogène, ce que nous avons fait jusqu'ici. La raison d'un solde naturel négatif en Auvergne ou en Limousin vient de ce que pendant des décennies, ces régions ont été des terres d'émigration. Ainsi, peut-on énoncer en partie que la fécondité d'aujourd'hui est le reflet des migrations d'hier. Pour livrer une vision longue de l'impact des migrations, il faut donc raisonner sur l'évolution de la population totale. Comme l'économie géographique met en avant les phénomènes de polarisation ou de dispersion induits par la baisse des coûts de transports, il importe de raisonner en tenant compte de la rareté de l'espace et de travailler sur les densités de population. Ici, nous avons calculé l'indice de Gini²⁰ (la moyenne des écarts de densité en valeur absolue rapportée à la densité moyenne) des densités régionales sur la période 1851-2004. Il importe de réaliser que cet indice d'inégalité de densité est anonyme. L'indice peut rester au même niveau alors que les régions pauvres en densité ne sont plus les mêmes. Ainsi, si l'inégalité reste constante, on ne respecte qu'une version anonymisée du critère d'équité fort. À partir de 2008, il sera possible de calculer l'évolution annuelle de cet indicateur grâce au recensement tournant. C'est un indice grossier mais qui permet d'enregistrer comment évolue la tension entre forces d'agglomération et de dispersion. C'est typiquement le genre d'indice dont aurait besoin le CIADT (Comité interministériel de l'aménagement et du développement du territoire) pour son calibrage de la politique d'aménagement du territoire. Nous avons procédé à deux calculs suivant que l'on pondère ou non par la surface (Figures 2 et 3).

20. L'indice de Gini est un cas très particulier d'un indice de polarisation à la Esteban Ray (1994). Il serait intéressant de calculer des indices de polarisation qui tiennent compte des distances inter-régionales en utilisant, par exemple, les outils proposés par Marcon et Puech (2003).

Figure 2
**Indice de Gini des densités de population régionales
(pondérées par la superficie régionale)**

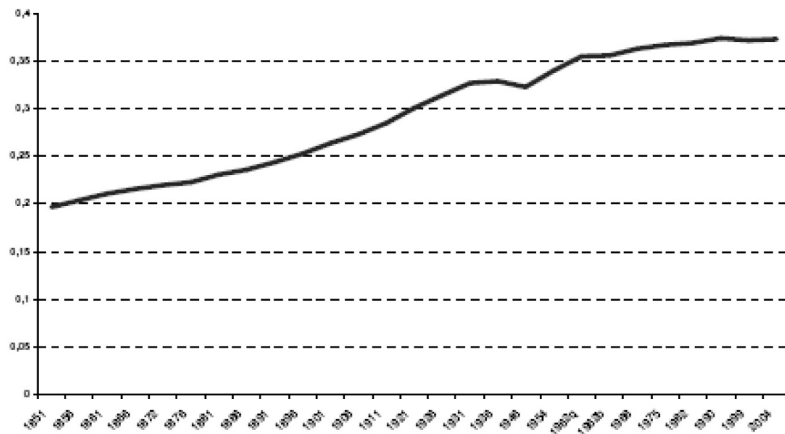
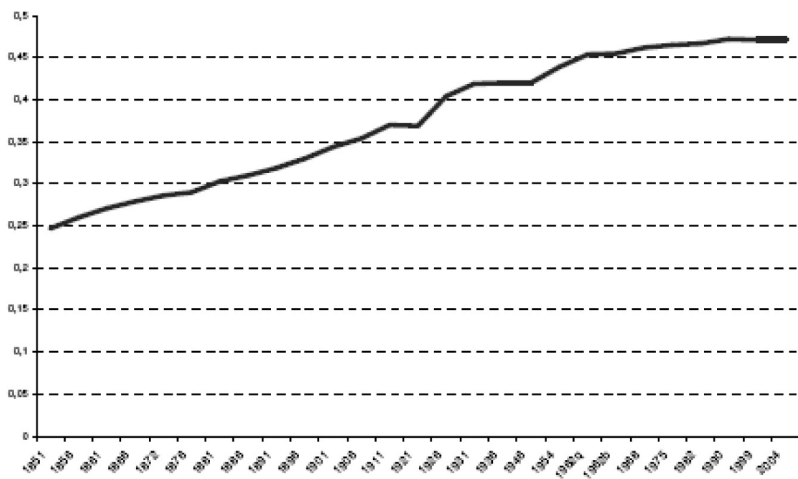


Figure 3
Indice de Gini des densités régionales de population (non pondérées)



La pondération par la surface est *a priori* la pondération idoine, car une répartition égalitaire de la population correspondrait bien à une densité constante sur tous les km² du territoire²¹. Cet indice a pratiquement doublé en un siècle et demi. Deux phénomènes marquants dans ce mouvement de hausse continue : l'arrêt et même la diminution du phénomène de concentration spatiale inter-régionale de la crise de 1929 à la sortie de la seconde guerre mondiale, marquée par l'exode de 1940 et le retour à la terre promu par le régime de Vichy et l'arrêt de la progression de la concentration depuis le début des années 1990. Il semble que l'on ait atteint un point d'équilibre entre forces d'agglomération et forces de dispersion, peut-être le point de retournement du U cher aux économistes-géographes. L'avenir dira si ce phénomène est durable. La période récente semble en tout cas marquée par une sorte de croissance homothétique des régions denses et vides, au moins en raisonnant d'une façon anonyme. Curieusement, le même constat peut être dressé pour l'évolution de la densité des départements (Figure 4). En effet, on aurait pu s'attendre, avec un partitionnement 5 fois plus fin, d'abord à une inégalité bien plus forte, ce qui n'est pas le cas, ensuite, à plus grande permanence du phénomène de polarisation. Il est toutefois intéressant de noter que la phase de stabilisation ne commence au niveau des départements qu'en 1999.

Quand la polarisation en termes de préférence augmente, il en est de même pour le risque de sécession (Haimanko *et alii*, 2005). Ici la localisation s'identifie à la préférence en termes d'accessibilité. Une société polarisée spatialement a des préférences polarisées en termes de bien public, ici le réseau de transport. C'est donc *a priori* plutôt une bonne chose que la polarisation se stabilise pour les conflits d'intérêt spatiaux en France.

Au vu de ces graphiques, on peut se demander si les émois qui entourent les CIADT ne sont pas plus un écho du passé qu'un témoignage résultant d'une saine appréciation et compréhension des inflexions récentes dans les décisions de localisation des ménages. En effet, l'évidence empirique accumulée à travers trois tentatives pour essayer de documenter la possible inégalité territoriale induite à court ou long terme par l'infrastructure de transport n'est pas de nature à susciter un sentiment d'inquiétude exagéré quant à la situation des régions peu denses. La thèse selon laquelle il faudrait compenser les

21. La pondération par la population correspondrait à une vision différente. Cela serait l'inégalité de densité ressentie par les individus. Cette façon de pondérer serait intéressante dans d'autres contextes, par exemple, celui de la mesure de l'inégalité du sentiment de congestion.

sion sur l'équité territoriale. Elle propose un tri parmi les raisonnements en termes d'équité spatiale entre ceux qui peuvent s'appuyer sur une démarche rationnelle et ceux qui en sont dépourvus. Elle permet donc de dégager un langage et des raisonnements qui, s'ils sont partagés par toutes les parties prenantes, ouvrent la voie à un compromis partagé par tous parce que basé sur une rationalité commune.

Cela dit, nous voudrions pointer une faiblesse actuelle dans le processus de décision des infrastructures publiques en France. Alors même que la multiplicité des acteurs a changé la donne, le processus de décision reste entièrement piloté par l'administration et donc par l'État central. Dans une certaine mesure, cela permet de préserver autant que faire se peut l'intérêt général. Mais cela autorise également les surenchères de la part des régions qui n'ont pas complètement abandonné une logique de guichetier. Au fond, le souci d'employer au mieux les deniers publics devrait être également le souci des édiles régionaux. S'ils dépensent l'argent de leurs contribuables sur des projets d'infrastructures peu utilisées, leur réélection n'en sera pas facilitée. La sanction politique d'un gaspillage local de fonds publics ne devrait pas être moins lourde au niveau local qu'au niveau national. Sauf que, si le projet est de toute façon très majoritairement financé par le contribuable national, le gaspillage est moins ressenti pour le contribuable local alors que le bénéfice reste visible. Le schéma de décision et de financement actuel entretient donc une certaine irresponsabilité de la part des édiles régionaux, tant il encourage ceux-ci à faire payer par le contribuable national des projets peu rentables qui ne seraient pas entrepris et financés sur une base locale. Par exemple, le barreau de ligne TGV entre Poitiers et Limoges qui se fixe en râteau sur la ligne Sud-Europe-Atlantique-Paris-Bordeaux serait-il entrepris par les conseils régionaux Poitou-Charentes et Limousin si seules ces deux collectivités territoriales en assuraient le financement ? On peut gager que ces deux collectivités seraient alors intéressées au partage du savoir concernant les études de rentabilité socioéconomique, au lieu d'en faire fi comme maintenant. Par conséquent, la nouvelle donne entre l'État et les régions appelle à un nouveau processus de décision peut-être plus proche de ce qu'un état fédéral comme l'Allemagne a mis en place (Rothengatter, 2000). Sans qu'un tel conseil puisse être pris au pied de la lettre, car il demanderait à connaître de l'intérieur comment les choses se passent, les régions seraient invitées à prendre toute leur part dans la phase d'évaluation du projet et à en faire sienne les conclusions, de telle manière que la décision soit vraiment copartagée au terme d'un processus. Ainsi, si la question d'acceptabilité appelle à élargir le champ de réflexion quant à l'impact à long terme de

l'infrastructure de transport, elle amène également à s'interroger sur un processus de décision où les régions sont devenues suffisamment puissantes pour contester la logique des décisions tout en pouvant continuer à adopter un comportement d'assistantat qui peut friser l'irresponsabilité.

Références

- Alesina A. et Spolaore E., *The Size of Nations*, MIT Press, Cambridge, MA, 2003.
- Benoît J.-L., *Tocqueville, un destin paradoxal*, Bayard, Paris, 2005.
- Bonnafeous A. et Masson S., « Politique des transports et équité spatiale », *Revue d'économie régionale et urbaine*, 4, 2003, p. 547-572.
- Blackorby C., Bossert W. et Donaldson D., *Population issues in social choice theory, Welfare economics and ethics*, Cambridge University Press, 2005.
- Combes P.-P. et Lafourcade M., « Transport Costs : Measures, Determinants and Regional Policy Implications for France », *Journal of Economic Geography*, Vol.5 (Issue 3), 319-3, 2005.
- Combes P.-P. et Lafourcade M., « Coûts de transport et inégalités régionales : une approche structurelle », *Revue Économique*, 53 (3), 2002, p. 505-516.
- Combes P.-P. et Lafourcade M., « Distribution spatiale des activités et politiques d'infrastructures de transport : l'économie géographique, un nouvel outil d'évaluation ? », dans *Aménagement du territoire*, Rapport n° 31 du Conseil d'Analyse Économique, La Documentation Française, Paris, 2001.
- Combes P.-P., Duranton G. et Gobillon L., « Origine et ampleur des inégalités spatiales de salaire en France », complément C du rapport n° 40 *Compétitivité*, Conseil d'Analyse Économique, 2003, p.163-183.
- Charlot S. et Lafourcade M., « Infrastructures publiques, coûts de transport et croissance régionale », dans *Économie Géographique, la théorie à l'épreuve des faits*, C. Beaumont, P.-P. Combes, H. Derycke et H. Jayet (éd.), Economica, Paris, 2000.
- Cornette J., *Histoire de la Bretagne et des Bretons*, Tome I et II, Seuil, Paris, 2005.
- Deyon P., *Paris et ses provinces*, Armand Collin, Paris, 1992.
- Esteban J.-M. et Ray D., « On the Measurement of Polarisation », *Econometrica*, Vol. 62, 1994, p. 819-851.
- Fleurbaey M., *Théories économiques de la justice*, Economica, Paris, 1996.

- Flory T., *Le mouvement régionaliste français, sources et développements*, PUF, Paris, 1996.
- Furet F., *La Révolution*, Tome I et II, Hachette, Paris, 1997.
- Furet F., Ozouf M., *La Gironde et les Girondins*, Payot, Paris, 1991.
- Gérard-Varet L.-A. et Thisse J.-F., « Économie publique locale et économie géographique », *Annales d'Économie et Statistique* 45, 1997, p. 1-18.
- Gobillon L., « Emploi, logement et mobilité résidentielle », *Économie et Statistiques*, n° 349-350, 2001, p. 77-98.
- Gobillon L. et Le Blanc D., *Migrations, incomes and unobserved heterogeneity*, Mimeo, 2003.
- Gravier J.F., *Paris et le désert français*, Paris, 1947.
- Habermas J., *Raison et légitimité : problèmes de légitimation dans le capitalisme avancé*, Payot, Paris, 1978.
- Haimanko O., Le Breton M. et Weber S., « Transfers in a polarized country : bridging the gap between efficiency and stability », *Journal of Public Economics* 89, 2005, p. 1277-1303.
- Harsanyi J., « Cardinal utility in welfare economics and in the theory of risk-taking », *Journal of Political Economy* 61, 1953, p. 309-321.
- Harsanyi J., « Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility », *Journal of Political Economy* 63, 1955, p. 309-321.
- Jayet H., « L'analyse économique des migrations : une synthèse critique », *Revue Économique*, 47, 1996, p. 193-226.
- Lafourcade M., *L'impact des infrastructures de transport sur la localisation des activités et la croissance locale*, Thèse Paris 1, 1998.
- Le Breton M. et Weber S., « The art of making people happy : how to prevent a secession », *IMF Staff Papers* 50, 2003, p. 403-435.
- Madival J. et Laurent E., *Archives parlementaires de 1787 à 1860 : recueil complet des débats législatifs et politiques des chambres françaises*, 1^{re} série : 1789-1799, Paris, Librairie Administrative de Paul Dupont, 47 Volumes, 1867.
- Marcon E. et Puech F., « Evaluating the Geographic Concentration of Industries Using Distance-Based Methods », *Journal of Economic Geography* 3, 2003, p. 409-428.
- Moulin H., *Fair division and collective welfare*, MIT Press, 2003.
- Meyer B., « Natural and Quasi-Natural Experiments in Economics », *Journal of Business and Economic Statistics*, 12, 1995, p. 151-162.
- Ozouf-Marignier M.-C., « De l'universalisme constituant aux intérêts locaux : le débat sur la formation des départements en France (1789-1790) », *Annales ESC*, t. 41, 1986, p. 1193-1213.
- Parfit D., *Reasons and Persons*, Oxford University Press, Oxford, 1984.

- Quinet E., « Evaluation methodologies of transportation projects in France », *Transport Policy* 7, 2000, p. 27-34.
- Rapport Boiteux I, *Transport et choix des investissements et coûts des investissements*, Rapport du CGP, La Documentation Française, Paris, 1994.
- Rapport Boiteux II, *Transport et choix des investissements et coûts des investissements*, Rapport du CGP, La Documentation Française, Paris, 2001.
- Rawls J., *Theory of Justice*, Harvard University Press, 1974.
- Roberts K., « Valued opinions or opinionated values : the double-aggregation problem », in Basu K., Patanaik PK. et Suzumura K. (eds), *Choice, welfare and development : a festschrift in honour of Amartya Sen*, 141-165, Oxford University Press, 1995.
- Roemer J., *Theories of distributive justice*, Harvard University Press, Cambridge, Mass, 1996.
- Rosanvallon P., *Le modèle politique français, la société civile contre le jacobinisme de 1789 à nos jours*, Seuil, Paris, 2004.
- Rota Graziosi G., « La fragmentation politique, une analyse de la littérature », *Revue française d'économie*, 18, 2004.
- Rothengatter W., « Evaluation of infrastructure investments in Germany », *Transport Policy* 7, 2000, p. 17-25.
- Sauvant A. et Salager J., « L'accessibilité des villes et des territoires aux réseaux de transport », *Note de synthèse du SES*, 2002.
- Sen A.K., « On weighs and measures : informational constraints in social welfare analysis », *Econometrica*, 45, 1977, p.1539-1572.
- Thisse J., « Équité, efficacité et acceptabilité dans la localisation des équipements collectives », 2007, Chapitre 12 dans ce volume.
- Toutain J.-C., « Le produit intérieur brut de la France en 1860 et 1930 par départements et par régions », *Économies et Sociétés, Série Histoire Économique Quantitative*, à paraître.
- Trannoy A., « À propos des évidences apportées par des expériences naturelles », *économie Publique*, 2003, 13, p. 159-163.

CHAPITRE 14

LES FIGURES DE L'ACCEPTABILITÉ

Charles Raux, Stéphanie Souche, Lucie Vaskova¹

L'acceptabilité des politiques de transport reste une « expression-valise » aux contours assez flous, où se pressent des définitions multiples et variées, malgré (ou à cause de ?) l'abondance des projets européens consacrés à la question, et de la littérature qui s'en est ensuivie. Ce courant de recherche assez récent fait suite aux prises de position de la Commission Européenne quant à la politique de transport et notamment en faveur de la tarification au coût marginal social (CE, 1998). Dans cette littérature on trouve souvent des inventaires à la Prévert des différents critères d'acceptabilité, où l'on a assez de mal à « faire son marché » pour aider à la prise de décision en matière de politique publique de transport.

L'expérience des débats publics autour des projets d'infrastructure ou de mise en place de changements tarifaires, que ce soit l'augmentation chaque année des tarifs des transports collectifs ou, plus exceptionnellement, la mise en place d'un péage urbain, montre que les arguments utilisés relèvent de l'équité ou plus trivialement de la justice, et rarement de l'efficacité économique supposée de ces projets.

On écartera d'emblée la vision cynique, qui consisterait à dire que tous ces débats autour de la « justice » ne sont qu'un rideau de fumée utilisé par les porteurs d'intérêts, qui utiliseraient les arguments de l'intérêt public pour justifier le leur. Cette vision cynique est réductrice car souvent ces porteurs d'intérêts sont sincères. Il faut donc prendre acte d'une justice perçue qui est une donnée de la politique publique.

Il est trivial de constater que ce qui semble efficace sur le plan économique (la raison d'être du calcul économique) n'est pas nécessairement considéré comme juste. À partir des travaux de Zajac (1995) nous pouvons identifier au moins cinq principes de justice positive qui pourraient s'appliquer avec pertinence au cas de la politique publique

1. Laboratoire d'Économie des Transports, CNRS, Université Lyon 2, ENTPE.

de transport : (1) bien qu'il n'y ait que peu de référence explicite à l'efficacité économique dans les débats publics, certaines inefficacités sont perçues comme injustes, surtout si elles consistent en une perpétuation de privilèges pour certains groupes qui s'opposent à leur remise en cause ; (2) dans un papier fondateur, Kahneman, Knetsch et Thaler (1986) ont établi à l'aide d'un support empirique l'existence d'un droit sur les termes de la « transaction de référence », i.e. les droits issus du statu quo : le sentiment d'injustice apparaît lorsque la décision de l'entreprise ou des pouvoirs publics entraîne un changement de contexte qui ne tient pas compte de la transaction de référence, formée par l'histoire des transactions précédentes entre l'entreprise ou la puissance publique d'une part, et l'employé, le consommateur ou l'utilisateur d'autre part. Dans le cas du transport, cette transaction de référence sera par exemple la gratuité d'usage de la route² ou le droit à ne pas voir débarquer une autoroute au fond du jardin de sa maison de campagne ; (3) on s'attend à ce que la société nous protège contre les conséquences négatives de changements économiques sur lesquels nous n'avons pas pris ; (4) on exige un droit de regard sur le service public quand il s'agit d'un monopole, surtout quand il sert ce qui est considéré comme un droit (par exemple le transport public) ; (5) on réclame l'égalité de traitement : cette égalité fait référence aux différentes dimensions de l'équité (verticale, horizontale et spatiale) que nous développons ci-après.

Ces principes vont être plus ou moins pertinents selon les situations, et parfois même contradictoires : comment préserver le statu quo tout en remettant en cause les privilèges pour certains groupes ? À travers une revue des résultats de travaux de recherches européennes et en nous aidant de la riche littérature académique sur les questions de justice et d'équité en tant que de besoin, nous tenterons de mettre un peu d'ordre dans une problématique, l'acceptabilité des politiques publiques dans les transports. Autrement dit, nous nous proposons de construire une sorte de grille de lecture.

L'état de l'art sur l'acceptabilité des politiques de transport peut s'organiser autour de quelques grands thèmes. Pour comprendre l'acceptabilité, il faut d'abord la définir (section 1). Mais l'opération est délicate. L'utilisation de ce terme sert à décrire des attitudes et des comportements d'acteurs souvent empreints de subjectivité, ce qui en

2. en apparence seulement car dans les faits cette gratuité n'en est pas une : en effet l'automobiliste paye des impôts et des taxes spécifiques à cet usage, mais le plus souvent indifféremment des lieux et moments de cet usage. Ce qui est en question est la remise en cause de l'apparente gratuité à travers la tarification routière sans échappatoire comme certains types de péage urbain.

rend difficile l'évaluation. Il convient d'ailleurs d'identifier ces acteurs (section 2) : parle-t-on de l'acceptabilité d'une nouvelle mesure par les automobilistes ou les usagers des transports collectifs, ou par les producteurs de services de transports (opérateurs) ou encore par d'autres acteurs ? Afin d'établir si les conditions de mise en œuvre de la nouvelle mesure sont favorables à la réussite du projet, il devient nécessaire de spécifier le contexte et la manière dont un « problème » est perçu (section 3). La question de la perception va également intervenir dans l'appréciation des solutions proposées par les acteurs (section 4). Elle s'appuie sur des arguments d'équité et de justice qui prennent souvent le pas sur ceux de l'efficacité au sens économique. Pour illustrer ce traitement de l'acceptabilité, nous faisons ensuite un bref exposé de trois expériences emblématiques de péage urbain (Oslo, Londres, Lyon) (section 5). Nous terminons enfin par une proposition de mise en ordre conceptuelle, mais provisoire, du processus menant à l'acceptabilité (section 6).

1. ACCEPTABILITÉ ET ACCEPTATION : UNE TENTATIVE DE COMPRÉHENSION ET DE DÉFINITION

Même si la question de l'acceptabilité fait l'objet de nombreux travaux de recherche, sa définition semble assez floue. Les projets européens PRIMA et CUPID cherchent à intégrer la question de l'acceptabilité dans l'étude de faisabilité et dans l'évaluation *ex ante* d'un projet. Ils ont une conception très large et générale de l'acceptabilité. Pour d'autres projets comme AFFORD et MC-ICAM, c'est plutôt le refus *ex post* – la non acceptation – qu'il faut comprendre puisqu'elle peut faire échouer la réalisation d'un projet. Notre objectif est ici de tenter de donner une définition de l'acceptabilité, ou tout au moins d'en préciser les contours à travers en particulier la distinction entre une mesure dite acceptable (*versus* non acceptable) et une mesure dite acceptée (*versus* non acceptée).

L'acceptabilité peut d'abord se comprendre dans le sens de ce qui est ou non acceptable *ex ante*. Avant qu'une nouvelle mesure ne soit expérimentée, les acteurs peuvent déjà porter sur elle un jugement favorable ou défavorable, en la jugeant acceptable ou non. Dans la littérature, on trouve l'expression « d'acceptation subjective » (*attitudinal acceptance* dans le projet PATS) ou « d'acceptabilité » d'une nouvelle mesure (Schade, 2003 ; Schade et Schlag, 2003).

L'acceptabilité peut ensuite être abordée sous l'angle de ce qui est ou non accepté, une fois la mesure mise en œuvre. Cette attitude *ex-post* peut être désignée par l'expression « d'acceptation comporte-

mentale » (*behavioural acceptance* dans le projet PATS) ou « d'acceptation » (Schade, 2003 ; Schade et Schlag, 2003). Dans cette logique, il est possible de définir l'acceptation comportementale comme « une absence de résistance à la mise en œuvre d'un système » (Van der Laan, 1998, p. 39, in Schade et Schlag, 2003).

Finalement, il semble se dégager deux pistes pour mieux comprendre l'acceptabilité : la première cherche à comprendre qui accepte ou n'accepte pas et s'intéresse donc aux acteurs du processus d'acceptation, l'autre se demande sous quelles conditions l'on accepte ou l'on n'accepte pas telle ou telle mesure.

2. LES ACTEURS CONCERNÉS

Les acteurs sont au cœur de la question de l'acceptabilité, encore faut-il savoir de quels acteurs il s'agit. Pour certains travaux, seuls les avis des automobilistes ou des passagers, c'est-à-dire des utilisateurs de transport, sont pris en compte (projet AFFORD). Pour d'autres il faut tenir compte non seulement des citoyens en général (Odeck et Brathen, 1997), mais également des opérateurs, des politiciens, et des groupes d'intérêt (CUPID, PATS).

Pour faciliter cette lecture complexe des jeux d'acteurs de la décision politique et de leurs arguments, l'approche proposée par Weck-Hannemann (1999) dans le cadre de la recherche européenne sur l'acceptabilité du péage (PATS, 2001) revêt une grande utilité. En s'inspirant d'une analyse du marché politique, l'auteur étudie les réactions des divers acteurs aux mesures de politique environnementale dans le secteur des transports. Elle identifie les différents acteurs – à savoir les politiques, les groupes d'intérêts³ et les citoyens-électeurs – et les organise selon qu'ils relèvent plutôt de l'offre ou de la demande de transport.

L'acceptabilité par les politiques : la théorie dite du Public Choice, initiée dans les années soixante par J. Buchanan et G. Tullock, étend au champ politique le raisonnement économique en terme de marché. La structure du vote (en particulier la date de la prochaine élection) et l'existence de groupes de pression, seraient susceptibles d'influencer les procédures de prise de décision publique. Ainsi les politiques eux-mêmes trouveraient intérêt à accroître leur pouvoir discrétionnaire et à affaiblir la contrainte budgétaire, en favorisant par exemple des

3. Qui désignent aussi bien les victimes d'externalités négatives (exemple, les associations de riverains d'une infrastructure en projet) que les bénéficiaires d'externalités positives (exemple, les organisations de transporteurs ou de chargeurs).

outils orientés sur la productivité fiscale. On peut supposer qu'ils préfèrent les revenus qui doivent aller dans le budget public général, leur fournissant la flexibilité d'usage de ces ressources additionnelles au lieu d'avoir à les retourner aux citoyens ou aux contribuables, sous forme de transferts d'argent ou de réduction de taxes. Cependant, cette politique n'est possible que lorsque la contrainte de réélection n'est pas trop forte. À l'approche des échéances électorales, le pouvoir discrétionnaire des politiciens diminue et ils doivent alors prendre en compte les intérêts des électeurs.

L'acceptabilité par les groupes d'intérêt : selon ce point de vue, ce sont les instruments en faveur des groupes les mieux organisés, qui seront favorisés au détriment des autres groupes moins structurés. Il est clair que les usagers individuels (automobilistes ou passagers des transports collectifs) sont dans une situation comparative défavorable par rapport aux opérateurs ou producteurs, quant à la capacité de s'organiser en groupes d'intérêts. Il en est de même pour les contribuables. Les instruments tarifaires peuvent servir l'intérêt de groupes différents selon l'usage des revenus, affectés par exemple au financement d'infrastructure ou à la fourniture de services. Les instruments tarifaires ou fiscaux sont d'autant plus acceptés par les groupes d'intérêts qu'ils leur permettent de modérer l'incitation à changer de comportement, d'obtenir des exonérations, de faire supporter la charge financière par d'autres, et d'affecter les recettes pour en retirer les bénéfices.

L'acceptabilité par les citoyens et électeurs : cette question renvoie à la résistance, généralement observée dans l'opinion, à la généralisation de l'usage des mécanismes de prix pour gérer des biens relevant de la sphère collective. Cette résistance recouvre les aspects de la redistribution des revenus, de l'équité du système de prix pour rationner la demande, et de l'éthique de ce système pour gérer des biens environnementaux non renouvelables.

Pour chaque groupe d'acteurs participant au processus de décision, il est alors possible de repérer leurs intérêts spécifiques et leur influence sur les résultats du processus de prise de décision, de façon à connaître qui sont les gagnants et perdants des mesures à mettre en place.

3. LE CONTEXTE ET LA PERCEPTION DU PROBLÈME

À travers la notion de contexte, c'est le cadre socio-économique, législatif, institutionnel, culturel, ainsi que les habitudes qui préexistent à la mise en œuvre de la mesure qui sont pris en compte. Ce

contexte n'est pas immuable, il est conditionné par l'évolution de longue durée. Les résistances semblent se réduire au cours du temps comme le montre l'exemple du péage à Oslo où la résistance s'est réduite à partir du moment où le système est devenu opérationnel (cf. *infra*).

Il apparaît que la connaissance du cadre juridique – qui borne l'éventail des possibles – et des pratiques de mobilité, est une des conditions nécessaires à la mise en œuvre d'une mesure : par exemple, en France, la loi n'autorise pas la mise en place d'un péage urbain sur une infrastructure existante. Par ailleurs, les informations sur les caractéristiques socio-démographiques des individus (niveaux de revenu et d'éducation, âge, sexe), sur les comportements de mobilité (motif de déplacement, actif/inactif, mode de transport, localisations résidentielle et professionnelle) sont utiles pour comprendre ce qui influence les réponses des acteurs. Ces caractéristiques vont pour partie déterminer les marges de manœuvre dont les individus disposent.

Une situation législative confuse et peu claire, des expériences négatives, des traditions peu favorables aux démarches proposées peuvent freiner le processus de décision (MC-ICAM). Pour Schade et Schlag (2003) les normes, voire les tensions sociales, influencent l'acceptabilité publique d'une stratégie de tarification. Le contexte intervient également à travers le niveau (local, régional, national, international) de la prise de décision. Si les nouvelles mesures concernent la vie locale, une prise de décision locale est perçue comme plus proche des attentes des individus, et peut ainsi paraître plus acceptable. L'exemple de Rotterdam montre comment l'expérimentation d'une tarification routière unique du Randstaad envisagée par l'échelon national, a été source de conflit entre l'échelon national et régional et a finalement échoué sous la pression des régions métropolitaines qui y étaient opposées (Harsman, 2003).

L'acceptabilité est également liée à la perception ou non d'un problème (PATS, 2000 ; Schade et Schlag, 2003 ; MC-ICAM, CUPID, INPHORMM). Si nous percevons qu'une situation pose problème nous allons chercher à trouver des solutions (Schade et Schlag, 2000). Pour ces auteurs, les habitants les plus contraints et ceux des régions densément peuplées perçoivent avec plus d'acuité les problèmes liés à la mobilité et attendent aussi davantage de solutions. Les résultats du projet AFFORD montrent que la pollution de l'air, la congestion automobile et la pénurie des places de stationnement, sont les contraintes perçues comme les plus importantes. Dans cette optique, une nouvelle mesure dont les objectifs seraient clairement affichés et

dont on percevrait l'efficacité, serait plus acceptable (Schade et Schlag, 2000 ; Weck-Hannemann, 1999).

Rietveld et Verhoef (1998) distinguent la composante individuelle et la composante collective (ou sociale) de la perception d'un problème. Quand les personnes disent être touchées individuellement par un problème, elles lui donnent également une dimension sociale. Pour ces mêmes auteurs, le niveau de revenu *via* la valeur du temps, joue un rôle déterminant dans l'évaluation de la mesure. Précisément, la sensibilité à la congestion serait corrélée positivement avec l'âge, un haut niveau d'éducation et l'utilisation de la voiture particulière.

4. LA PERCEPTION DES SOLUTIONS

Si la question de la perception se pose lorsqu'il s'agit de caractériser l'acuité d'un problème, elle intervient également lorsqu'il s'agit de trouver des solutions.

4.1. *L'efficacité de la mesure et sa perception*

Il faut distinguer l'*efficacité* (l'adéquation entre résultats et objectifs) et l'*efficience* (l'articulation entre moyens et résultats). La tarification est par exemple considérée comme moins efficiente que les mesures de restriction quand il s'agit de réduire la circulation automobile (Schade et Schlag, 2003).

Fernandez et Rodrik (1991) expliquent que si les gouvernements n'adoptent pas certaines réformes, alors que les économistes les préconisent pour accroître l'efficacité de la politique publique, c'est notamment à cause de l'incertitude sur la distribution des gains et des pertes consécutifs à ces réformes : autrement dit, les futurs gagnants ou perdants de la réforme ne peuvent pas être identifiés au préalable. Cette incertitude expliquerait le biais en faveur du statu quo, et ainsi contre toute réforme qui accroîtrait l'efficacité.

Un autre élément d'explication de ce décalage nous ramène à la question de la perception. Il peut en effet exister une différence entre les effets d'une nouvelle mesure affichés par ses promoteurs, et leur perception par les acteurs : une nouvelle tarification – comme l'introduction d'une tarification à la distance – estimée plus efficace par l'opérateur de transports collectifs urbains, peut être perçue négativement par les usagers qui la jugent inéquitable pour ceux de la périphérie. Les arguments d'équité et de justice occupent alors le devant de la scène.

4.2. L'équité de la mesure

Les acteurs ont donc recours aux arguments d'équité et de justice pour débattre de l'acceptabilité d'une nouvelle mesure. Une nouvelle difficulté apparaît puisque l'équité elle-même est polysémique. En suivant la littérature⁴, l'équité semble s'articuler autour des thématiques suivantes : les gagnants et les perdants à la mise en œuvre d'une nouvelle mesure, la redistribution des recettes issues de la réforme tarifaire, et les compensations possibles.

La mise en œuvre d'une nouvelle mesure créerait des gagnants et des perdants. L'introduction d'un changement tarifaire comme le péage urbain serait inéquitable puisqu'elle toucherait d'abord les plus pauvres (Evans, 1992 ; Emmerink *et al*, 1995). Dans cette optique, les travaux de Else (1986), Guiliano (1992) et Langmyhr (1997), mettent en évidence que les gagnants seraient ceux qui bénéficient des plus hauts revenus. Pour Jakobsson *et al* (2000), l'inéquité du péage routier serait d'autant plus forte que ce dernier porterait atteinte à la liberté de se déplacer (selon le mode de transport et le lieu). La liberté de se déplacer serait un droit fondamental selon Schlag et Teubel (1997). L'introduction d'un péage en milieu urbain conduirait également à accroître le coût de la localisation dans les régions ou les villes qui le mettraient en œuvre, créant ainsi une inéquité territoriale (Emmerink *et al*, 1995).

La question de la redistribution des recettes du péage semble jouer un grand rôle dans l'acceptabilité d'une nouvelle mesure. Goodwin (1989) a proposé une règle d'affectation des recettes du péage urbain entre améliorations des routes, de l'offre en transports collectifs et de l'environnement physique urbain ; il a également esquissé (1995) les principes d'une convergence et non plus d'une opposition entre amélioration de l'environnement et efficacité économique. Small (1992) a proposé une stratégie de distribution des impacts de programmes financés avec les recettes du péage qui intègre les différents groupes d'intérêt et qui verse une compensation aux usagers affectés. D'autres auteurs voudraient que ces recettes soient utilisées pour les modes les plus respectueux de l'environnement (Schlag et Teubel, 1997). Pour Harrington *et al* (2001), la redistribution des recettes peut prendre la forme d'une compensation *via* la diminution des taxes existantes. Thorpe (2002) s'intéresse à la capacité libérée sur les infrastructures existantes grâce à la réduction du trafic. Il suggère que le bénéfice

4. Pour un survey sur l'équité et la question du péage urbain voir Souche (2003). Pour une approche normative de la justice, et de la justice comme équité en particulier, sans application au cas spécifique des transports voir Fleurbaey (1996).

d'une allocation en faveur des mesures d'amélioration de l'environnement (voies réservées pour les bus notamment) soit ajouté aux gains de temps dont bénéficient les automobilistes qui continuent à utiliser l'infrastructure à péage.

On en vient naturellement à la question de la compensation. Parmi les principes de justice économique qu'énonce E. Zajac – évoqués en introduction –, il y a la revendication d'un droit à l'assurance : on s'attend à ce que la collectivité assure les individus contre des pertes entraînées par un changement des conditions économiques et, si ce n'est pas le cas, cette situation est considérée comme injuste.

On sait depuis les travaux de Hicks et Kaldor, qu'une compensation hypothétique suffit à justifier une politique dès lors que les avantages des gagnants l'emportent sur les pertes des perdants. On peut donc *a priori* penser qu'une compensation effective contrebalancerait le rejet du mécanisme d'allocation par les prix.

Néanmoins il apparaît que ce principe de compensation serait rejeté au motif que les « perdants compensés » ont l'impression que l'on « achète » leurs votes pour permettre aux riches de profiter des biens rendus ainsi disponibles (Frey *et al.*, 1996 ; Kunreuther and Easterling, 1996). C'est pourquoi Oberholzer-Gee et Weck-Hannemann (2002) suggèrent que la compensation doit ressembler le moins possible à un achat de votes et doit viser à récompenser, dans la même « dimension » que la perte, ceux qui acceptent de réduire leur consommation du bien rare parce qu'ils contribuent à l'amélioration du bien-être collectif. L'une des conséquences est qu'il est nécessaire que les sommes qui sont prélevées dans le secteur des transports soient utilisées pour des compensations effectuées dans le transport et non pas dans un autre secteur.

Enfin, il faut faire face au soupçon d'un enrichissement indu du gouvernement sur le dos des automobilistes (Emmerink *et al.*, 1995). L'un des principes de justice économique énoncé par Zajac est relatif au pouvoir abusif du monopole : « *The fewer the substitutes for a regulated or monopoly firm's output, and the more the output is considered an economic right, the more the public expects to exert control over the firm. Denial of control is considered unjust* » (Zajac, 1995, p. 127).

À partir des principes explicités par Rawls dans sa théorie de la justice (1971), Raux et Souche (2004) ont développé un cadre d'analyse de l'acceptabilité des changements dans le secteur des transports, qui combine l'efficacité économique (orienter efficacement la demande) avec trois dimensions de l'équité : l'équité verticale au sens de l'attention à porter aux personnes les plus fragiles économiquement ; l'équité horizontale au sens de la répartition de la charge entre

contribuables et usagers des transports, et entre usagers des différents modes ; enfin et surtout, l'équité territoriale qui se réfère à la garantie d'accessibilité aux aménités, et à la liberté d'aller et venir. Une première démarche peut être alors d'identifier les groupes de perdants ou de gagnants selon ces trois dimensions.

4.3. *La perception de la justice*

Ce que l'on perçoit des caractéristiques d'un projet permet de comparer les différentes alternatives et de les rendre ainsi plus ou moins acceptables.

Schade et Schlag (2000) estiment que la justice perçue peut être assez différente de la distribution réelle des coûts et des avantages d'une nouvelle mesure. Il semble exister deux niveaux de perception de l'équité quand on compare l'évolution de notre situation entre avant et après la mise en œuvre de la nouvelle mesure. Le niveau d'équité intrapersonnelle peut être résumé par la question suivante, « ma situation personnelle est-elle ou non améliorée par la nouvelle mesure ? ». Le niveau d'équité interpersonnelle revient à poser la question, « ma situation personnelle est-elle plus ou moins améliorée – ou détériorée – que celle des autres par la nouvelle mesure ? ».

Baron et Jurney (1993) dissocient le consentement à voter pour une mesure et le jugement que l'on porte sur cette mesure. Frey et Jegen (2001) nous permettent de comprendre pourquoi. Ils expliquent en effet, que les attitudes sont différentes des comportements : il est tout à fait possible d'afficher une sensibilité aux questions environnementales, tout en utilisant chaque jour son automobile. Toutefois, Baron et Jurney (1993) affirment que sur les décisions qui conduisent à un dilemme social, c'est-à-dire une situation où les préférences individuelles sont différentes de la préférence collective, il faut exercer un pouvoir de coercition. Dans cette logique, les auteurs du rapport européen TAPESTRY (2001) identifient différentes phases à respecter pour qu'un changement de comportement se produise à l'issue d'une campagne publicitaire. Ils insistent en particulier sur : l'identification des comportements, des attitudes et de l'acceptation avant la campagne publicitaire ; l'identification des groupes cibles ; la mesure des changements dans les comportements, les attitudes et l'acceptation après la campagne ; l'identification du contexte et des facteurs d'influence externes.

De plus, il semble qu'une nouvelle mesure soit mieux acceptée si nous avons le choix entre plusieurs alternatives (Schade et Schlag, 2000).

Le degré d'acceptabilité est également influencé par l'attitude des individus par rapport aux alternatives et aux mécanismes de régulation qui peuvent être proposés. Frey et Pommerehne (1993) se sont posé la question suivante : pourquoi la régulation d'une demande excessive par les prix est-elle considérée comme injuste alors que les économistes préconisent ce principe ? Face à une situation de rareté en eau, ils identifient et testent plusieurs procédures d'allocation des ressources qui renvoient à autant de principes de justice. Ils trouvent que pour rationner l'excès de demande, une procédure traditionnelle de file d'attente ou une procédure administrative sont jugées plus justes qu'une procédure tarifaire qui consiste à faire payer plus cher les ressources devenues rares. Ils en concluent que les économistes devraient inclure des aspects moraux ou éthiques dans leur analyse s'ils veulent améliorer leur capacité à conseiller des politiques. Dans un contexte de bien public, Frey et Oberholzer-Gee (1996) montrent également que, quand il s'agit de trouver un lieu d'implantation pour un bien public localement non désiré, la compensation ne suffit pas et que les procédures de décision perçues comme justes jouent un rôle essentiel.

5. TROIS EXPÉRIENCES : OSLO, LONDRES, LYON

Nous avons choisi trois cas emblématiques de réussites et de défaillances, sur le thème critique du péage urbain, à savoir les expériences d'Oslo, de Londres et de Lyon. Pour chacun de ces cas, nous esquissons le contexte et les acteurs, d'abord au niveau national, puis au niveau local, puis nous soulignons certains aspects qui ont pu influencer l'acceptabilité de façon positive et négative.

5.1. *Le péage de cordon d'Oslo*

Le péage de cordon d'Oslo, qui a été mis en fonctionnement en 1990, est souvent présenté comme un exemple de réussite. Son but était de financer un système de rocade et de tunnels routiers, et par là même de décongestionner le centre ville : le péage est donc appliqué sur plusieurs voies d'accès à la zone centrale, où réside la moitié de la population.

5.1.1. *Le contexte national*

La Norvège a une longue tradition de financement des investissements dans le domaine de transport au moyen de redevances. Selon Odeck et Brathen (1997) plus de 26 % du budget pour la construction

d'infrastructures provient des recettes de péage. Cela s'explique par la topographie compliquée de ce pays qui implique des coûts très élevés de construction des routes (Certu *et al*, 2002).

Il n'existe pas de système de concession en Norvège, et c'est l'administration publique des routes qui s'occupe de tout le processus de planification, et, au moyen de filiales contrôlées par les pouvoirs publics, de la construction et de l'exploitation des projets financés par le péage. La loi norvégienne rend possible la mise en place d'une tarification routière sur les réseaux existants pour financer une infrastructure nouvelle à intégrer dans ce réseau. En outre, il n'est pas nécessaire que la nouvelle infrastructure préexiste pour pouvoir introduire le péage. Les péages urbains norvégiens sont mis en place pour une durée limitée. Les projets débutent souvent par des initiatives locales et sont ensuite approuvés par le Parlement.

Jusqu'en 2002, la seule forme de péage urbain autorisée était un péage de financement avec réinvestissement des recettes dans les infrastructures routières. La loi autorise désormais le péage de régulation par le biais de la tarification de la congestion, mais ce type de péage n'a pas encore été mis en place en Norvège.

5.1.2. *Le contexte local*

Du fait du relief assez compliqué d'Oslo, les accès au centre ville sont limités à trois corridors et la congestion routière devenait de plus en plus problématique déjà dans les années 70 – 80. En 1984 le Parlement Norvégien s'est engagé à accroître les subventions dans le cas où les collectivités locales adopteraient le plan de transport proposé par la Direction des Routes, dont les coûts seraient financés par péage (Certu *et al*, 2002). La discussion politique fut assez ardue et deux variantes principales furent débattues : les conservateurs souhaitaient que seule la nouvelle infrastructure (le tunnel) soit payante, les travaillistes étaient pour un péage de zone. Le compromis fut un péage de cordon. Les écologistes étaient contre le développement des routes mais ont accepté le compromis du moment que les automobilistes devaient payer.

Les industries, les entreprises et les opérateurs liés au transport étaient également favorables au péage. Même si le péage urbain était une mesure impopulaire, les gouvernements locaux en dehors d'Oslo étaient motivés par sa mise en œuvre car 40 % des revenus du péage sont investis en faveur du comté d'Akershus, en périphérie d'Oslo, ce qui correspond à la proportion des résidents de ce comté qui utilisent le réseau d'Oslo (Certu *et al*, 2002).

5.1.3. *La problématique de l'acceptabilité*

En 1989, un an avant la mise en œuvre du péage à Oslo, les sondages montraient une opposition majoritaire (65 % contre, 28 % pour). Au moment où le système est devenu opérationnel, les résultats furent tout de suite visibles. Il faut aussi souligner que le tunnel d'Oslo (2 km) qui assure le principal trafic de transit est-ouest (sous le centre-ville et le port) fut ouvert quelques semaines avant la mise en service du péage, rendant visible aux automobilistes les effets positifs du programme à venir. La congestion diminua de 20 % (Certu *et al*, 2002) et en conséquence la résistance s'est amoindrie : en 1991, seuls 57 % des sondés étaient contre et 36 % pour (Tretvik, 2003). En outre, le niveau de tarif du péage est relativement bas (Certu *et al*, 2002).

Pourquoi les projets urbains ont-ils été acceptés dans les quatre principales villes du pays⁵ ? Cinq raisons essentielles peuvent être avancées :

Les avantages de ces programmes d'amélioration des routes étaient évidents pour tout le monde et en particulier les automobilistes.

- Les opposants au développement du réseau routier appréciaient que les automobilistes aient à payer pour circuler en ville et qu'une part du financement global du programme aille aux transports publics.

- Les recettes de péage étaient complétées par des fonds gouvernementaux qui autrement n'auraient pas été obtenus.

- Les partis politiques s'accordèrent pour ne pas faire de cette question un enjeu de controverse politique entre eux.

- Le péage était mis en œuvre pour un objectif limité (financer un programme d'infrastructures précis) et donc avec une limite dans le temps.

Bergen (qui a ouvert son péage en 1986) a déjà décidé la prolongation. La discussion est en cours à Trondheim (échéance 2005) et Oslo (échéance 2007).

5.2. *Le péage de zone de Londres*

Le péage urbain de Londres est un autre cas qui peut être considéré comme une application réussie d'un système de tarification. Le système de *Congestion Charging* est en œuvre depuis février 2003. Son objectif principal était de réduire la congestion par l'instauration du péage et par l'amélioration des transports publics à l'aide des recettes du péage.

5. Bergen (1986), 240 000 hab ; Oslo (1990), 800 000 hab ; Trondheim (1991), 180 000 hab ; Stavanger (2001), 110 000 hab.

5.2.1. *Le contexte national*

La Grande-Bretagne est un pays quasiment sans péage routier. Cette situation est en partie due au fait que ce n'est qu'à partir de 1999 que la législation a autorisé les pouvoirs locaux à introduire un système de péage. En 1999, le gouvernement britannique a approuvé la loi « Greater London Authority Act » qui permet au maire et aux autorités de Londres d'imposer un système de péage routier. Les revenus du péage devaient être destinés à l'amélioration du transport local pendant au moins les dix ans qui suivent (Feasibility Study Report, 2004).

Un an plus tard, en 2000, le « Transport Act » a donné la même possibilité aux autorités locales en dehors de Londres mais uniquement en Angleterre et au pays de Galles. Toutes les mesures doivent être acceptées par le Secrétaire d'État. L'Écosse a adopté⁶ la loi « Transport (Scotland) Act » en 2001 avec certaines nuances. L'Irlande du Nord ne dispose pas de législation similaire (Feasibility Study Report, 2004).

5.2.2. *Le contexte local*

Le contexte local à Londres se caractérisait par une situation très mauvaise en matière de transports après des décennies de sous-investissement. Le débat est ancien autour du « *road congestion pricing* », il date de 1964 avec le rapport Smeed. Il y avait un consensus général chez les décideurs et dans l'opinion sur la nécessité de « faire quelque chose » pour réduire le trafic automobile. Du point de vue institutionnel, c'est l'établissement public *Transport for London* (TfL) qui est l'autorité organisatrice du système de transport londonien. TfL a pour mission de réaliser la stratégie de transport du maire et de gouverner les services de transport dont le maire est responsable.

Le péage de Londres représente une partie de la stratégie du maire Ken Livingston concernant le transport urbain et était un point important de sa politique et même de sa première (2000) puis seconde campagne électorale (2004).

Le péage a démarré en février 2003. C'est un péage de zone qui fonctionne de 7 h à 18 h 30 du lundi au vendredi, avec un tarif forfaitaire de 5 £ (7 € environ) par véhicule et par jour⁷, pour tout véhicule circulant dans la zone (détection par caméras lisant les plaques). Les

6. La population de l'agglomération d'Édimbourg vient de refuser par référendum, en 2005, le projet de péage de double cordon qui lui était proposé.

7. une augmentation de 5 £ à 8 £ du péage a été décidée par le maire de Londres le ... 1^{er} avril 2005.

recettes nettes du péage sont affectées prioritairement aux transports collectifs.

Les effets positifs attendus se sont manifestés immédiatement (cf. TfL, 2003). En ce qui concerne le trafic, les résultats ont dépassé les prévisions puisque le trafic dans la zone à péage a baissé de 20 % (en avril 2003) contre 10 à 15 % attendus. La congestion (mesurée comme le temps supplémentaire de trajet par rapport à la situation de fluidité la nuit) a diminué de 30 %. La fréquentation des bus, qui enregistrent une plus grande ponctualité et une baisse des coûts d'exploitation, est en hausse de 14 %. Il résulte de cette baisse inattendue du trafic dans la zone à péage que les recettes (100 M£ en année pleine) sont très inférieures aux prévisions (130 M£).

Le nombre de voitures entrantes dans la zone payante a été réduit d'environ 60 000 véhicules par jour sur approximativement 380 000 en 2002. On estime que 20 à 30 % de ces voitures circulent à l'extérieur de la zone et que 50 à 60 % de ces voyageurs utilisent désormais le transport en commun, lequel a gagné en efficacité (Abadie, 2004).

5.2.3. *La problématique de l'acceptabilité*

Ce projet a au départ rencontré une forte opposition, notamment du parti conservateur, des automobilistes, des syndicats et des habitants des quartiers concernés. L'un des facteurs déterminant de l'acceptabilité a été la réduction de 90 % accordée aux résidents de la zone à péage et l'exemption à 100 % pour les handicapés. Les résidents enregistrés sont plus de 22 000, les handicapés plus de 100 000. Il faut ajouter également, facteur supplémentaire d'acceptabilité, l'affectation des recettes du péage principalement à l'amélioration des transports en commun.

Quant à l'acceptabilité, en 2003 plus de 50 % des Londoniens (du Grand Londres, donc bien au-delà de la zone centrale) soutenaient cette démarche et seulement 30 % des habitants se prononçaient contre. 73 % des habitants reconnaissent l'efficacité dans la réduction de la congestion (TfL, 2003 ; Abadie, 2004).

Cette expérience montre comment une volonté politique forte conjuguée à un consensus sur le caractère critique de la situation a permis de passer outre l'impopularité de cette mesure qui paraissait initialement insurmontable. Le maire de Londres a été élu sur ce projet en 2000 mais également réélu en 2004, c'est-à-dire un an après la mise en fonctionnement du péage, avec son projet annoncé d'extension de la zone à péage.

5.3. *Le boulevard périphérique nord de Lyon*

Le boulevard périphérique Nord qui fut ouvert en 1997, est un bon exemple de non-acceptabilité publique. Il s'agit d'une nouvelle infrastructure à péage qui, dès son ouverture, a fait l'objet de fortes résistances sociales et politiques par la population concernée. Après deux épisodes judiciaires dont l'annulation de la concession, aujourd'hui seul le tunnel central est à péage et les tarifs en ont été réduits (Raux et Souche, 2001).

5.3.1. *Le contexte national*

La France est un pays avec une longue tradition de péage interurbain qui date de 1955 où une loi autorisant la concession d'autoroutes à péage a été approuvée, en dérogation au principe de gratuité d'usage des routes, issu de la Révolution de 1789. Ce n'est qu'en 1986 qu'a été instaurée la possibilité de faire appel au même mode de financement dans les agglomérations urbaines. Cependant, en milieu urbain, cette pratique reste réservée aux seuls ouvrages d'art (tunnels, ponts) nouvellement construits. Néanmoins, aussi bien en urbain qu'en interurbain, l'obligation de maintien d'un itinéraire alternatif gratuit demeure.

Du point de vue institutionnel, ce sont les autorités locales qui montent les projets de péage urbain dans lesquels le gouvernement central n'intervient que très peu. Néanmoins, le projet doit être autorisé par le Conseil d'État, la plus haute juridiction administrative nationale (Raux et Souche, 2001, 2004).

5.3.2. *Le contexte local*

Le boulevard périphérique nord (dénommé à l'époque TEO – Trans Est-Ouest) fut ouvert en 1997 : il relie un boulevard urbain périphérique, à l'est, à l'entrée de l'autoroute A6 au nord de l'agglomération, en passant par une zone très urbanisée. Il s'en est ensuivi des coûts de construction assez élevés (environ 900 M€) car il fallait enterrer près des deux tiers de l'ouvrage. Sur une longueur totale de 10 km, nous trouvons donc trois tunnels dont le plus long mesure 3,5 km.

Au vu des possibilités limitées des fonds publics, les autorités locales ont fait le choix de la concession avec péage, dont les recettes étaient supposées financer quasiment la moitié des coûts. Néanmoins, au fil des études précédant l'ouverture de l'ouvrage, il est apparu que le trafic serait insuffisant pour équilibrer financièrement l'opération. Le contrat de concession fut donc modifié par avenant pour inclure une réduction de capacité de certaines voiries existantes, parallèles et gratuites.

5.3.3. *La problématique de l'acceptabilité*

L'ouverture de l'infrastructure en août 1997 provoqua dès le départ un important mouvement de refus de la part des automobilistes. En effet ceux-ci découvrirent en même temps la nouvelle infrastructure à péage et les restrictions sur les voies parallèles. En outre, la signalisation et la configuration technique du périphérique Est étaient pensées pour diriger le trafic dans cette infrastructure à péage. Il s'en est ensuivi un mouvement de boycott de la nouvelle infrastructure, accompagné de manifestations chaque semaine aux barrières de péage, empêchant le paiement par les usagers, et parfois avec des destructions de ces barrières.

Parallèlement, deux actions en justice ont été menées par les opposants à ce péage :

- La première a débouché en septembre 1997 sur une annulation par le tribunal administratif de la clause de rétrécissement des voies parallèles gratuites. Le tribunal ordonna le rétablissement partiel et immédiat de la capacité d'écoulement du trafic sur le boulevard concerné.

- L'autre a concerné une annulation de la concession par le Conseil d'État en 1998 pour non-respect des obligations de publicité européenne et de mise en concurrence pour l'attribution de la concession.

Après quelques semaines de fermeture, suite à cette annulation de concession, l'ouvrage a été réouvert en exploitation par la Communauté Urbaine de Lyon, avec un péage réduit et restreint au tunnel central. L'infrastructure a été rachetée par la Communauté Urbaine de Lyon en 1999 qui en a concédé l'exploitation.

6. LE PROCESSUS MENANT À L'ACCEPTABILITÉ

Les expériences brièvement présentées ci-dessus, les études de cas réalisées par ailleurs ainsi que la revue de la littérature ont montré que l'acceptabilité n'était pas uniquement un résultat mais également un processus. Une nouvelle mesure est soit acceptée soit refusée, mais cet aboutissement n'est que la conséquence d'un long cheminement. En effet, l'acceptabilité est aussi le produit d'un jeu d'acteurs qui poursuivent des intérêts qui leurs sont propres, et entre lesquels se font et se défont les alliances.

Ces intérêts sont fortement dépendants du contexte et du moment dans lesquels la nouvelle mesure est introduite. Ils vont être relayés par des leaders d'opinion dont le leadership est positivement corrélé à leur capacité d'expertise, leur expérience et leur créativité (Childers,

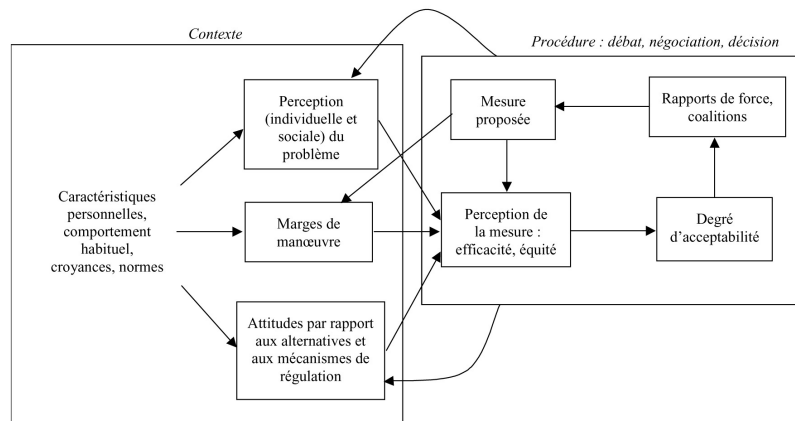
1986) mais également à la forme et aux moyens de pression dont ils disposent pour s'exprimer (Champagne, 1990).

Le processus d'acceptabilité met en œuvre un mécanisme de concertation et de prise de décision au cours duquel des coalitions se forment. Le débat sur l'acceptabilité, les mesures de compensation peuvent alors devenir des outils de persuasion pour parvenir à l'acceptabilité ou au rejet de la mesure.

Parvenus à ce stade, nous proposons une mise en ordre conceptuelle des différents aspects analysés plus haut (cf. Figure 1).

Figure 1

Une proposition de schéma conceptuel du processus d'acceptabilité



Dans ce schéma sont synthétisés les différents éléments de l'analyse précédente. Ce schéma distingue deux grands ensembles, l'un ayant trait au « contexte », l'autre à la « procédure » de débat et négociation menant au résultat final. Le « contexte » distingue trois éléments, qui sont la perception individuelle et sociale du problème, les marges de manœuvre individuelles, et les attitudes face aux alternatives et mécanismes de régulation. Ces trois éléments, d'une part sont conditionnés par les caractéristiques socio-économiques personnelles – comportements habituels, croyances et normes –, d'autre part conditionnent la perception de la mesure, notamment dans ses dimensions d'efficacité et d'équité. La « procédure » menant au résultat final (acceptation ou non du projet) fait interagir dynamiquement la modification éventuelle de la mesure proposée, dont les caractéristiques influent sur la perception de cette dernière, puis sur son degré d'acceptabilité. En fonction de cette perception, des coalitions se forment, des rapports de force s'établissent qui, en retour, vont ou non modifier la mesure.

Enfin, dernière rétroaction, la procédure de débat et négociation vient modifier la perception individuelle et sociale du problème, et les attitudes face aux alternatives et mécanismes de régulation.

7. CONCLUSION

Dans l'état de l'art que nous avons dressé, nous avons tenté de définir l'acceptabilité. Nous avons passé en revue l'identification des acteurs, nous avons souligné l'importance du contexte et de la perception du problème pour établir des conditions favorables à la réussite du projet. Nous avons montré comment, dans la perception des mesures par les acteurs, les arguments d'équité et de justice prennent souvent le pas sur ceux de l'efficacité au sens économique.

Ce traitement de l'acceptabilité a été illustré par trois expériences emblématiques de péage urbain, à Oslo, Londres et Lyon. Ces expériences montrent que la réussite de ce type de projet passe par les conditions nécessaires suivantes : (a) une volonté politique forte (cas de Londres) ou une convergence politique sur le projet (cas de la Norvège) ; (b) une compensation clairement affichée envers les « perdants » apparents (cas de Londres) ou un compromis gagnant-gagnant (cas de la Norvège) ; (c) la nécessité de porter au préalable le débat sur la place publique (contre exemple de Lyon).

Nous avons enfin conclu par une proposition de mise en ordre conceptuelle du processus menant à l'acceptabilité. Ce processus relève d'une alchimie complexe, et si le calcul économique peut prendre en compte des éléments de l'acceptabilité à travers l'intégration de l'équité dans son analyse (cf. chapitre 10), l'insertion de ce calcul dans le processus d'acceptabilité reste à éclairer.

Références

- Abadie G., « Le péage urbain de Londres, un an après », *Notes de synthèse du SES*, n° 153, mai-juin 2004.
- Baron J. et Journey J., « Norms against voting for coerced reform », *Journal of Economic Personality and Social Psychology*, 64, 1993, p. 347-355.
- Commission Européenne, « *Des redevances équitables pour l'utilisation des infrastructures : une approche par étapes pour l'établissement d'un cadre commun en matière de tarification des infrastructures de transport dans l'Union Européenne* », Livre Blanc, Bruxelles, Juillet 1998, 58 p.

- CERTU, DREIF, IAURIF, LET, « Les Péages urbains en Norvège : Oslo et Trondheim », Rapport des visites des 27 et 28 mai, 2002.
- Champagne P., *Faire l'opinion. Le nouveau jeu politique*, Collection Le sens commun, 1990, 312 p.
- Childers T.L., « Assessment of the psychometric properties of an opinion leadership scale », *Journal of Marketing Research*, Vol XXIII, 184-188, may 1986.
- Department for Transport. Feasibility study of road pricing in the UK, <http://www.dft.gov.uk/stellent/groups/dft—roads/documents/page/dft—roads—029788.hcsp>,
- Else P.K., « No entry for congestion taxes ? », *Transportation Research*, 20A, (2), 1986, p. 99-107.
- Emmerink R.H.M., Nijkamp P. et Rietveld P., « Is congestion pricing a first-best strategy in transport policy ? A critical review of arguments », *Environment and Planning B*, 22, 1995, p. 581-602.
- Evans A.W., « Road congestion pricing : when is it a good policy ? », *Journal of Transport Economics and Policy*, 26, 1992, p.213-244.
- Fernandez R. et Rodrik D., « Resistance to reform : status quo bias in the presence of individual specific uncertainty », *American Economic Review*, vol 81, n° 5, 1991, p. 1146-1155.
- Fleurbaey M., *Théorie économique de la justice*, Economica, Paris, 1996, 250 p.
- Frey B.S. et Jegen R., « Motivation crowding theory », *Journal of Economic Surveys*, 15, n° 5, 2001, p. 589-611.
- Frey B.S. et Oberholzer-Gee F., « Fair Siting Procedures : An Empirical Analysis of Their Importance and Characteristics », *Journal of Policy Analysis and Management*, vol 15, n° 3, 1996, p. 353-376.
- Frey B.S. et Pommerehne W.W., « On the fairness of pricing. An empirical survey among the general population », *Journal of Economic Behaviour and Organization*, 20, 1993, p. 295-307.
- Giuliano G., « An assessment of the political acceptability of congestion pricing », *Transportation*, 19 (4), 1992, p. 335-358.
- Goodwin, P.B., « The Rule of Three : a possible solution to political problem of competing objectives for road pricing », *Traffic Engineering and Control*, 1989, p. 495-497.
- Harrington W., Krupnick A. et Albertini A., « Overcoming public aversion to congestion pricing », *Transportation Research Part A*, 35, 2001, p. 91-111.
- Harsman B., « Success and Failure : Experiences from Cities », in Schade J. et Schlag B., *Acceptability of transport Pricing Strategies*, Elsevier, 2003.

- Hicks J.R., *A revision of demand theory*, Oxford, Clarendon Press, 1956.
- Jakobsson C., Fujii S. et Garling T., « Determinants of private car user's acceptance of road pricing », *Transport Policy*, 7, 2000, p.153-158.
- Kahneman D., Knetsch J.L. et Thaler R., « Fairness as a constraint on profit seeking : entitlements in the market », *American Economic Review*, vol. 76, n° 4, 1986, p. 728-741.
- Kunreuther H. et Easterling D., « The role of compensation in siting hazardous facilities », *Journal of Policy Analysis and Management*, 15, 1996, p. 601-622.
- Langmyrh T., « Managing equity. The case of road pricing », *Transport Policy*, vol. 4, n° 1, 1997, p. 25-39.
- Oberholzer-Gee F. et Weck-Hannemann H., « Pricing road use : politico-economic and fairness considerations », *Transportation Research Part D*, 7, 2002, p. 357-371.
- Odeck J. et Brathen S., « On public attitudes toward implementation of toll roads – the case of Oslo toll ring », *Transport Policy*, Vol. 4, No. 2, 1997, p. 73-83.
- Raux C. et Souche S., « L'acceptabilité des changements tarifaires dans le secteur des transports : comment concilier efficacité et équité ? », *Revue d'Économie Régionale et Urbaine*, n. 4, 2001, p. 539-558.
- Raux C., Souche S. et Andan O., *Les conditions d'acceptabilité de nouvelles mesures tarifaires en transports urbains*, Rapport pour l'ADEME, 2003, 255 p.
- Raux C. et Souche S., « The acceptability of urban road pricing : A theoretical analysis applied to experience in Lyon », *Journal of Transport Economics and Policy*, Vol. 38, Part 2, 2004, p. 191-216.
- Rawls J., *A Theory of Justice*, Harvard University Press, 1971, 607 p.
- Rietveld P. et Verhoef E.T., « Social feasibility of policies to reduce externalities in transport », in Button K.J et Verhoef E.T., *Road pricing, traffic congestion and the environment*, Edgar Elgar, 1998, 316 p.
- Schade J., « European Research Results on Transport Pricing Acceptability », p. 109-124, in Schade J. et Schlag B., *Acceptability of transport Pricing Strategies*, Elsevier, 2003.
- Schade J. et Schlag B., *Acceptability of Urban transport Pricing*, Helsinki 2000, Deliverable 2, AFFORD, 2000.
- Schade J. et Schlag B., « Acceptability of urban transport pricing strategies », *Transportation Research Part F* 6, 2003, p. 45-61.

- Schlag B. et Teubel U., « Public acceptability of transport pricing », *IATSS Research*, 21, N° 2, 1997, p. 134-142.
- Small K.A., « Using the revenues from congestion pricing », *Transportation*, 19, 1992, p. 359-381.
- Small K.A. et Rosen H.S., « Applied Welfare Economics with Discrete Choice Models », *Econometrica*, vol. 49, n° 1, 1981, p. 105-130.
- Souche S., « Péage urbain et équité : une revue de la littérature », *Les Cahiers Scientifiques du Transport*, janvier, 2003, 20 p.
- Transport For London, « Congestion Charging Central London. Impacts Monitoring », First Annual Report, June, 2003, 120 p.
(see <http://www.tfl.gov.uk/tfl/cclondon/cc—monitoring-1st-report.shtml>)
- Transport For London, « Congestion Charging Central London. Impacts Monitoring », Second Annual Report, April, 2004, 120 p.
(see <http://www.tfl.gov.uk/tfl/cclondon/cc—monitoring-2nd-report.shtml>)
- Transport For London, « Central London Congestion Charging Scheme. Impacts Monitoring », Third Annual Report, April, 2005. 162 p.
(see <http://www.tfl.gov.uk/tfl/cclondon/pdfs/ThirdAnnualReportFinal.pdf>)
- Thorpe N., « Public acceptance of road-user charging. A case-study of the toll-rings in Norway », *IATSS Research*, vol 26, N° 1, 2002, p. 17-27.
- Tretvik T., « Urban Road Pricing in Norway : Public Acceptability and Travel Behaviour », p. 77-92 in Schade J. et Schlag B., 2003, *Acceptability of transport Pricing Strategies*, Elsevier. 2003.
- Weck-Hannemann H., « Acceptance of Pricing Instruments in the Transport Sector », Annex 3, *PATS*, 1999, p. 65-92.
- Zajac E.E., *Political economy of fairness*, The MIT Press, Cambridge, Mass, 1995, 325 p.

Programmes européens

AFFORD

<http://europa.eu.int/comm/transport/extra/web/downloadfunction.cfm?docname=200310/afford.pdf&apptype=application/pdf>, janvier – février 2005.

CUPID

<http://transport-pricing.net/what22.html>, janvier – février 2005.

INPHORMM

<http://www.wmin.ac.uk/transport/inphormm/inphormm.htm>

PRIMA

<http://labs2.itcube.be/extra/final—reports/road/PRIMA.pdf>, janvier – février 2005.

MC-ICAM

<http://www.strafica.fi/mcicam/index.html>, janvier – février 2005.

PATS

<http://www.tis.pt/proj/pats/pats.html>, janvier – février 2005.

TAPESTRY, *State-of-the-Art Review*, Deliverable 2. Project funded by the European Commission, Transport RTD Programme of the 5th Framework Programme, 2001, 90 p.