



HAL
open science

Modélisation micro-économétrique des choix de production des agriculteurs

Elodie Letort

► **To cite this version:**

Elodie Letort. Modélisation micro-économétrique des choix de production des agriculteurs. Sciences de l'Homme et Société. Université de Rennes 1, 2009. Français. NNT : . tel-02824542

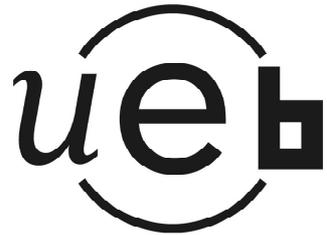
HAL Id: tel-02824542

<https://hal.inrae.fr/tel-02824542>

Submitted on 6 Jun 2020

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.



THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Européenne de Bretagne

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

Mention : Sciences Economiques

Ecole doctorale de Sciences Economiques et de Gestion

présentée par

Elodie Letort

préparée à l'unité mixte de recherche 1302 SMART
Structures et Marchés Agricoles, Ressources et Territoires

**Modélisation micro-
économétrique des
choix de production
des agriculteurs**

**Thèse soutenue à Rennes
le 2 juillet 2009.**

devant le jury composé de :

Isabelle CADORET

Professeur Université Rennes 1 / *président*

Philippe POLOME

Professeur Université Lyon 2 / *rapporteur*

Valentin PATILEA

Professeur INSA Rennes / *rapporteur*

Chantal LE MOUËL

Directeur de recherche INRA-UMR SMART Rennes
/ *directeur de thèse*

Jean-Marc MEYNARD

Directeur de recherche INRA-INA-PG Grignon / *co-
directeur de thèse*

Je tiens en tout premier lieu à remercier Alain Carpentier pour son implication dans ce travail de thèse. Tous les travaux de recherche présentés dans cette thèse n'auraient pu être menés à bien sans son expérience et ses compétences. Je le remercie pour tout le temps qu'il m'a consacré et pour toutes les connaissances techniques qu'il a su m'apporter.

Je tiens également à remercier Chantal Le Mouël et Jean-Marc Meynard pour m'avoir guidée dans la détermination de ce projet de thèse, ainsi que Bernard Rolland, Arnaud Gauffreteau et Chantal Loyce, qui ont montré un intérêt particulier à mes travaux de recherche et qui m'ont ouverte à de nouvelles perspectives.

Un merci tout particulier à Céline Nauges pour sa disponibilité, son soutien et ses nombreux conseils. Enfin, un grand merci à toutes les personnes de l'INRA de Rennes pour leur soutien amical et leur bonne humeur quotidienne.

SOMMAIRE

INTRODUCTION	7
Références.....	19
ARTICLE 1 : MODÉLISER LES CHOIX DE PRODUCTION AGRICOLE : FONCTIONS DE PRODUCTION ET GESTION DES ASSOLEMENTS	23
Introduction	24
1. Une brève revue de littérature.....	28
<i>Les modèles duaux (Weaver, 1983 ; Chambers, 1989).....</i>	<i>29</i>
<i>Le modèle de Just, Zilberman et Hochman (1983).....</i>	<i>32</i>
<i>Les modèles de Chambers et Just (1989).....</i>	<i>35</i>
<i>Le modèle de Moore et Negri (1992).....</i>	<i>36</i>
<i>Les jointures technologiques.....</i>	<i>37</i>
2. Les fonctions de production.....	42
<i>Modéliser les choix des agriculteurs : combiner les apports de l'économie et de l'agronomie.....</i>	<i>44</i>
<i>Les fonctions de production.....</i>	<i>46</i>
<i>Fonctions de demande d'intrants, d'offre de production et fonctions de profit indirectes.....</i>	<i>53</i>
3. Les déterminants des choix d'assolement	58
<i>Choix des assolements et motifs de diversification.....</i>	<i>58</i>
<i>Rendements marginaux de la terre décroissants et coûts implicites des assolements.....</i>	<i>63</i>
4. Modèles micro-économétriques : formes fonctionnelles simples et mise en œuvre	74
<i>La fonction de rendement « quadratique translatée ».....</i>	<i>75</i>
<i>Modèles de choix d'assolement MNL et fonction de coût « entropique ».....</i>	<i>79</i>
<i>Modèles de choix d'assolement fondés sur des fonctions de coût quadratiques.....</i>	<i>83</i>
<i>Quelques aspects liés à l'utilisation empirique des modèles proposés.....</i>	<i>88</i>
Conclusion.....	92
Références	95
Annexe	101

ARTICLE 2 : MODELING ACREAGE DECISIONS WITHIN THE MULTINOMIAL LOGIT FRAMEWORK

..... 105

Introduction 106

1. Main assumptions on the multi-crop production technology 109

2. Acreage decisions within MNL framework: the cost function approach 112

3. Acreage decisions within the MNL framework: the discrete choice approach 117

4. Empirical illustration..... 120

Data 120

Multi-crop econometric models..... 120

Estimation issues..... 124

Main results..... 125

5. Generalized MNL acreage share models..... 127

"Cost function approach": the Nested multinomial logit model..... 128

"Discrete choice approach": MNL acreage share models accounting for crop rotation effects... 130

Concluding remarks 132

References 135

Tables 138

Appendix 140

ARTICLE 3 : LAND AND VARIABLE INPUT ALLOCATION : WHY HETEROGENEITY MATTERS?..... 145

Introduction 146

1. Literature review 147

Approaches based on single input allocation equations..... 148

Approaches based on multicrop econometric models..... 150

Outline of the control function approach..... 152

2. Econometric model specification..... 154

Yield functions..... 155

Acreage functions..... 157

<i>“Complete” multi-crop econometric model</i>	159
3. Control function approach	160
<i>Distributional assumptions</i>	160
<i>Control function approach under normality assumptions</i>	162
<i>A three-stage inference procedure</i>	163
4. Empirical application	164
<i>Data</i>	165
<i>Main results</i>	166
Conclusion	168
References	170
Tables	175
Appendix	178

ARTICLE 4 : UN MODÈLE ÉCONOMÉTRIQUE FLEXIBLE POUR L’ALLOCATION DE LA TERRE ET DES INTRANTS VARIABLES	179
Introduction	180
1. Le modèle économétrique multi-produits	182
<i>Les fonctions de rendement</i>	182
<i>Les fonctions de parts de surface</i>	186
<i>Le modèle économétrique multi-produits « complet »</i>	188
2. L’approche par les fonctions de contrôle	189
<i>Hypothèses de distribution</i>	190
<i>Fonctions de contrôle basées sur la projection linéaire</i>	191
<i>Procédure d’estimation en trois étapes</i>	192
3. Application empirique	194
<i>Données</i>	194
<i>Les principaux résultats</i>	195
<i>Rendements marginaux de la terre décroissants versus coûts implicites des assolements</i>	198
Conclusion	200

References	202
Tableaux	204
Annexe	210
CONCLUSION	211
Références	217

Introduction

Après plus de 50 ans de modernisation de l'agriculture française, le rôle des agriculteurs se retrouve au cœur du débat social et économique. L'après-guerre les avait incités à améliorer leur performance afin de garantir la sécurité alimentaire et de relancer l'économie française. Aujourd'hui les enjeux sont différents.

Les réformes de la Politique Agricole Commune (PAC) de 1992 et de 1999 se sont traduites par une forte baisse du soutien des prix, et par suite une baisse des prix à la production, compensée par des aides directes à l'hectare. La réforme de la PAC de juin 2003 fait un pas supplémentaire dans la direction du découplage des aides de soutien aux revenus agricoles et des quantités produites. Même si les tractations politiques de dernière minute ont réussi à maintenir un certain degré de couplage au produit et un lien fort à la terre, il est plus que vraisemblable que les politiques de soutien des revenus agricoles seront, à l'avenir, de plus en plus déconnectées des volumes produits et des facteurs de production. Parallèlement, pour répondre aux attentes croissantes de la société en matière de protection de l'environnement et des ressources naturelles, de sécurité et de qualité des produits alimentaires, se développent des régulations visant à contraindre les agriculteurs à respecter les normes minimales dans ces domaines (conditionnalité des aides directes de soutien des revenus) et à les compenser voire

les rémunérer pour les efforts réalisés au-delà de ces normes minimales (aides directes dites du second pilier). En 2008, le Grenelle de l'environnement a décidé de réduire l'utilisation des pesticides de moitié d'ici 2018, « *en accélérant la diffusion des méthodes alternatives et sous réserve de leur mise au point* ».

Ces nouvelles conditions conduisent à s'interroger sur de nouveaux systèmes de cultures, qui combindraient un objectif de rendement et un objectif de réduction des intrants polluants. Les systèmes existants sont basés sur une utilisation intensive d'intrants chimiques, dans une logique de maximisation du rendement.

« Ils ont été mis au point dans les années 1980 parce qu'ils étaient rentables ; ils perdurent bien que leur rentabilité ait diminué, parce que c'est sur ces systèmes [intensifs] que se sont construits progressivement un savoir-faire collectif et un dispositif de conseil technique financé par la vente de pesticides » (Meynard et Savini, 2003).

Ce processus d'intensification des pratiques culturales s'auto-entretient en raison de toutes les complémentarités techniques entre pesticides, engrais, types de semences, date et densité de semis. Ces systèmes de production agricole intensifs nécessitent l'utilisation de beaucoup d'intrants, amènent à pratiquer des semis précoces et denses et à choisir les variétés, avant tout sur un critère de productivité. Les semis précoces et denses sont favorables au parasitisme et ce d'autant plus que les variétés les plus productives sont peu résistantes aux maladies. Ces risques élevés entraînent une fréquence élevée de traitements phytosanitaires. Cette tendance est renforcée par la surconsommation d'intrants, en particulier d'engrais azoté qui permet, par exemple, d'augmenter la teneur en protéines du grain de blé. La grande cohérence technico-économique des systèmes intensifs ne permet pas de changements ponctuels d'un élément de l'itinéraire technique. Pour diminuer l'utilisation des intrants chimiques, il faut réviser

globalement les systèmes existants et accompagner les agriculteurs pour qu'ils adoptent de nouveaux systèmes de cultures car ils impliquent des changements de pratiques importants.

Depuis le début des années 1980, des recherches plus directement orientées vers la mise au point de systèmes de culture moins intensifs ont été entrepris. Des expérimentations comparant deux itinéraires techniques, dont les objectifs de rendement différaient de 15 quintaux par hectare, ont permis de montrer, avec les prix de l'époque, que l'on pouvait réduire le rendement sans diminuer les marges brutes. Au début des années 1990, plusieurs programmes de recherche ont été mis en place pour étudier si la baisse des prix provoquée par la réforme de la PAC ne pouvait pas être compensée par une réduction des intrants dans le cadre de systèmes agricoles intégrés. Tous les essais réalisés dans plusieurs départements en France sont convergents et montrent que l'intérêt des itinéraires techniques à intrants réduits s'accroît quand baisse le prix du blé. La variabilité des rendements et des marges n'est pas accrue par la réduction des intrants, car celle-ci s'accompagne d'une réduction maîtrisée de verse et maladies.

La crédibilité des itinéraires techniques à bas niveaux d'intrants a été renforcée ces dernières années par la mise en place sur le marché de variétés rustiques multi-résistantes aux maladies. Ces variétés ont un potentiel de rendement qui reste un peu inférieur de l'ordre de 5 à 10 quintaux par hectare à celui des variétés de même génération plus productives mais sensibles. Un réseau de test d'itinéraires techniques adaptés aux variétés rustiques de blé tendre, regroupant l'Institut National de Recherche Agronomique (INRA), l'Institut Technique des Céréales et des Fourrages (ITCF) et les sélectionneurs privés du groupement d'intérêt économique « Club des 5 » a été lancé en 1999 à l'initiative de l'INRA pour tester les performances agronomiques et économiques de différents couples « variété-itinéraires techniques ». Ces analyses montrent qu'il est possible à la fois, de faire face à la baisse du prix des céréales décidée dans le cadre des réformes de la PAC de 1992 et 1999, et de répondre aux

exigences environnementales, par une double adaptation, le choix des variétés rustiques multi-résistantes, d'une part, et l'utilisation d'itinéraires techniques plus économes en intrants, d'autre part (Rolland et al., 2003).

L'intérêt économique des couples variétés rustiques-conduite à coûts réduits a été fort ces dernières années dans un contexte de prix bas, mais est fortement tempéré aujourd'hui par la hausse des cours. Dans une perspective de changement de pratiques agricoles, le réseau expérimental se pose aujourd'hui la question de l'acceptabilité d'une innovation (Bouchard et al., 2008).

Ce projet de thèse est né d'interrogations communes entre économistes et agronomes sur le développement de ces nouveaux systèmes de cultures, et particulièrement sur l'intérêt économique des couples variétés rustiques de blé tendre-itinéraires techniques à bas niveaux d'intrants. Des réflexions ont été menées sur les outils disponibles pour évaluer les impacts de ces nouveaux systèmes sur les choix de production des agriculteurs. Celles-ci ont abouti à un double constat : les modèles économiques standards ne sont pas adaptés pour répondre à ces enjeux, et ils sont difficilement utilisables dans le cadre de recherches pluridisciplinaires, et ce pour principalement deux raisons : 1) les fonctions de production utilisées par les économistes et surtout les approches qu'ils utilisent couramment pour caractériser les technologies agricoles (approches duales via des fonctions de coût ou de profit) ne prennent en compte que de façon implicite les contraintes agronomiques et sont peu lisibles et interprétables par les agronomes ; 2) les données issues de modèles agronomiques ou d'expérimentations produites par les agronomes et à l'échelle des parcelles sont difficilement utilisables par les économistes via des fonctions de production, de coût ou de profit usuelles, qui sont adaptées pour traiter des données à l'échelle d'une exploitation, d'une région ou d'un état.

L'objectif de cette thèse est donc de proposer une modélisation des choix des agriculteurs qui soit à la fois cohérente avec la théorie économique et utilisable dans le cadre de collaborations avec des agronomes. C'est cet objectif de rapprochement avec l'agronomie qui a dicté nos choix en termes de modélisation du comportement de l'agriculteur. Cette thèse apporte trois contributions principales dans la modélisation micro-économétrique des choix de production des agriculteurs.

L'utilisation de l'approche primale constitue une première originalité de ce travail de thèse. L'approche primale, ou approche par les fonctions de production, permet d'intégrer de façon simple des contraintes de type agronomique et par conséquent de produire des résultats lisibles et facilement interprétables par des non-économistes. Dans la littérature récente (depuis les années 90), les technologies de production sont plutôt représentées à partir d'une approche duale (fonction de profit ou fonction de coût), plus simple d'utilisation que l'approche primale. L'approche duale constitue, par exemple, un outil pratique pour analyser les impacts des réformes de la PAC sur l'allocation des terres (Moore et Negri, 1992 ; Oude Lansink et Peerlings, 1996 ; Guyomard et al., 1996, Moro et Sckokai, 1999). L'intérêt de cette approche est qu'il n'est pas nécessaire comme dans l'approche primale, de définir avec précision les relations entre les niveaux de production et les niveaux d'utilisation d'intrants. En particulier, elle permet de ne considérer que des données agrégées au niveau des exploitations et de négliger les allocations d'intrants (Paris, 1989), ce qui la rend applicable avec les données habituellement disponibles (Shumway, 1995).

Une deuxième originalité de ce travail de thèse est de combiner l'approche micro-économétrique et la programmation mathématique (positive), et en particulier la représentation implicite des contraintes agronomiques, de capital et de travail, à travers une fonction de coût implicite (Howitt, 1995). La structure simple des modèles de programmation mathématique permet de prendre en compte des contraintes technologiques, biophysiques,

politiques ou environnementales, « d'accueillir » facilement des paramètres estimés par ailleurs, notamment des résultats d'essais agronomiques, et de fournir des résultats facilement interprétables. Par contre, les paramètres de ces modèles sont soit intégrés directement lorsqu'ils sont disponibles, ou calibrés, ce qui les prive de tests d'inférence et de validations statistiques (De Frahan et al., 2007). En plus, en raison de leurs discontinuités, ces modèles produisent des résultats de simulation avec d'importants « basculements » ou « sauts » dans les choix des agriculteurs (Heckeleï, 1997). Ces limites ont motivé le développement de nouvelles approches. D'une part, la programmation mathématique positive (Howitt, 1995) permet la construction de modèles fondés sur les principes de la programmation mathématique standard tout en « lissant » leurs discontinuités. D'autre part, l'inférence statistique par la méthode du maximum d'entropie s'avère particulièrement bien adaptée à l'estimation des paramètres des modèles de programmation mathématique positive (Paris et Howitt, 1998 ; Heckeleï et Wolff, 2003).

La troisième originalité de ce travail de thèse réside dans le traitement des termes aléatoires. Il n'y a pas de différence entre nos modèles « économiques » et nos modèles « économétriques ». Les modèles proposés contiennent des termes aléatoires qui font partie intégrante de la spécification choisie (Zellner et al., 1966 ; Mundlak, 1963). Dans la plupart des modèles de choix de production agricole multi-produits, les termes d'erreur sont simplement « ajoutés » et interprétés comme des erreurs d'approximation. Dans nos modèles, les termes aléatoires représentent soit l'hétérogénéité des conditions de production entre les exploitations, soit les aléas inhérents à la production agricole (aléas climatiques ou sanitaires). Si les aléas de productions sont explicitement considérés par les modèles dédiés à l'analyse de ces aléas ou l'attitude de l'agriculteur face au risque (Antle, 1983 ; Just et Pope, 2001 ; Feder, 1979), l'hétérogénéité des conditions de production est rarement considérée. Or, les résultats obtenus par l'économétrie des données de panel en ce qui concerne l'estimation de modèles micro-économiques de fonctions de production en général (Griliches et Mairesse, 1995 ;

Blundell et Bond, 2000) ou de fonctions de production agricole en particulier (Carpentier et Weaver, 1996), tendent à montrer que la majeure partie de la variation des niveaux de production est due aux variations inter-firmes ou inter-exploitations et que cette hétérogénéité ne peut en général être contrôlée par les variables disponibles.

Ce travail de thèse contribue donc à la littérature micro-économétrique sur la modélisation des choix des agriculteurs. Tous les modèles qui seront développés dans la suite sont construits à partir de développements théoriques, et sont estimables économétriquement sous des hypothèses que nous détaillerons. Ils permettent d'apporter des éclairages nouveaux sur des questions telles que la modélisation des choix d'assolement et l'allocation des intrants entre les cultures, et offrent des perspectives intéressantes en termes de modélisation dynamique des choix d'assolement et de traitement économétrique des solutions en coin. Cette thèse se décline sous la forme de quatre articles, chacun représentant un chapitre.

Dans le chapitre 1, nous présentons de nouveaux modèles de choix de production agricole dans un contexte de productions multiples. Ce chapitre est particulièrement dense et détaillé par rapport aux autres chapitres, et ce pour deux raisons. D'une part, il introduit de nouveaux modèles, ce qui nécessite de les comparer à la littérature existante et d'expliquer les motivations qui ont dicté nos choix de modélisation. D'autre part, la structure générale de ces modèles et les hypothèses posées sont communes à tous les modèles présentés dans ce travail de thèse. Les descriptions de ces modèles seront donc moins détaillées dans les autres chapitres.

Ces modèles ont été conçus pour être employés dans le cadre de recherches impliquant des économistes, des agronomes et des conseillers agricoles, et pour être facilement généralisable (afin, entre autres, de prendre en compte les effets des rotations culturales). Ils ont donc une structure simple et leurs paramètres admettent des interprétations directes, en termes agronomiques ou en matière de comportement des producteurs agricoles. Pour cela, ils

reposent sur une représentation primale de la technologie de production des cultures. Les modèles économétriques utilisés dans la littérature pour décrire les choix de production agricole s'appuient généralement sur une approche duale (Oude Lansink et Peerlings, 1996 ; Guyomard et al., 1996, Moro et Skokai, 1999). Les principaux modèles existants dans la littérature relative aux choix de production des agriculteurs sont présentés et comparés.

La première étape de ce travail est de présenter la structure générale de nos modèles et les hypothèses associées. Nos modèles combinent à des degrés divers, des éléments des deux grandes approches de la modélisation empirique des choix de production agricole : la programmation mathématique positive (Howitt, 1995 ; Heckelevi et Wolff, 2003) et la modélisation économétrique (Just, Zilberman et Hochman, 1983 ; Chambers et Just, 1989 ; Moore et Negri, 1992). En particulier, nous utilisons ce que nous nommons la « fonction de coût implicite de gestion de l'assolement », qui est à la base des techniques de la programmation mathématique positive.

La deuxième étape de ce travail est de donner des formes fonctionnelles aux fonctions de production et à la fonction de coût implicite. Ces formes fonctionnelles sont choisies pour leur flexibilité et/ou leur facilité d'interprétation. Des termes aléatoires sont intégrés explicitement dans les modèles pour représenter, entre autres l'hétérogénéité des conditions de production entre les exploitations. L'hétérogénéité est une des sources majeures de variabilité des niveaux de production agricole entre les exploitations et entre les périodes (Carpentier et Weaver, 1996, 1997), et est donc importante à prendre en compte dans les modèles destinés à être estimés à partir de données microéconomiques. Ces modèles microéconométriques construits à partir d'une approche primale ont plusieurs possibilités d'application et d'extension.

Dans le chapitre 2, nous décrivons plus précisément un des modèles cités dans le chapitre 1. La particularité de ce modèle est que la forme Logit multinomial est utilisée pour décrire les choix d'assolements. Le modèle Logit multinomial a été souvent utilisé pour décrire les choix

d'assolement (Caswell et Zilberman, 1985 ; Lichtenberg, 1989 ; Wu et Segerson, 1995) mais jamais il n'avait été démontré que cette forme particulière pouvait être dérivée d'un programme de maximisation du profit. La particularité de notre modèle est sa fonction de coût implicite utilisée pour représenter des contraintes non-observées par les économistes et auxquelles l'agriculteur doit faire face (disponibilité de la main d'œuvre et du matériel agricole, agronomiques) (Howitt, 1995). C'est la forme spécifique de cette fonction de coût, basée sur une fonction d'entropie, qui permet de dériver des équations de parts de surfaces de forme Logit multinomial.

Dans ce chapitre, nous dérivons également des équations de parts de surface de forme Logit multinomial à partir d'un modèle de choix discrets. Ce modèle est une extension des travaux de Caswell et Zilberman (1985) qui ont dérivé un modèle de choix d'assolement en agrégeant des décisions faites à l'échelle de la parcelle. Le problème de cette approche est que les décisions d'assolement sont supposées être indépendantes les unes des autres. Nous appliquons alors le concept d'ajustement partiel pour approximer les choix dynamiques des parts de surfaces (Treadway, 1971 ; Oude Lansink et Stefanou, 2001). Outre le fait que les équations de parts de surface de forme Logit multinomial ont une forme simple et des paramètres facilement interprétables, nous montrons que de nombreuses généralisations sont possibles. En particulier, ce type de modèle peut facilement être étendu pour tenir compte des effets des rotations culturales dans les choix d'assolement (Rust, 1987 ; Livingston et *al.*, 2008).

Les modèles économétriques obtenus sont estimés à partir de données françaises du Réseau d'Information Comptable Agricole (RICA) de 1988 à 2006 sur les exploitations spécialisées en céréales et grandes cultures. Bien que le modèle de choix discret fournisse de meilleurs résultats, les résultats obtenus montrent que les deux modèles économétriques représentent de manière satisfaisante les décisions de production des agriculteurs.

Dans le chapitre 3, toujours sur la base des modèles et des hypothèses décrits dans le chapitre

1 et 2, nous proposons une méthode d'allocation des intrants variables entre les cultures. L'allocation des intrants variables entre les cultures est un problème récurrent en économie agricole à cause du manque de données disponibles sur les dépenses d'intrants par culture. Connaître la répartition des facteurs de production par culture est nécessaire lorsque l'on souhaite analyser les marges brutes culture par culture et étudier les effets des politiques publiques sur l'utilisation d'intrants.

Les modèles économétriques basés sur une approche duale utilisent, en général, les données disponibles sur les quantités d'intrant totales appliquées par l'exploitant, sans tenter de répartir ces intrants entre les cultures. Des méthodes d'allocation ont cependant été développées pour différentes raisons : pour estimer simplement les dépenses d'inputs par activité ou par culture (Just et *al.*, 1983 ; Chambers et Just, 1989), pour calculer des coefficients inputs-outputs (Peeters et Surry, 1993; Dixon et al., 1984 ; Hornbaker et al., 1989), ou pour évaluer les effets de politiques agro-environnementales sur l'utilisation d'inputs (Lence et Miller, 1998).

Cet article a trois objectifs. Les approches utilisées jusqu'ici s'appuient principalement sur des méthodes de régression standard. Le premier objectif est de montrer que les allocations des intrants par les méthodes standards de régression sont vraisemblablement biaisées du fait de l'endogénéité des surfaces dans l'équation de répartition des intrants. En effet, puisque la quantité utilisée d'un intrant variable sur une culture détermine la profitabilité de cette culture qui est elle-même un déterminant des choix de surfaces, les surfaces sont vraisemblablement corrélées au terme d'erreur dans l'équation d'allocation des intrants. Cette corrélation est due à l'hétérogénéité non observée des exploitations qui influence en même temps les décisions d'utilisation d'intrants et les choix d'assolement. La qualité du sol, par exemple, peut influencer l'agriculteur dans ses choix de cultures, mais aussi dans ses applications d'engrais.

Le deuxième objectif de cet article est de construire un modèle économique qui permette de répartir les intrants entre les cultures tout en corrigeant l'endogénéité des surfaces. L'idée est

d'intégrer explicitement les termes d'erreur liés à l'hétérogénéité non observée dans le modèle économétrique, afin de « localiser » la corrélation entre les surfaces et le terme d'erreur dans l'équation de répartition des intrants.

Le troisième objectif de cet article est de proposer une approche basée sur les fonctions de contrôle pour contrôler ce problème d'endogénéité. La méthode des Variables Instrumentales ne peut pas être appliquée dans ce cas parce que les termes d'erreur contiennent la variable endogène et sont donc corrélés avec les instruments potentiels. L'approche par les fonctions de contrôle, généralement utilisée pour contrôler l'endogénéité de variables explicatives (Wooldridge, 2007) ou pour tenir compte des biais de sélection dans les échantillons (Heckman, 1974, 1979), est ici la seule méthode adaptée. Une procédure de mise en œuvre (en trois étapes) de cette méthode est proposée.

L'application empirique est réalisée sur les mêmes données françaises (de 1988 à 2006). Plusieurs estimations sont réalisées pour mettre en évidence l'importance de corriger le problème d'endogénéité des parts de surface et le rôle des variables contrôlant l'hétérogénéité des exploitations sur ce biais d'endogénéité. Les résultats montrent que ce biais modifie les allocations d'intrants entre les cultures, mais qu'il est atténué par l'utilisation de variables de contrôle.

Dans le chapitre 4, nous proposons de construire un modèle économique plus général et plus flexible que celui développé dans les chapitres 2 et 3. Jusqu'à présent une forme construite à partir d'une fonction d'entropie était utilisée pour spécifier la fonction de coût implicite du modèle et permettait de représenter les équations des parts de surface par des fonctions de type Logit multinomial (chapitres 2 et 3). Néanmoins, l'utilisation de cette forme spécifique pour la fonction de coût implicite s'appuyait sur certaines hypothèses qui peuvent s'avérer restrictives dans certains cas. Nous proposons donc de spécifier la fonction de coût implicite de gestion des assolements sous une forme quadratique. Cette forme a deux principaux avantages : elle est plus flexible et permet de dériver des équations de parts de surfaces avec

plus de paramètres ; elle permet de relâcher certaines hypothèses (détaillées ci-dessous).

Dans les chapitres précédents, la marge brute d'une culture était supposée indépendante de la surface allouée à cette culture, ce qui revient à faire l'hypothèse de rendement marginal constant de la terre. Cette hypothèse est courante dans les modèles de programmation mathématiques (Heckeleï et Wolff, 2003). En revanche, les modèles micro-économétriques couramment utilisés pour décrire les choix de production, et caractérisés par l'existence d'intrants fixes allouables (Moore et Negri, 1992), considèrent que le profit d'une culture décroît avec la quantité de terre allouée à cette culture (soit une hypothèse de rendement marginal décroissant de la terre). Le premier objectif du chapitre 4 est de relâcher l'hypothèse de rendement marginal constant en spécifiant la fonction de coût implicite sous une forme quadratique et en intégrant des effets d'échelles dans les fonctions de rendements (et donc dans les marges brutes par culture).

La répartition des intrants entre les cultures a nécessité de corriger un biais d'endogénéité des surfaces (chapitre 3). Nous utilisons l'approche par les fonctions de contrôle, qui revient à déterminer explicitement la forme de la corrélation entre les surfaces et les termes d'erreurs dans l'équation d'allocation des intrants, sous une hypothèse de normalité des termes d'erreur (Imbens et Wooldridge, 2008). Le deuxième objectif de ce chapitre est d'appliquer l'approche par les fonctions de contrôle pour corriger ce biais d'endogénéité, mais en utilisant les techniques de la projection linéaire (Chamberlain, 1982 ; Wooldridge, 2004). La forme quadratique de la fonction de coût permet d'éviter de poser une hypothèse sur la distribution des termes d'erreurs parce que les équations de parts des surfaces dérivées ont des termes d'erreurs additifs.

L'application empirique est réalisée à partir des mêmes données françaises. Plusieurs spécifications du modèle sont estimées : avec ou sans fonction de coût implicite de gestion des assolements et avec ou sans les effets d'échelle. Les résultats montrent que l'utilisation de la fonction de coût pour tenir compte implicitement des contraintes agronomiques, de travail et de capital, a un rôle important dans la détermination des paramètres et des élasticités-prix.

Références

- Antle, J.M. 1983. "Sequential Decision Making in Production Models." *American Journal of Agricultural Economics* 65:282-290.
- Blundell, R. et S. Bond. 2000. "GMM Estimation with Persistent Panel Data: An Application to Production Functions," *Econometric Reviews* 19, 321-340.
- Bouchard, C., Bernicot, M.H., Félix, I., Guérin, O., Loyce, C., Omon, B. et B. Rolland. 2008. "Associer des Itinéraires Techniques de Niveau d'Intrants Variés à des Variétés Rustiques de Blé Tendre : Evaluation Economique, Environnementale et Energétique." *Courrier de l'environnement de l'INRA* 55:53-77.
- Buysse, J., Fernagut, B., Harmignie, O., Henry de Frahan, B., Lauwers, L., Polomé, P., Van Huylenbroeck, G. et J. Van Meesel. 2004. "Modelling the Impact of Sugar Reform on Belgian Agriculture." Selected Paper presented at the International Conference on Policy Modelling, Paris, 30 June – 2 July, 2004.
- Carpentier, A. et R.D. Weaver. 1997. "Damage Control Productivity: Why specification matters." *American Journal of Agricultural Economics* 52:11-22.
- Carpentier, A. et R.D. Weaver. 1996. "Intertemporal and Interfirm Heterogeneity: Implications for Pesticide Productivity." *Canadian Journal of Agricultural Economics* 44:219-236.
- Caswell, M. et D. Zilberman. 1985. "The Choice of Irrigation Technologies in California." *American Journal of Agricultural Economics* 67:224-234.
- Chamberlain, G. 1982. "Multivariate Regression Models for Panel Data." *Journal of Econometrics* 18:5-46.
- Chambers, R. et R. Just. 1989. "Estimating Multioutput Technologies." *American Journal of Agricultural Economics* 71:980-95.
- Dixon, B. L., Batte, M. T. et S. T. Sonka. 1984. "Random Coefficients Estimation of Average Total Product Costs For Multiproduct Firms." *Journal of Business Economics and Statistics* 2: 360-366.

- Feder, G. 1979. "Pesticides, Information and Pest Management under Uncertainty." *American Journal of Agricultural Economics* 61:97-103.
- Griliches, Z. et J. Mairesse. 1995 "Production Functions: the Search for Identification" National Bureau of Economic Research, Cambridge, MA, NBER Working Paper No. 5067.
- Guyomard, H., Baudry, M. et A. Carpentier. 1996. "Estimating crop supply response in the presence of farm programmes: Application to the CAP." *European Review of Agricultural Economics* 23:401-420.
- Heckelei T. 1997. "Positive Mathematical Programming: Review of the Standard Approach." CAPRI Working paper 97-03. Bonn : University of Bonn.
- Heckelei, T., et W. Britz., 2001. "Concept and Explorative Application of an EU-wide, Regional Agricultural Sector Model.", in Heckelei T., Witzke, H.P., and Henrichsmeyer, W. (eds) *Agricultural Sector Modelling and Policy Information Systems*, Vauk Verlag, Kiel.
- Heckelei, T. et W. Britz. 2005. "Models Based On Positive Mathematical Programming: State of the Art and Further Extensions." Paper presented at the *89th EAAE Seminar on 'Modelling agricultural policies: state of the art and new challenges'*, Parma.
- Heckelei, T. et H. Wolff. 2003. "Estimation of constrained optimisation models for agricultural supply analysis based on generalised maximum entropy." *European Review of Agricultural Economics* 30:27-50.
- Heckman, J. J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47:53-61.
- Heckman, J. J. 1974. "Shadow Prices, Market Wages, and Labor Supply." *Econometrica* 42:679-693.
- Hornbaker, R. H., Dixon, B. L. et S. T. Sonka. 1989. "Estimating Production Activity Costs for Multioutput Firms with a Random Coefficient Regression Model." *American Journal of Agricultural Economics* 71:167-177.
- Howitt, E. 1995. "Positive Mathematical Programming." *American Journal of Agricultural Economics* 77:329-342.

- Imbens, G. et J. M. Wooldridge. 2007. "Control Functions and Related Methods" Lecture notes NBER Summer Institute '07, What's new in econometrics?, NBER.
- Judez, L., Chaya, C., Martinez, S., et A.A Gonsalez. 2001. "Effects of the Measures Envisaged in 'Agenda 2000' on Arable Crop Producers and Beef and Veal Producers: an Application of Positive Mathematical Programming to Representative Farms of a Spanish Region." *Agricultural Systems*, 67:121-138.
- Just, R.E. et R.D. Pope. 2001. "The Agricultural Producer: Theory and Statistical Measurement." In B.L. Gardner and G.C. Raussler, ed. *Handbook of Agricultural Economics*, Vol. 1A. Amsterdam, Elsevier.
- Just, R.E., Zilberman, D., et E. Hochman. 1983. "Estimation of Multicrop Production Functions." *American Journal of Agricultural Economics* 65:770-780.
- Lence, S. H. et D. J. Miller. 1998a. "Estimation of Multi-Output Production Functions with Incomplete Data: A generalised Maximum Entropy Approach." *European Review of Agricultural Economics* 25:188-209.
- Lence, S. H. et D. J. Miller. 1999. "Recovering Output-Specific Inputs from Aggregate Input Data: A Generalized Cross-Entropy Approach." *American Journal of Agricultural Economics* 80:852-867.
- Lichtenberg, E. 1989. "Land Quality, Irrigation Development and Cropping Pattern in the Northern High Plains." *American Journal of Agricultural Economics* 71:187-94.
- Livingston, M., Roberts, M.J. et J. Rust. 2008. "Optimal Corn and Soybean Rotations." Paper presented at AAEA annual meeting, Orlando, Florida, 27-29 July.
- Moore, M. R. et D. H. Negri. 1992. "A Multicrop Production Model of Irrigated Agriculture, Applied to Water Allocation Policy of the Bureau of Reclamation." *Journal of Agricultural and Resource Economics* 17:29-43.
- Moro, D. et P. Sckokai. 1999. "Modelling the CAP Arable Crop Regime in Italy: Degree of Decoupling and Impact of Agenda 2000." *Cahiers d'économie et sociologie rurales* 53:50-73.

- Paris, Q. et R.E. Howitt. 1998. "An Analysis of Ill-posed Production Problems using Maximum Entropy." *American Journal of Agricultural Economics* 80(1):124-138.
- Peeters, L. et Y. Surry. 1993. "Estimating Feed Utilisation Matrices Using a Cost Function Approach." *Agricultural Economics* 9:109-126.
- Oude Lansink, A. et S. Stefanou. 2001. "Dynamic Area Allocation and Economies of Scale and Scope." *Journal of Agricultural Economics* 52(3):38-52.
- Rolland, B., Bouchard, C., Loyce, C., Meynard, J.M., Guyomard, H., Lonnet, P. et G. Doussinault. 2003. "Des Itinéraires Techniques à Bas Niveaux d'Intrants pour des Variétés Rustiques de Blé Tendre : Une Alternative pour Concilier Economie et Environnement." *Courrier de l'environnement de l'INRA* 49:53-77.
- Rust, J. 1987. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica* 55:999-1033.
- Shumway, R. 1995. "Recent Duality Contributions in Production Economics." *Journal of Agricultural and Resource Economics* 20:178-194.
- Treadway, A.B. 1971. "On the Rational Multivariate Flexible Accelerator." *Econometrica* 39:845-855.
- Viaux, P. 1999. Une troisième voie en grande culture, Environnement, Qualité, Rentabilité. Editions Agridécisions Livre. 211p.
- Wooldridge, J. M., 2008. "Instrumental Variables Estimation of the Average Treatment Effect in Correlated Random Coefficient Models," *Advances in Econometrics* 21:93-116.
- Wooldridge, J. M. 2004. "Estimating Average Partial Effects under Conditional Moment Independence Assumptions." *The Institute for Fiscal Studies, Dpt of Economics, UCL. Cemmap*, WP CWP03/04.
- Wu, J. et K. Segerson. 1995. "The Impact of Policies and Land Characteristics on Potential Groundwater Pollution in Wisconsin." *American Journal of Agricultural Economics* 77:1033-47.

Chapitre I

Modéliser les Choix de Production Agricole :

Fonctions de Production et Gestion des Assolements¹

Résumé

L'objectif de cet article est de construire de nouveaux modèles de choix de production, conçus pour être employés dans le cadre de recherches impliquant des économistes, des agronomes et des conseillers agricoles. Nos modèles ont trois principales caractéristiques. Tout d'abord, ils sont construits à partir d'une représentation primale de la technologie de production agricole même s'ils s'appuient par la suite sur des concepts duaux. L'approche primale est moins flexible que l'approche duale mais elle est plus transparente pour des économistes et des non-économistes. Ensuite, ils combinent des éléments des deux grandes approches de la modélisation empirique, la programmation mathématique et la modélisation économétrique. La structure générale de ces modèles repose sur des fonctions de production et sur un concept emprunté à la programmation mathématique positive, la fonction de coût implicite de la gestion de l'assolement. Enfin, l'hétérogénéité des conditions de production et le caractère aléatoire de la production agricole sont considérés directement dans la spécification de nos modèles.

¹ Ce chapitre est basé sur un papier écrit conjointement avec Alain Carpentier.

1.1 Introduction

Deux méthodes principales sont utilisées pour construire des modèles de choix de production des agriculteurs : la programmation mathématique et la modélisation économétrique. La programmation mathématique a longtemps été prédominante pour représenter le fonctionnement des exploitations agricoles et simuler les impacts de chocs exogènes (des changements de politiques publiques, par exemple) sur les choix de production des agriculteurs, due à sa relative simplicité de mise en œuvre et d'interprétation. Ces modèles offrent des avantages uniques par rapport aux autres méthodes : ils sont bien adaptés pour caractériser des contraintes technologiques, biophysiques, politiques ou environnementales ; ils « accueillent » facilement des paramètres estimés par ailleurs, notamment des résultats d'essais agronomiques ; et leur structure est facilement interprétable par les décideurs politiques. En revanche, les paramètres de ces modèles sont difficiles à estimer à partir des méthodes d'inférence statistique classiques. Ces dernières sont délicates à utiliser pour des modèles caractérisés par d'importantes discontinuités et des contraintes à l'inégalité. De fait, ces paramètres sont soit intégrés directement lorsqu'ils sont disponibles, soit calibrés. Ceci prive les modèles de la programmation mathématique de tests d'inférence et de validations statistiques (De Frahan et *al.*, 2006). En outre, en raison de leurs discontinuités ces modèles produisent des résultats de simulation avec d'importants « basculements » ou « sauts » dans les choix des agriculteurs (Heckeleï, 1997).

Ces limites ont motivé le développement de nouvelles approches. D'une part, la programmation mathématique positive (Howitt, 1995) permet la construction de modèles fondés sur les principes de la programmation mathématique standard tout en « lissant » leurs discontinuités. D'autre part, l'inférence statistique par la méthode du maximum d'entropie, de par ses caractéristiques, s'avère particulièrement bien adaptée à l'estimation des paramètres des modèles de la programmation mathématique ou de la programmation mathématique positive (*e.g.*, Paris et Howitt, 1998 ; Heckeleï et Wolff 2003). Heckeleï et Wolff (2003)

proposent par exemple d'utiliser directement les conditions d'optimalité du modèle de programmation pour estimer, et non calibrer, les paramètres à partir des estimateurs du maximum d'entropie. Bien qu'aucune application de cette nouvelle approche n'ait encore été réalisée du fait du manque de données disponibles et de problèmes de résolution numérique (Buysse et *al.*, 2007), cette approche apparaît prometteuse. De fait, elle adapte les approches employées pour la modélisation économétrique, en particulier l'utilisation des conditions d'optimalité ou plus généralement de la théorie duale du comportement du producteur, dans le cadre théorique de la programmation mathématique. Les concepts et outils de la programmation mathématique positive procurent un « pont » théorique entre la modélisation économétrique et la programmation mathématique « standard ».

Contrairement aux modèles de programmation mathématique, les modèles micro-économétriques sont conçus à partir de représentations flexibles et « lisses » de la technologie de production agricole afin d'être estimés à partir des méthodes d'inférence statistique classiques. Ces modèles ont notamment été utilisés dans les années 1980 et 1990 pour tester des hypothèses sur la forme de la technologie agricole, et notamment des hypothèses de jointure (*e.g.*, Just, Zilberman et Hochman, 1983 ; Chambers et Just, 1989 ; Moore et Negri, 1992 ; Asunka et Shumway, 1996). Les modèles de la programmation mathématique ne peuvent être utilisés dans cette optique puisqu'ils postulent des formes de technologie rigides mais très faciles d'utilisation, notamment en matière de simulation.

Les modèles micro-économétriques sont généralement conçus pour définir des fonctions de réponse flexibles et « lisses » représentant les choix des agriculteurs. Dans ces modèles, les technologies de production sont souvent représentées à partir de formes implicites flexibles et utilisent une approche duale de la modélisation du comportement du producteur (Chambers 1988). L'utilisation de ces modèles s'est développée depuis les années 90, à la fois aux Etats-Unis et en Europe, notamment pour analyser les impacts des réformes de la PAC sur l'allocation des terres (Lansink et Peerlings, 1996 ; Guyomard et *al.*, 1996 ; Moro et Sckokai, 1999). L'intérêt de l'approche duale par rapport à l'approche primale réside dans sa simplicité

de calcul et d'utilisation. Les résultats de la théorie duale du producteur peuvent être utilisés avec des formes de technologie très générales, en particulier parce qu'il est possible de construire des modèles empiriques cohérents du point de vue de la théorie économique directement à partir de la spécification de fonctions de profit indirectes. Il n'est pas nécessaire, comme dans l'approche primale, de définir avec précision les relations entre niveaux de production et niveaux d'utilisation d'intrants. En particulier, l'approche duale permet de ne considérer que des données agrégées au niveau des exploitations et de négliger les allocations de surface (Paris, 1989), ce qui la rend applicable avec les données habituellement disponibles (Shumway, 1995). En outre, les modèles décrivant les choix des agriculteurs sont exprimés comme des fonctions des paramètres de la technologie et des prix. Ces prix sont des variables exogènes, ce qui permet l'utilisation de procédures d'estimation simples. En revanche, ces modèles ne permettent pas d'intégrer ou d'imposer facilement certaines propriétés de la technologie de production agricole ou des contraintes telles que des contraintes de rotation. Ils sont également « opaques ». Leurs paramètres n'ont en général pas d'interprétation directe, notamment en termes agronomiques, ce qui les rend difficilement utilisables dans le cadre de recherches pluridisciplinaires.

L'objectif de cet article est de présenter de nouveaux modèles de choix de production agricole dans un contexte de productions multiples. Ces modèles ont été conçus 1) pour être employés dans le cadre de recherches impliquant des économistes, des agronomes et des conseillers agricoles, 2) pour être intégrés en tant que modèles « de base » des choix de production dans des modèles économétriques plus complexes et 3) comme des « plate-formes » de modélisation potentiellement généralisables. Trois types de généralisations sont plus particulièrement envisagés : la prise en compte des solutions en coin dans les choix d'assolement, la prise en compte des effets de rotation dans l'optique d'une modélisation des choix « dynamiques » des agriculteurs et l'analyse (*ex ante* ou *ex post*) des changements de pratiques culturales. Ces

deux dernières généralisations requièrent une modélisation primale de la technologie de production culture par culture.

Les modèles proposés ici ont une structure simple et leurs paramètres admettent des interprétations directes, en termes agronomiques ou en matière de comportement des producteurs agricoles. Ces modèles ont quatre principales caractéristiques. 1) Ils reposent sur une formalisation primale de la technologie de production des cultures. 2) Les modèles de technologie de production sont « structurellement » aléatoires. 3) Ils combinent, à des degrés divers, des éléments des deux grandes approches de la modélisation empirique des choix de production agricole : la programmation mathématique (positive) et la modélisation économétrique. 4) Ces modèles ont une structure simple qui repose sur des hypothèses limitant leur utilisation pour des choix de production de court terme. Ces hypothèses sont explicitées et discutées dans l'article.

De fait, les modèles de choix de production présentés ici sont explicitement interprétés comme des approximations des « véritables » modèles de choix des agriculteurs, si tant est que ces modèles « véritables » existent. Ces hypothèses justifient d'ailleurs l'intérêt des généralisations envisagées. De manière générale, les modèles proposés ici reposent sur une spécification de la technologie agricole beaucoup moins générale ou flexible que celles considérées dans les modèles économétriques fondés sur la théorie duale du producteur. Ceci est dû aux contraintes imposées sur la forme de la technologie par la spécification de fonctions de production primales pour les différentes cultures considérées. En ce sens, les modèles proposés sont plus proches de ceux considérés par les économistes agricoles utilisant la programmation mathématique. Ils partagent avec ces modèles leur principale qualité : ils ont une structure facilement interprétable. Les modèles proposés ici sont cependant plus flexibles pour la modélisation des choix de court terme et pour intégrer des aspects

spécifiques de la technologie de production agricole que les approches habituellement considérées par les modèles économétriques dans un contexte multi-produits.

L'article est organisé comme suit. La première partie présente brièvement et compare les principaux modèles économétriques utilisés dans la littérature pour construire des modèles économiques de choix de production agricole dans un contexte de productions multiples. Ensuite la structure générale des modèles est présentée et les hypothèses posées sont discutées. La seconde partie présente les modèles de fonction de rendement qui fondent les choix des pratiques culturales des agriculteurs et la troisième présente les déterminants de leurs choix d'assolement tels qu'ils sont définis dans l'approche proposée ici. La quatrième partie présente les modèles de choix de production multi-produits que nous proposons.

1.2 Une brève revue de littérature

Cette partie présente la structure de la modélisation économique sous-jacente aux modèles économétriques des choix de production agricole considérant la terre comme un intrant fixe allouable. Elle a deux objectifs : 1) présenter les outils de modélisation économique mobilisés dans le reste de l'article et 2) présenter les limites et avantages des modèles habituellement utilisés pour ensuite les comparer à celles et ceux des modèles proposés ici.

Les notations suivantes seront utilisées tout au long de l'article. Les agriculteurs considérés produisent C cultures, $c = 1, \dots, C$. Y_c désigne la quantité produite de la culture c et y_c son rendement. Une unité de la culture c est vendue au prix p_c . \mathbf{X}_c est le vecteur des quantités d'intrants variable utilisées pour la culture c et \mathbf{x}_c sa contrepartie à l'hectare. Les agriculteurs utilisent K intrants variables et $\mathbf{X}_c \equiv (X_{1c}, \dots, X_{Kc})$. $\bar{\mathbf{X}}$ est le vecteur des quantités d'intrants variables utilisées sur l'exploitation, $\bar{\mathbf{X}} \equiv \sum_{c=1}^C \mathbf{X}_c$. Le vecteur des prix des intrants variables est noté \mathbf{w} . Le vecteur $\bar{\mathbf{X}}$ est généralement disponible dans les bases de données à la

disposition des économistes agricoles, l'allocation des intrants variables aux cultures $\mathbf{X} \equiv (\mathbf{X}_1, \dots, \mathbf{X}_C)$ l'est très rarement. Les quantités d'intrants variables sont généralement observées au niveau des exploitations mais pas leur allocation aux cultures, ce qui sera supposé dans la suite.

Un intrant est considéré comme fixe si sa quantité est fixée à court terme, *i.e.* si cette quantité ne s'ajuste qu'à la suite d'investissements. Deux types d'intrants fixes sont considérés ici : la terre et les autres intrants fixes, *i.e.* essentiellement le travail, les machines et les bâtiments (le capital). Cette distinction se justifie pour deux raisons : la terre est un intrant essentiel pour la production végétale et son allocation aux cultures de l'exploitation est généralement une donnée disponible, que les données soient au niveau des exploitations ou à un niveau agrégé. La surface de l'exploitation est notée \bar{S} et l'assolement est désigné par le vecteur \mathbf{S} avec $\sum_{c=1}^C S_c = \bar{S}$. L'assolement en parts est noté \mathbf{s} avec $s_c \equiv S_c / \bar{S}$ et $\sum_{c=1}^C s_c = 1$. Les quantités totales des L autres intrants fixes sont définies par $\bar{\mathbf{Z}} \equiv (\bar{Z}_1, \dots, \bar{Z}_L)$. L'allocation de l'intrant fixe ℓ est définie par le vecteur $\mathbf{Z}_\ell \equiv (Z_{\ell 1}, \dots, Z_{\ell C})$ pour $\ell = 1, \dots, L$ et $\mathbf{Z} \equiv (\mathbf{Z}_1, \dots, \mathbf{Z}_L)$. Le profit total des exploitation sera noté $\bar{\Pi}$. Ce profit se décompose en la somme de profits par culture : $\bar{\Pi} \equiv \sum_{c=1}^C \Pi_c$. La contre-partie de Π_c à l'hectare est notée π_c .

Les modèles de choix de production décrits dans cette section sont développés dans un cadre statique et sans incertitude, à l'instar de ceux développés dans la plupart des études citées. Les travaux de Just, Zilberman et Hochman (1983) font ici figure d'exception.

1.2.1 Les modèles duaux (Weaver, 1983 ; Chambers, 1989)

A partir du début des années 1980, beaucoup d'économistes agricoles ont utilisé les concepts de la dualité pour spécifier des modèles économiques de production à intrants et productions multiples (Weaver, 1983 ; Shumway, 1983 ; Chambers, 1989). L'intérêt de cette approche

réside dans sa simplicité d'utilisation : il suffit de spécifier une forme de fonction de profit indirecte et d'en dériver les fonctions d'offre de produits et de demande d'intrants à partir du lemme de Hotelling. Ces fonctions peuvent ensuite être utilisées pour définir des systèmes d'équations utilisables empiriquement. Ces équations sont « automatiquement » cohérentes du point de vue de la théorie de la production, par construction (voir Shumway, 1995, pour plus de détails).

La fonction de profit multi-produits indirecte est définie par :

$$(1) \quad \bar{\Pi}(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S}) \equiv \underset{\mathbf{Y}, \bar{\mathbf{X}}}{\text{Max}} [\mathbf{p}'\mathbf{Y} - \mathbf{w}'\bar{\mathbf{X}}] \quad \text{s.c.} \quad F(\mathbf{Y}, \bar{\mathbf{X}}; \bar{\mathbf{Z}}, \bar{S}) = 0$$

où $\mathbf{Y} \equiv (Y_1, \dots, Y_C)$ et $\mathbf{p} \equiv (p_1, \dots, p_C)$. Les contraintes de positivité des utilisations d'intrants et des niveaux de production sont ici négligées : toutes les cultures sont produites et tous les intrants, fixe ou variables, sont utilisés pour toutes les cultures. La contrainte d'utilisation de l'intrant fixe est supposée saturée, la terre étant nécessaire à la production des cultures. La technologie de production est représentée par la fonction de production implicite $F(\cdot)$. Cette fonction n'est en général pas explicitée. L'approche duale consiste à spécifier la forme de la fonction de profit indirecte, $\bar{\Pi}(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S})$, en choisissant une forme flexible. Plus la forme de $\bar{\Pi}(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S})$ l'est, plus la technologie implicitement décrite par $F(\mathbf{Y}, \bar{\mathbf{X}}; \bar{\mathbf{Z}}, \bar{S})$ est de forme générale. Il est important de noter ici que les allocations aux cultures des intrants variables ou de l'intrant fixe ne sont pas considérées. C'est ce qui confère sa simplicité à cette approche et son intérêt pratique : elle n'utilise que des variables généralement disponibles. En outre, $\bar{\mathbf{Z}}$ et \bar{S} étant fixes leur rôle se limite à « paramétrer » la forme de $F(\mathbf{Y}, \bar{\mathbf{X}}; \bar{\mathbf{Z}}, \bar{S})$, et donc celle de $\bar{\Pi}(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S})$.

Le lemme d'Hotelling permet alors déterminer les fonctions d'offre de produits et de demande d'intrants variables :

$$(2a) \quad \frac{\partial \bar{\Pi}(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S})}{\partial \mathbf{p}} = \mathbf{Y}^*(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S})$$

$$(2b) \quad \frac{\partial \bar{\Pi}(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S})}{\partial \mathbf{w}} = -\bar{\mathbf{X}}^*(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S}).$$

Ces fonctions d'offre de produits et de demande d'intrants variables sont fonctions des prix de toutes les cultures et des prix des intrants variables, et sont « paramétrées » par les quantités totales d'intrants fixes, ces données étant généralement disponibles.

Les modèles économétriques de choix de production sont dérivés de cette approche en ajoutant des termes d'erreur aux fonctions d'offre et de demande pour en déduire un système d'équations utilisable empiriquement. Les prix et les quantités d'intrants fixes étant supposés exogènes, les techniques de régression peuvent être employées.

Ces modèles qui ont été très utilisés à partir du début des années 1980 le sont encore, principalement avec des données agrégées au niveau des Etats pour calculer des élasticités utilisables dans des modèles d'équilibre général ou partiel. Ces modèles ont cependant plusieurs inconvénients, notamment pour leur utilisation sur données micro-économiques. Ils sont « opaques » en termes de technologie de production. Ils ne permettent pas de déterminer les allocations d'intrants aux cultures. En outre, ils n'exploitent pas l'information apportée par les allocations de surface aux cultures. Ce dernier point a motivé une série de travaux sur les modèles de choix de production avec la terre en tant que « facteur fixe allouable ».

Les intrants variables sont supposés être alloués par les agriculteurs à des activités de production spécifiques. Les intrants considérés fixes à court et moyen terme dans une exploitation mais allouables entre les activités de production, comme la terre, apparaissent comme une source de jointure potentielle importante qu'il faut prendre en compte

spécifiquement dans les modèles de production. Des technologies de production agricole non-jointes par les intrants mais caractérisées par l'existence de facteurs fixes mais allouables sont alors utilisées et décrites par des fonctions de production spécifiques à chaque culture ou par les relations duales correspondantes.

Les modèles définis à partir de ces hypothèses sont les plus couramment employés pour des études empiriques sur données micro-économiques. Trois modèles, en particulier, sont à la base des modélisations actuelles, l'approche primale de Just, Zilberman et Hochman (1983), les approches duales de Chambers et Just (1989) et de Moore et Negri (1992).

1.2.2 Le modèle de Just, Zilberman et Hochman (1983)

Just, Zilberman et Hochman (1983) construisent un modèle de production à partir d'une approche primale. La technologie de production agricole est caractérisée en trois points : la plupart des intrants agricoles (engrais, produits phytosanitaires, tracteurs, travail...) sont alloués par les agriculteurs à des activités de production spécifiques, des contraintes physiques limitent la quantité totale de certains intrants, comme la terre, et les combinaisons de productions sont déterminées uniquement par les allocations d'intrants aux différentes activités de production (mis à part les événements aléatoires). Le modèle des choix de production est alors défini par la maximisation d'une fonction de profit sous contrainte :

$$(3a) \quad \underset{Y, X, S}{\text{Max}} [\mathbf{p}'\mathbf{Y} - \mathbf{w}'\bar{\mathbf{X}}] \quad \text{s.c.} \quad Y_c = F_c(\mathbf{X}_c, \bar{\mathbf{Z}}, S_c) \quad \text{pour } c = 1, \dots, C \quad \text{et} \quad \sum_{c=1}^C S_c = \bar{S}$$

où :

$$(3b) \quad \bar{\mathbf{X}} = \sum_{c=1}^C \mathbf{X}_c .$$

La première série de contraintes correspond à la fonction de production de la culture c , qui dépend des quantités de l'intrant variable, de la terre allouée à cette culture et des quantités

disponibles des autres intrants fixes. La seconde contrainte correspond à la disponibilité de la terre en quantité fixe. L'équation (3b) définit $\bar{\mathbf{X}}$. Ce programme peut être résolu en deux étapes. La première consiste en la maximisation des profits par culture à assolements donnés :

$$(4) \quad \underset{Y_c, \mathbf{X}_c}{\text{Max}} [p_c Y_c - \mathbf{w}' \mathbf{X}_c] \quad \text{s.c.} \quad Y_c = F_c(\mathbf{X}_c, \bar{\mathbf{Z}}, S_c) \quad \text{pour } c = 1, \dots, C.$$

Cette étape permet alors de définir les fonctions de demande d'intrants et d'offre de produits conditionnelle en S_c :

$$(5a) \quad \mathbf{X}_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c) \equiv \underset{\mathbf{X}_c}{\text{Argmax}} [p_c F_c(\mathbf{X}_c, \bar{\mathbf{Z}}, S_c) - \mathbf{w}' \mathbf{X}_c],$$

$$(5b) \quad Y_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c) = F_c[\mathbf{X}_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c), \bar{\mathbf{Z}}, S_c]$$

et la fonction de profit indirecte restreinte en S_c (et en $\bar{\mathbf{Z}}$) de la culture c :

$$(5c) \quad \Pi_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c) \equiv p_c Y_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c) - \mathbf{w}' \mathbf{X}_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c)$$

pour $c = 1, \dots, C$. Cette approche est considérée comme primale puisqu'elle dérive les fonctions $Y_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c)$, $\mathbf{X}_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c)$ et $\Pi_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c)$ à partir d'un programme de maximisation du profit, *i.e.* à partir des fonctions de production $F_c(\cdot)$. Just, Zilberman et Hochman (1983) utilisent des fonctions de Cobb-Douglas. La seconde étape utilise les fonctions de profit par culture pour définir l'allocation de la terre :

$$(6) \quad \mathbf{S}^*(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S}) \equiv \underset{\mathbf{S}}{\text{Argmax}} \left\{ \sum_{c=1}^C \Pi_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c) \quad \text{s.c.} \quad \sum_{c=1}^C S_c = \bar{S} \right\}.$$

L'assolement optimal est caractérisé par les conditions du premier ordre du programme de maximisation :

$$(7) \quad \frac{\partial \Pi_c [p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c^*(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S})]}{\partial S_c} = \frac{\partial \Pi_1 [p_1, \mathbf{w}; \bar{\mathbf{Z}}, S_1^*(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S})]}{\partial S_1} \quad \text{pour } c = 2, \dots, C.$$

Les relations de dualité permettent alors de définir les fonctions de demande d'intrants variables et d'offre de produits :

$$(8a) \quad Y_c^*(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S}) = Y_c [p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c^*(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S})]$$

et :

$$(8b) \quad \mathbf{X}_c^*(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S}) = \mathbf{X}_c [p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c^*(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S})]$$

pour $c = 1, \dots, C$. Lorsque l'assolement optimal $\mathbf{S}^*(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S})$ est sous forme explicite, le modèle économétrique est relativement simple. Il est composé des équations d'assolement, d'offre et des demandes agrégées d'intrants :

$$(9a) \quad \begin{cases} S_c = S_c^*(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S}) + u_c^S, & c = 2, \dots, C \\ Y_c = Y_c^*(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S}) + u_c^Y, & c = 1, \dots, C \\ \bar{\mathbf{X}} = \sum_{c=1}^C \mathbf{X}_c^*(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S}) + \mathbf{u}^X \end{cases}$$

Ce système définit un système de régressions empilées et peut donc être estimé simplement. Lorsque $\mathbf{S}^*(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S})$ n'admet pas de forme explicite, le système d'équations considéré empiriquement est composé des équations d'offre conditionnelle, des équations de demande conditionnelle agrégée d'intrants et des conditions du premier ordre des choix d'assolement :

$$(9b) \quad \begin{cases} \partial \Pi_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c) / \partial S_c - \partial \Pi_1(p_1, \mathbf{w}; \bar{\mathbf{Z}}, S_1) / \partial S_1 = u_c^S, & c = 2, \dots, C \\ Y_c = Y_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c) + u_c^Y, & c = 1, \dots, C \\ \bar{\mathbf{X}} = \sum_{c=1}^C \mathbf{X}_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c) + \mathbf{u}^X \end{cases}$$

L'estimation de ce système requiert l'utilisation de techniques à variables instrumentales. En effet ce système est un système d'équations simultanées en raison de l'endogénéité de \mathbf{S} .

1.2.3 Les modèles de Chambers et Just (1989)

Chambers et Just (1989) développent deux modèles multi-produits à partir d'une approche duale, reprochant à la méthode utilisée par Just, Zilberman et Hochman (1983) l'impossibilité de considérer des technologies complexes.

Dans leur premier modèle, la production est supposée non-jointe par les intrants mais caractérisée par l'existence d'intrants fixes allouables. Bien que cette approche ait été utilisée pour présenter les travaux de Just, Zilberman et Hochman (1983), Chambers et Just (1989) sont les premiers à avoir défini les prises de décisions des producteurs en deux étapes : le producteur maximise le profit associé à chaque culture étant donné une allocation des intrants fixes donnée ; et il alloue ensuite les intrants fixes entre les cultures, il en résulte alors une fonction de profit multi-produits. L'approche de Chambers et Just (1989) se distingue de celle de Just, Zilberman et Hochman (1983) simplement parce que Chambers et Just utilisent une approche duale : ils spécifient les fonctions de profit indirectes des cultures, les $\Pi_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, S_c)$ pour $c = 1, \dots, C$ (ils choisissent des formes Translog) en en dérivent les fonctions de demande d'intrant et d'offre de production conditionnelle aux assolements.

Dans leur second modèle, ils supposent des jointures technologiques par les assolements. Dans ce modèle, la fonction de profit indirecte restreinte d'une culture dépend de la surface allouée à cette culture et des surfaces allouées aux autres cultures. En outre la fonction de profit indirecte (restreinte en \mathbf{S}) de l'exploitation prend une forme plus générale en intégrant des effets de \mathbf{S} non spécifiques aux cultures. Cette fonction de profit s'écrit alors :

$$(10) \quad \sum_{c=1}^C \Pi_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, \mathbf{S}) + H(\bar{\mathbf{Z}}, \mathbf{S}).$$

Le rôle du terme $H(\bar{\mathbf{Z}}, \mathbf{S})$ n'est pas clairement défini même s'il donne de la « flexibilité » aux modèles de choix dérivés de la fonction de profit décrite en (10). Il rapproche cependant la

forme de cette fonction de profit indirecte de celle utilisée dans l'approche « purement » duale, de la fonction de profit indirecte définie dans l'équation (1). Cette généralisation ne remet pas en cause la résolution du modèle en deux étapes. Les fonctions d'offre de production et de demande d'intrants variables conditionnelles sont alors définies par :

$$(11a) \quad \frac{\partial \Pi_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, \mathbf{S})}{\partial \mathbf{w}} = -\mathbf{X}_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, \mathbf{S})$$

et :

$$(11b) \quad \frac{\partial \Pi_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, \mathbf{S})}{\partial p_c} = Y_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, \mathbf{S})$$

pour $c = 1, \dots, C$. La non-jointure complique essentiellement la seconde étape qui devient :

$$(12a) \quad \text{Max}_{\mathbf{S}} \left[\sum_{c=1}^C \Pi_c(p_c, \mathbf{w}; \bar{\mathbf{Z}}, \mathbf{S}) + H(\bar{\mathbf{Z}}, \mathbf{S}) \right] \quad \text{s.c.} \quad \sum_{c=1}^C S_c = \bar{S}.$$

Et les conditions du premier ordre associées à ce programme sont alors définies par :

$$(12b) \quad \sum_{d=1}^c \frac{\partial \Pi_d(p_d, \mathbf{w}; \bar{\mathbf{Z}}, \mathbf{S}^*)}{\partial S_c} + \frac{\partial H(\bar{\mathbf{Z}}, \mathbf{S}^*)}{\partial S_c} = \sum_{d=1}^c \frac{\partial \Pi_d(p_d, \mathbf{w}; \bar{\mathbf{Z}}, \mathbf{S}^*)}{\partial S_1} + \frac{\partial H(\bar{\mathbf{Z}}, \mathbf{S}^*)}{\partial S_1}$$

pour $c = 2, \dots, C$ ou $\mathbf{S}^* \equiv \mathbf{S}^*(\mathbf{p}, \mathbf{w}; \bar{\mathbf{Z}}, \bar{S})$ désigne le choix optimal d'assolement.

1.2.4 Le modèle de Moore et Negri (1992)

Les modèles de choix de production agricole où la terre est un intrant fixe allouable sont encore peu utilisés dans les années 80 en raison de leur relative complexité. Les modèles de Just, Zilberman et Hochman (1983) et de Chambers et Just (1989) sont relativement difficiles à utiliser empiriquement. Leur principal défaut dans ce contexte est d'utiliser des fonctions de production de Cobb-Douglas ou des fonctions de profit indirectes de forme Translog. Ces formes fonctionnelles définissent les ln des niveaux de productions ou les ln des fonctions de

profit des cultures. Elles sont d'utilisation peu pratique dans des fonctions de profit au niveau de l'exploitation qui sont linéaires dans les niveaux de production et dans les niveaux de profit des cultures.

Moore et Negri (1992) ont proposé des modèles économétriques simples définis à partir de l'approche duale définie par Chambers et Just (1989). Leur principale contribution est de spécifier les fonctions de profit des cultures sous la forme d'une fonction quadratique (normalisée). L'avantage de cette forme fonctionnelle est double. Elle définit directement les niveaux de profit des cultures et ses dérivées premières sont linéaires dans leurs paramètres et leurs variables. Les assolements optimaux et les fonctions de demande d'intrants et d'offre de production peuvent s'écrire sous forme explicite. Les fonctions de demande d'intrants variables et d'offre de production conditionnelles sont en outre linéaires dans les assolements. Le modèle développé par Moore et Negri a beaucoup été utilisé dans les années 90, dans une série de travaux menés par Moore lui-même (Moore et al., 1994 ; Moore et Dinar, 1995). Ce modèle a également été utilisé pour analyser les impacts des réformes de la PAC sur l'allocation des terres pour les céréales et les oléoprotéagineux (Guyomard et *al.*, 1996 ; Moro et Sckokai, 1999). Dans le même but Oude Lansink et Peerlings (1996) généralisent l'approche de Moore et Negri en la combinant avec le modèle avec jointures par les assolements de Chambers et Just (1989).

1.2.5 Les jointures technologiques

Considérer la terre comme un facteur fixe allouable a soulevé une question qui a été à l'origine d'une série de travaux initiées par Shumway, Pope et Nash (1984) : les différentes productions d'une exploitation sont-elles jointes ? Les discussions relatives aux jointures dans les modèles de choix de production ont cependant eu tendance à être relativement confuses, d'une part à cause de l'absence d'une terminologie bien définie, et d'autre part en raison de certaines incohérences ou approximations dans la définition des concepts utilisés. Lynne

(1988) a en partie clarifié ces discussions en reformulant l'idée originale de Shumway, Pope et Nash (1984) qui peut être résumée ainsi : deux raisons non nécessairement exclusives expliquent que le niveau de production d'une culture puisse dépendre des prix des autres cultures : les interdépendances techniques dans la technologie multi-produits et l'existence de contraintes sur un intrant fixe allouable. Cette remarque soulève un problème d'identification de la source de l'interdépendance des niveaux de production par les prix : jointures technologiques ou/et existence d'intrants fixes allouables.

Dans le cas des intrants fixes allouables, le niveau de production d'une culture dépend des prix des autres cultures pour une raison simple : les agriculteurs allouent leur quantité fixe de terre (et des autres intrants fixes) aux différentes cultures en fonction de la profitabilité relative de ces cultures. Les cultures les plus profitables, grâce en particulier à des prix élevés, auront des assolements plus importants et donc des niveaux de production plus importants. Cependant, se pose la question des raisons pour lesquelles les agriculteurs ne produisent pas qu'une culture, la plus rentable. Ce point, lié à la diversification des productions, sera discuté en détail dans la suite car il est au cœur des travaux présentés ici.

Dans le cas des jointures technologiques, les choix de production des agriculteurs sont interdépendants car les technologies de production des différentes cultures le sont. Les jointures technologiques sont définies comme ce qui lie les niveaux de production de cultures différentes. Formellement, on parle de jointures technologiques lorsque la fonction de production d'une culture dépend des intrants et/ou des niveaux de production des autres cultures. Soit $F_c(.)$ la fonction de production de la culture c . La technologie de production multi-produits est non jointe si $Y_c = F_c(\mathbf{X}_c, \mathbf{Z}_c, S_c)$ pour $c = 1, \dots, C$. La technologie de production est jointe dans les intrants variables si $Y_c = F_c(\mathbf{X}, \mathbf{Z}_c, S_c)$ pour $c = 1, \dots, C$, *i.e.* lorsque les quantités d'intrants variables utilisées pour une culture peuvent affecter la production d'autres cultures. La technologie est jointe dans les intrants fixes si $Y_c = F_c(\mathbf{X}_c, \mathbf{Z}, \mathbf{S})$ pour $c = 1, \dots, C$. Enfin, la technologie est jointe dans les productions si

$Y_c = F_c(\mathbf{X}_c, \mathbf{Z}_c, S_c, \mathbf{Y}_{(-c)})$ pour $c = 1, \dots, C$ où $\mathbf{Y}_{(-c)}$ désigne le vecteur des Y_d pour $d = 1, \dots, C$ et $d \neq c$. Bien entendu ces hypothèses de jointures peuvent être combinées. Comme l'a noté Leathers (1991), ces jointures technologiques s'apparentent à des économies de gamme.

Les hypothèses de non-jointures dans les intrants variables, les intrants fixes autres que la terre et les niveaux de production sont généralement acceptées alors que, curieusement, la non-jointure ou la jointure de la technologie de production dans les assolements \mathbf{S} est plus débattue (Chambers et Just, 1989 ; Asunka et Shumway 1996). Deux arguments sont généralement avancés. Le premier argument est que d'autres intrants fixes allouables (travail, machines, ...) que la terre engendrent des jointures de production. Cet argument relève plus de considérations pratiques que théoriques : il est difficile de considérer l'allocation de plusieurs intrants fixes et seule l'allocation de la terre est observée généralement. Puisque \mathbf{Z} n'est pas observée et est difficile à identifier, les fonctions de production sont généralement (tout au moins implicitement) écrites sous la forme : $Y_c = G_c(\mathbf{X}_c, \mathbf{S}, \bar{\mathbf{Z}})$. Ceci provient de ce que la fonction de production est implicitement définie sous la forme $Y_c = F_c[\mathbf{X}_c, \mathbf{Z}_c(\mathbf{X}_c, \mathbf{S}, \bar{\mathbf{Z}}), S_c]$. La fonction $\mathbf{Z}_c(\mathbf{X}_c, \mathbf{S}, \bar{\mathbf{Z}})$ détermine la quantité d'intrants fixes (autres que la terre) allouée à c . \mathbf{Z}_c étant une fonction de \mathbf{S} , Y_c l'est également. Le second argument est lié aux effets des rotations culturales : les agriculteurs diversifient leurs cultures pour exploiter les effets bénéfiques des rotations culturales. L'invocation de cet argument repose cependant sur une incohérence liée à ce que le cadre d'analyse utilisé ici est statique alors que les effets des rotations culturales relèvent d'effets dynamiques de la production. En outre, les effets des rotations culturales ne dépendent pas que de la nature de la culture précédente, *i.e.* de l'assolement \mathbf{S} de l'année précédente, mais également de la manière dont elle a été produite. Par exemple, les reliquats d'azote laissés par une culture dépendent de la culture elle-même, des quantités d'engrais utilisées sur cette culture et de son rendement puisque ce rendement détermine les quantités d'azote « exportées ». Aussi, si les jointures technologiques proviennent des effets des rotations culturales ces jointures s'opèrent non

seulement par les assolements mais également par les intrants variables et les niveaux de productions. Malgré ces incohérences relevant essentiellement de considérations agronomiques, de nombreuses études empiriques ont cherché à tester des hypothèses de jointure technologique en production agricole (Chambers et Just, 1989 ; Asunka et Shumway 1996).

Arnberg et Hansen (2007) ont récemment proposé de résoudre cette incohérence par une approche originale. Schématiquement, cette approche consiste à définir la technologie de production non plus sur une seule saison mais sur un cycle complet de successions culturales. Les choix d'assolement se raisonnent alors sur un cycle complet et non pas année par année. Dans ce contexte les effets des rotations culturales peuvent être analysés comme des jointures technologiques pour des plans de production interannuels. Malheureusement, Arnberg et Hansen (2007) exploitent peu ce cadre théorique pour se concentrer sur des questions empiriques, intéressantes par ailleurs, mais finalement assez éloignées du cadre d'analyse présenté. En fait, ces auteurs font l'hypothèse que les schémas de rotation employés par les agriculteurs sont stationnaires (ou tout au moins stables) ce qui simplifie *in fine* l'analyse. Il suffit alors de raisonner « en moyenne » sur le cycle de rotation comme cela est réalisé dans un cadre statique. Cette hypothèse est en tous cas « pratique » pour définir des modèles empiriques qui tiennent compte des effets des rotations culturale sans avoir ni à en spécifier, ni à en mesurer explicitement les effets. Elle est certes restrictive mais apparaît somme toute raisonnable. Nos discussions avec les agronomes que nous consultons tendent à montrer que les agriculteurs utilisent des schémas de rotation relativement rigides, ou tout au moins assez stables. Ces schémas sont dictés par les caractéristiques de la distribution des conditions pédo-climatiques des exploitations et reposent sur des résultats établis par la recherche et le développement ou encore l'expérience des agriculteurs.

De cette discussion des jointures technologiques dans la production agricole (végétale ici), deux points nous semblent à ressortir. Tout d'abord, dans un cadre statique ces jointures ne peuvent être associées qu'à l'existence d'intrants fixes allouables, tels que le travail ou les machines non spécifiques à une culture (tracteurs, ...). Cependant, cette jointure n'est qu'indirecte dans le sens où le niveau de production de la culture c ne dépend du niveau de la culture d que *via* la concurrence de ces cultures pour un intrant disponible en quantité limitée. Ces jointures interviennent *via* les fonctions $Z_c(\mathbf{X}_c, \mathbf{S}, \bar{\mathbf{Z}})$ pour $c = 1, \dots, C$. Ce point s'avère important ici. Les fonctions de productions de la forme $Y_c = G_c(\mathbf{X}_c, \mathbf{S}, \bar{\mathbf{Z}})$ sont utilisées par les économistes. Les agronomes définissent des fonctions de production (de rendement) en n'utilisant en facteurs explicatifs de la production que des intrants ayant un rôle (productif ou protectif) sur la croissance et le développement de la plante cultivée, *i.e.* principalement (\mathbf{X}_c, S_c) . Les différences entre les fonctions de production des agronomes et des économistes sont d'autant plus importantes que la fonction décrivant les allocations d'intrants fixes, $Z_c(\mathbf{X}_c, \mathbf{S}, \bar{\mathbf{Z}})$, n'est pas toujours qu'une simple relation technique. Cette différence dans les notions de fonction de production des agronomes et des économistes s'avèrera essentielle dans la suite.

Ensuite, la discussion précédente montre que les arguments en faveur d'un effet direct de $\mathbf{S}_{(-c)}$ dans la fonction de production de la culture c sont pour le moins discutables. Il est cependant possible de justifier des fonctions de production intégrant des effets « directs » de $\mathbf{S}_{(-c)}$ pour approximer les effets des rotations culturales. Cette approximation repose sur l'hypothèse que les schémas de rotation utilisés par les agriculteurs sont relativement « stables ». Arnberg et Hansen (2007) utilisent explicitement ce type d'approximation. Dans ce cas, on sort du cadre « purement » théorique des jointures technologiques pour aborder celui de la spécification empirique des modèles de choix de production.

1.3 Les fonctions de production

Les modèles considérés ici sont développés dans un cadre simple dans la mesure où les agriculteurs 1) sont considérés comme neutres vis-à-vis du risque et 2) sont supposés raisonner dans un cadre statique. L'hypothèse de neutralité vis-à-vis du risque est bien entendu discutable et les modèles proposés ici sont *a priori* difficiles à adapter pour tenir compte d'éventuels importants niveaux d'aversion face au risque des agriculteurs. Les modèles développés ici peuvent être utilisés dans un cadre dynamique, notamment pour prendre en compte les effets des successions culturales et des comportements d'anticipation des agriculteurs. Enfin, il est supposé que les agriculteurs connaissent les prix des productions (\mathbf{p}) et des intrants (\mathbf{w}). L'information parfaite sur le prix des productions est une hypothèse discutable même si elle dépend des pays et des périodes concernés. Les agriculteurs étant considérés neutres vis-à-vis du risque, l'approche considérée ici peut cependant être généralisée pour tenir compte d'incertitudes sur les niveaux des prix des produits agricoles. Enfin, compte-tenu de la discussion précédente, la technologie de production multi-produits est supposée non-jointe dans les intrants variables, les niveaux de production et les assolements. Le cadre d'analyse utilisé reconnaît cependant l'existence d'autres intrants fixes allouables que la terre, essentiellement le capital et le travail.

L'approche utilisée ici se distingue de celle généralement utilisée sur plusieurs points.

Tout d'abord, les modèles construits ici visent à « accueillir » des résultats d'essais agronomiques ce qui impose certaines contraintes. Nos modèles sont construits à partir d'une représentation primale de la technologie de production agricole même s'ils reposent par la suite sur des concepts duaux, à l'instar des modèles de Just, Zilberman et Hochman (1983). Ce cadre d'analyse est plus rigide que le cadre d'analyse « purement » dual de Weaver (1983) ou Chambers (1989). De fait, nos modèles reposent sur des hypothèses qui limitent leur spectre d'application. Ces hypothèses sont explicitées et discutées.

Ensuite, le secteur de la production agricole, notamment celui de la production végétale, est caractérisé par une forte hétérogénéité. Cette dernière est largement inobservée, *i.e.* dans les bases de données usuelles il n'existe pas de variables observées qui permettent de transformer un modèle « théorique » valide pour une exploitation en un « modèle empirique » valide pour un échantillon d'exploitations. Cette hétérogénéité des conditions de production et le caractère aléatoire de la production agricole sont considérés directement dans la spécification des modèles « théoriques ». De fait, il n'existe pas de différences entre nos modèles « théoriques » et les modèles que nous utilisons dans les applications empiriques. La pratique usuelle des économistes agricoles a longtemps été de spécifier des modèles théoriques de choix de production valides pour une exploitation et de les transformer en modèles empiriques par l'introduction de quelques variables de contrôle (de l'hétérogénéité et de l'évolution du progrès technique) et l'addition de termes d'erreur. Les termes aléatoires des modèles proposés ici sont des parties intégrantes de ces modèles, ils sont analysés comme des variables inobservées ou des paramètres aléatoires à l'instar de ce préconise McElroy (1987) ou de ce qui fondait déjà l'originalité des travaux de Zellner, Kmenta et Drèze (1966). En ce sens, l'approche adoptée ici s'inscrit dans le courant de ce qu'on nomme désormais « l'économétrie structurelle ». Il est cependant important de noter ici que cette approche tend à minimiser le caractère nécessairement imparfait des modèles spécifiés.

Cette partie s'articule en trois sections. La première présente les difficultés associées à la spécification de modèles permettant un dialogue entre les économistes d'une part, et les agronomes et les techniciens qui conseillent les agriculteurs d'autre part. Cette section a pour objectif de dégager des pistes visant à poser les bases d'un dialogue entre disciplines autour de modèles des choix de production des agriculteurs. La seconde section présente la notion de fonction de production qui sera utilisée dans la suite. La troisième section décrit les modèles de choix d'intrants variables et des niveaux de production, culture par culture.

1.3.1 Modéliser les choix des agriculteurs : combiner les apports de l'économie et de l'agronomie

Les discussions entre les économistes agricoles d'une part, et les agronomes et les techniciens qui conseillent les agriculteurs d'autre part, sont souvent difficiles pour trois raisons principales.

Tout d'abord, les agronomes et les techniciens agricoles raisonnent en termes de marges brutes, de rendements objectifs, d'utilisations d'intrants, ... Les économistes raisonnent plus généralement en termes d'effets marginaux, ces derniers définissant les conditions d'optimalité des choix des agriculteurs. L'utilisation de résultats d'essais agronomiques dans des modèles de choix de production suppose un travail d'adaptation soit des résultats d'essais pour en dériver des effets « à la marge », soit des modèles des économistes pour intégrer des effets à la fois en niveau et à la marge.

Ensuite, la notion de fonction de production n'est pas utilisée en agronomie, tout au moins dans son acception en économie agricole. Les notions de frontières de production ou de fonction implicite de technologie multi-produits sont considérées comme des concepts très théoriques par les agronomes, et ce à juste titre à certains égards. Pour caractériser les choix des agriculteurs ces derniers font essentiellement appel aux notions de marges de brutes, de règles de décision et de contraintes. Ceci explique d'ailleurs l'utilisation quasi-exclusive des modèles de la programmation mathématique dans les programmes de recherche impliquant des économistes et des membres des disciplines biotechniques. Les modèles des agronomes qui se rapprochent le plus des fonctions de production des économistes agricoles sont des modèles décrivant comment les pratiques des agriculteurs (choix d'intrants, des dates d'intervention, ...) et les conditions de production (conditions pédo-climatiques, conditions sanitaires...) contribuent à l'élaboration des éléments du rendement d'une culture. Ces modèles, plus ou moins complexes, visent à représenter les processus biologiques et bio-

physiques gouvernant la croissance et le développement des plantes cultivées. Ces processus sont sensés être résumés (*via* les quantités d'intrants variables : engrais, pesticides et semences essentiellement) dans les fonctions de production des économistes. Ces dernières incluent, de manière toute aussi approximative, d'autres opérations des agriculteurs : les applications d'intrants, le travail du sol, (*via* les quantités de carburant et le capital matériel).

Enfin, les agronomes et les techniciens agricoles, ont souvent beaucoup de mal à admettre l'aspect « statique » de la plupart des modèles des économistes agricoles. Spécifiquement dans le cas des grandes cultures, celui considéré ici, les agronomes et les techniciens agricoles, et donc les agriculteurs, raisonnent en termes de schéma de rotations culturales ou de successions culturales.

La plupart des rares économistes ayant abordé la question des rotations culturales (Eckstein, 1984 ; Ozarem et Miranowski, 1994 ; Thomas, 2003) ont utilisé les principes de l'optimisation dynamique dans une optique marginaliste, i.e. en fondant leurs modèles empiriques sur les équations d'Euler. Cette approche est complexe et ne permet de traiter que de problèmes spécifiques, essentiellement la gestion des reliquats d'engrais. Or les effets des rotations culturales concernent également la gestion du risque phytosanitaire et la gestion de la structure du sol. Considérer ces effets selon les principes de l'optimisation dynamique s'avère très complexe. Si la gestion des reliquats d'engrais peut s'appuyer sur une approche marginaliste, la gestion du risque phytosanitaire ou de la structure du sol repose sur des choix partiellement discrets. La gestion des reliquats d'engrais relève des mêmes principes quelles que soient les cultures, seule importe la quantité d'engrais mise à la disposition de la plante cultivée par la culture précédente. La gestion de l'état phytosanitaire des parcelles ou de la structure du sol dépend du couple culture actuelle/culture précédente. Dans ce cas, les choix d'assolement optimaux sont caractérisés par des choix discrets, des « coudes » et des « trous ». Seuls Livingston, Roberts et Rust (2008) considèrent des effets généraux de rotation. Leur

approche est essentiellement normative et se focalise uniquement sur des choix à la parcelle dans le cadre d'un modèle de choix discret dynamique. Les analyses citées ont cependant deux défauts importants. Elles ne tiennent pas compte de la volatilité des prix des productions agricoles et ne considèrent que des effets dynamiques d'ordre 1. Si la prise en compte des variations de prix paraît envisageable, celle des effets de rotation dépassant un an paraît beaucoup plus difficile.

De cette discussion ressortent deux aspects de la dynamique des choix d'assolement des producteurs. Les schémas de rotations utilisés par les agriculteurs sont relativement stables et une solution appliquant *stricto sensu* le principe de rationalité est difficilement envisageable. Les économistes agricoles utilisant la programmation mathématique intègrent ces deux points en spécifiant dans leurs modèles des contraintes sur les choix d'assolement. Ces contraintes visent à représenter les effets des règles de décision imposées par la gestion des effets des rotations culturales. Par exemple, certaines rotations sont très fortement déconseillées en raison de leurs effets négatifs sur la structure du sol ou sur l'état sanitaire des parcelles. Certaines cultures ont un important délai de retour sur les parcelles pour les mêmes raisons. Certaines successions culturales n'ont aucun intérêt car d'autres leur sont toujours préférables. Les modèles spécifiés ici reposent sur ces bases, mais utilisent d'autres outils de modélisation que la spécification de contraintes.

1.3.2 Les fonctions de production

Les modèles spécifiés ici s'articulent tout d'abord autour d'un ensemble de fonctions de production, une fonction de production étant définie pour chaque culture. Aussi, pour $c = 1, \dots, C$, le rendement de la culture c , y_c , est défini à partir de la fonction de rendement déterministe :

$$(13) \quad y_c = f_c(\mathbf{x}_c, s_c; \mathbf{m}_c, \boldsymbol{\mu}_c, \zeta_c)$$

dépendant du vecteur des quantités d'intrants variables utilisées par unité de surface pour la culture c , \mathbf{x}_c , et s_c la part de surface allouée à la culture c . Afin de rapprocher les fonctions de production utilisées ici de celles considérées en agronomie, ne sont considérés dans \mathbf{x}_c que les quantités d'intrants jouant un rôle direct dans les processus de croissance et de développement de la plante cultivée : les engrais, les produits phytosanitaires et les semences. L'introduction d'un facteur d'échelle représenté par s_c est discuté dans la suite.

Les termes $(\mathbf{m}_c, \boldsymbol{\mu}_c, \zeta_c)$ représentent des éléments, inobservés par les économistes mais observés par les agriculteurs, ayant un rôle dans le processus de production. Bien entendu, ces éléments ne peuvent être utilisés dans une modélisation empirique mais la spécification en amont de la fonction de rendement « théorique » (13) permet d'analyser la fonction de rendement « empirique » qui sera spécifiée dans la suite. Les termes $(\mathbf{m}_c, \boldsymbol{\mu}_c, \zeta_c)$ caractérisent la forme des relations entre le rendement et les intrants utilisés.

Le vecteur $\boldsymbol{\mu}_c$ décrit les éléments connus de l'agriculteur au moment où il met en place sa culture. Ces éléments peuvent être des conditions de production de l'exploitation considérées permanentes, ou tout au moins évoluant lentement, comme la qualité des sols ou être spécifiques à la campagne de production à venir comme l'état sanitaire initial des parcelles ou le niveau des reliquats laissés par les cultures précédentes. Le vecteur ζ_c décrit les événements aléatoires affectant les niveaux de productions : éléments climatiques, infestations, maladies, déprédations, ... L'agriculteur « découvre » ζ_c , plus ou moins progressivement, après avoir mis en place sa culture. Dans certains cas il peut « corriger » ou « exploiter » les effets de ζ_c , dans d'autres cas il subit ces effets.

Le vecteur \mathbf{m}_c décrit les pratiques culturales des agriculteurs non décrites par \mathbf{x}_c : dates et fractionnement des apports d'azote, dates des semis, ... Les décisions décrites par ces variables

n'ont en général pas de coût comptable. Comme cela sera vu par la suite, cette spécification pose la question des dépenses de carburants (et lubrifiants) associées aux applications des intrants variables dont les quantités sont mesurées dans \mathbf{x}_c et associées au travail du sol ou à la récolte. Contrairement à $\boldsymbol{\mu}_c$ et ζ_c qui sont des variables exogènes, la variable \mathbf{m}_c est endogène puisqu'elle décrit des choix des agriculteurs.

Les économistes agricoles ont depuis longtemps reconnu l'existence des aléas de la production agricole, ceux représentés par ζ_c . Cependant, ils l'ont fait pour l'analyse de points spécifiques tels que l'utilisation des pesticides (Feder, 1979) ou la gestion des risques de production afin d'analyser les effets de l'attitude face au risque des agriculteurs sur leurs choix de production (Antle, 1983a ; Moschini et Hennessy, 2001). L'hétérogénéité des conditions de production est plus rarement considérée. Or, et c'est là l'un des apports majeurs de l'économétrie des données de panel, cette hétérogénéité s'avère très importante pour l'analyse des choix de production (Griliches et Mairesse, 1995 ; Blundell et Bond 2000). Carpentier et Weaver (1996) ont montré, à partir d'un panel de 700 céréaliculteurs français suivis de 1987 à 1991, que 85% de la variation des niveaux de production est dû aux variations entre exploitations, *i.e.* dû aux variations de $\boldsymbol{\mu}_c$ et des effets moyens (à l'échelle des exploitations) de ζ_c (et des variations de \mathbf{m}_c associées). Ils ont par ailleurs montré (Carpentier et Weaver, 1997) que négliger cette hétérogénéité conduit à biaiser les estimations de la productivité des intrants agricoles. Dans le cadre de la formalisation adoptée ici ces biais s'expliquent par les relations existant entre \mathbf{x}_c , \mathbf{m}_c et $\boldsymbol{\mu}_c$: les agriculteurs adaptent leurs choix d'intrants et leurs pratiques de production aux conditions de leur exploitation. Ce résultat illustre empiriquement les travaux théoriques de Zellner, Kmenta et Drèze (1966).

Les termes $(\mathbf{m}_c, \boldsymbol{\mu}_c, \zeta_c)$ distinguent les données utilisées par les agronomes de celles utilisées par les économistes agricoles (Just et Pope 2001). De fait, les effets de $(\mathbf{m}_c, \boldsymbol{\mu}_c, \zeta_c)$ sont

rarement disponibles pour les économètres mais leur existence doit être reconnue dans la spécification des fonctions de production empiriques, et ce même si ces fonctions ne peuvent être que des approximations sommaires de l'effet des intrants sur les rendements. Ces dernières doivent essentiellement tenir compte 1) de l'hétérogénéité des conditions de production, 2) des aléas affectant les niveaux de production et 3) des possibilités d'adaptation des choix des agriculteurs aux conditions de production (hétérogénéité « permanente » ou aléas « spécifiques » d'une campagne de production).

Bien entendu, la fonction de rendement $y_c = f_c(\mathbf{x}_c, s_c; \mathbf{m}_c, \boldsymbol{\mu}_c, \boldsymbol{\zeta}_c)$ n'a que peu d'intérêt d'un point de vue empirique en raison des effets des termes $(\mathbf{m}_c, \boldsymbol{\mu}_c, \boldsymbol{\zeta}_c)$. Nous proposons ici une spécification de fonction de rendement permettant, de manière approximative, de tenir compte des effets de $(\mathbf{m}_c, \boldsymbol{\mu}_c, \boldsymbol{\zeta}_c)$ tout en étant d'utilisation relativement aisée dans le cadre de travaux économétriques. La spécification proposée a la forme générique suivante :

$$(14a) \quad y_c = a_c(\mathbf{q}_c, s_c) + e_c^y + \varepsilon_c^y - g_c(\mathbf{x}_c - \mathbf{b}_c(\mathbf{q}_c, s_c) - \mathbf{e}_c^x - \boldsymbol{\varepsilon}_c^x; \mathbf{q}_c)$$

avec :

$$(14b) \quad g_c(\mathbf{x}_c - \mathbf{b}_c(\mathbf{q}_c, s_c) - \mathbf{e}_c^x - \boldsymbol{\varepsilon}_c^x; \mathbf{q}_c) \geq 0 \text{ et } g_c(\mathbf{0}, \mathbf{q}_c) = 0.$$

Le vecteur \mathbf{q}_c est un vecteur de variables explicatives contenant les mesures disponibles des effets des termes $(\mathbf{m}_c, \boldsymbol{\mu}_c, \boldsymbol{\zeta}_c)$. Ce sont typiquement des variables de contrôle de l'hétérogénéité selon la terminologie de l'économétrie (Wooldridge, 2003). Les fonctions $a_c(\cdot)$, $\mathbf{b}_c(\cdot)$ et $g_c(\cdot)$ sont supposées de formes connues. Les termes $\mathbf{e}_c \equiv (e_c^y, \mathbf{e}_c^x)$ et $\boldsymbol{\varepsilon}_c \equiv (\varepsilon_c^y, \boldsymbol{\varepsilon}_c^x)$ représentent les effets de conditions de production non observées.

Les conditions (14b) sont des conditions de normalisation imposées pour faciliter l'interprétation de la forme de la fonction de rendement retenue. En particulier, lorsque les intrants variables sont utilisés en quantités $\mathbf{x}_c = \mathbf{b}_c(\mathbf{q}_c, s_c) + \mathbf{e}_c^x + \boldsymbol{\varepsilon}_c^x$ le rendement obtenu est

donné par $y_c = a_c(\mathbf{q}_c, s_c) + e_c^y + \varepsilon_c^y$. La fonction $g_c(\cdot; \mathbf{q}_c, s_c)$ étant positive pour tout \mathbf{x}_c , le terme $a_c(\mathbf{q}_c, s_c) + e_c^y + \varepsilon_c^y$ peut s'interpréter comme le rendement maximum qui peut être atteint avec les conditions de production décrites par \mathbf{q}_c , \mathbf{e}_c et $\boldsymbol{\varepsilon}_c$ (où $\mathbf{e}_c \equiv (e_c^y, \mathbf{e}_c^x)$ et $\boldsymbol{\varepsilon}_c \equiv (\varepsilon_c^y, \boldsymbol{\varepsilon}_c^x)$). Les quantités d'intrants variables nécessaires pour atteindre ce rendement maximum sont alors définies par $\mathbf{b}_c(\mathbf{q}_c, s_c) + \mathbf{e}_c^x + \boldsymbol{\varepsilon}_c^x$.

Ces notions de « rendement maximum » et de « quantités d'intrants » nécessaires pour atteindre ce rendement maximum ont des contre-parties quasi-directes en agronomie. Le rendement maximum renvoie à la notion de potentiel d'une culture pour des conditions de production données. Les quantités d'intrants associées renvoient aux quantités d'intrants utilisées dans le cadre de pratiques de production très intensives. Les quantités d'intrants \mathbf{x}_c et \mathbf{m}_c sont adaptées aux aléas de production (ζ_c) selon des règles de décision : traiter une infestation si nécessaire, irriguer si nécessaire, ajuster les apports d'azote en fonction des conditions climatiques, ... Les quantités d'intrants \mathbf{x}_c et \mathbf{m}_c sont également adaptées aux conditions de production spécifiques de l'exploitation. Les « rendements objectifs » des cultures, et donc les quantités d'intrants et les pratiques utilisées, sont établis en fonction des « potentialités » du milieu (profondeur, structure et nature des sols ; conditions climatiques habituelles) et des précédents cultureux. De même, les pratiques culturales et les quantités d'intrants utilisées dépendent de l'histoire culturelle de la parcelle : un blé succédant à un blé sera en général plus protégé contre les maladies, une culture de pois permet de réduire les apports d'azote pour la culture suivante, une culture de betterave ou de pomme de terre autorise des rendements objectifs élevés pour le blé de l'année suivante grâce aux effets bénéfiques de ces cultures sur la structure du sol et sur l'état sanitaire des parcelles.

Les maxima des rendements et les quantités d'intrants associées sont définies comme des fonctions des conditions de production (et de s_c ce qui sera discuté dans la suite) observées,

(\mathbf{q}_c, s_c) , ou non, $(\mathbf{e}_c, \varepsilon_c)$. La fonction de rendement décrite par les équations (14) ne peut être dérivée en toute généralité de celle décrite dans l'équation (3). Cependant, les termes $(\mathbf{q}_c, \mathbf{e}_c, \varepsilon_c)$ sont très reliés aux termes $(\mathbf{m}_c, \boldsymbol{\mu}_c, \zeta_c)$.

Le contenu du vecteur des conditions de production observées \mathbf{q}_c est très limité dans les bases de données utilisées par les économistes. Il se limite généralement à un ensemble d'indicateurs régionales. Dans le cas de données de panel, il peut également contenir des observations des choix passés. Il peut être défini comme une partie de $\boldsymbol{\mu}_c$. Les termes \mathbf{e}_c et ε_c ne sont pas « identifiables » séparément mais les distinguer est important pour la modélisation des choix d'assolement.

Le terme \mathbf{e}_c est supposé connu des agriculteurs lors de la mise en place de la culture c , donc au moment de ses choix d'assolement. Le terme ε_c est supposé inconnu de l'agriculteur au moment de ses choix d'assolement. Schématiquement, le terme \mathbf{e}_c contient les effets de $\boldsymbol{\mu}_c$ et les effets « en moyenne » de \mathbf{m}_c et ζ_c alors que ε_c contient les effets « hors moyenne » de \mathbf{m}_c et ζ_c . Les termes \mathbf{e}_c et ε_c représentent des effets très différents. Le terme \mathbf{e}_c « capture » essentiellement l'hétérogénéité des conditions de production, ce qui distingue « en moyenne » les conditions de production des exploitations. Le terme ε_c représente les effets « hors-moyenne » des aléas affectant la production de la culture c . Les termes \mathbf{e}_c et ε_c sont considérés comme aléatoires dans la spécification du modèle parce que qu'ils ne sont pas observés des économètres.

Le terme \mathbf{e}_c est analogue aux « effets individuels fixes » de l'économétrie des données de panel, même si ici il n'est pas nécessairement constant dans le temps pour une même exploitation. Le terme ε_c est lui analogue aux chocs de productivité utilisés dans les spécifications des fonctions de production dans l'économétrie des données de panel.

L'originalité de cette fonction de production est double. D'une part, cette spécification considère un ensemble de termes aléatoires et non pas un seul terme d'erreur. Ceci distingue la fonction de rendement (13) de celle de Just et Pope (1979) qui est celle généralement employée pour tenir compte des aléas de la production agricole. D'autre part, cette spécification distingue explicitement les effets de l'hétérogénéité des conditions de production des aléas affectant la production.

En utilisant la fonction de rendement (13) comme référence, il apparaît que les formes des fonctions $a_c(\cdot)$, $\mathbf{b}_c(\cdot)$ et $g_c(\cdot)$ dépendent de la forme de la fonction $f_c(\cdot)$ mais également des termes $(\mathbf{m}_c, \boldsymbol{\mu}_c, \zeta_c)$. La principale approximation liée à la transformation de la fonction de rendement donnée en (13) en celle donnée en (14) concerne la fonction $g_c(\cdot)$ qui est supposée « fixe » *modulo* les effets d'hétérogénéité capturés (généralement modestement) par \mathbf{q}_c . Dans la fonction de rendement (14), les effets inobservés des conditions de production sont concentrés dans les termes aléatoires \mathbf{e}_c et ε_c dans une simple logique de « translation » des niveaux de production et des niveaux d'utilisation d'intrant.

Une fonction de rendement avec la même forme que la fonction (13) mais avec une autre construction des fonctions $a_c(\cdot)$, $\mathbf{b}_c(\cdot)$ et $g_c(\cdot)$ et des termes aléatoires \mathbf{e}_c et ε_c offre une alternative possible. Dans le modèle considéré ici le terme \mathbf{e}_c représente des effets « en moyenne » et le terme ε_c des effets « hors moyenne » à l'instar des modèles spécifiés en économétrie des données de panel. Dans cette spécification alternative les termes ε_c représentent toujours les effets des aléas de production. Si on impose que ε_c^y est négatif alors que ε_c^x est positif, le terme $a_c(\mathbf{q}_c, s_c) + e_c^y$ s'interprète comme le rendement maximum pour une exploitation caractérisée par les conditions $(\mathbf{q}_c, s_c, \mathbf{e}_c)$ et ε_c^y mesure la perte de rendement liés aux aléas de la campagne considérée. Le terme $\mathbf{b}_c(\mathbf{q}_c, s_c) + \mathbf{e}_c^x$ s'interprète alors comme la quantité d'intrant minimale requise pour obtenir le rendement maximum sur l'exploitation

considérée et le terme ε_c^x donne les quantités supplémentaires d'intrants nécessaires pour obtenir le rendement maximum au cours de la campagne considérée. Cette spécification est à rapprocher de celles utilisées pour la mesure des inefficacités techniques.

La forme fonctionnelle décrite dans l'équation (14) peut paraître restrictive à certains égards. Elle s'avère pourtant plus générale que celles généralement utilisées par les économistes agricoles et, surtout, s'avère très pratique d'utilisation. Deux spécificités de la fonction de rendement (14) s'avèrent cruciales : l'additivité des termes aléatoires ε_c d'une part, et la distinction entre les termes aléatoires spécifiques des intrants ε_c^x et du terme aléatoire spécifique du niveau de rendement ε_c^y .

1.3.3 Fonctions de demande d'intrants, d'offre de production et fonctions de profit indirectes

Dans l'équation de rendement (14), le terme ε_c^y mesure l'effet des aléas que l'agriculteur ne peut contrôler ou compenser. Les effets des événements aléatoires résumés dans ε_c^y sont intégralement subis par l'agriculteur. Par construction, ils s'ajoutent ou se retranchent au rendement moyen pour donner le rendement obtenu. Les effets des événements aléatoires résumés dans ε_c^x sont des effets qui peuvent être compensés par des utilisations d'intrants ou peuvent compenser des utilisations d'intrants. Par construction, les effets ε_c^x peuvent être contrôlés par les agriculteurs. Il est cependant à noter ici que si les termes ε_c^x et ε_c^y ont des rôles très différents dans la fonction de rendement ils sont par nature très fortement liés.

Ces propriétés de la fonction de rendement (14) ont des implications directes pour la modélisation des choix des producteurs. Il est ici supposé que $g_c(\cdot; \mathbf{q}_c)$ est strictement décroissante et convexe en \mathbf{x}_c ce qui assure que la fonction de rendement est strictement croissante et concave en \mathbf{x}_c . Cette fonction est également supposée continûment

différentiable d'ordre de 2. Les agriculteurs étant supposés neutres vis-à-vis du risque, leurs choix d'intrants et de rendement objectif résultent d'un programme de maximisation de l'espérance de profit. Cependant ce programme doit tenir compte de ce que les agriculteurs vont observer ε_c^x et ε_c^y en cours de campagne et pourront donc ajuster leurs choix d'intrants. Afin de simplifier la suite, les arguments (\mathbf{q}_c, s_c) sont omis et $\boldsymbol{\beta}_c$ et α_c sont définis par $\boldsymbol{\beta}_c \equiv \mathbf{b}_c(\mathbf{q}_c, s_c) + \mathbf{e}_c^x$ et $\alpha_c \equiv a_c(\mathbf{q}_c, s_c) + e_c^y$.

Il est ici supposé que \mathbf{x}_c est composé de deux sous-vecteurs, $\mathbf{x}_{b,c}$ représente les quantités des intrants qui doivent être utilisés en début de campagne (*e.g.*, les semences et les herbicides) et $\mathbf{x}_{a,c}$ représente les quantités des intrants qui doivent être utilisés ou qui peuvent être ajustés en cours de campagne (*e.g.*, les engrais et les pesticides). Les vecteurs $\boldsymbol{\beta}_c$, \mathbf{e}_c^x et $\boldsymbol{\varepsilon}_c^x$ sont scindés en sous-vecteurs correspondants. Le programme de maximisation s'écrit donc :

$$(15) \quad \text{Max}_{\mathbf{x}_{b,c}} \left\{ E_{\boldsymbol{\varepsilon}_c} \left[\text{Max}_{\mathbf{x}_{a,c}^*(\cdot)} \left[p_c \left(\alpha_c + \varepsilon_c^y - g_c \left[\begin{array}{c} \mathbf{x}_{b,c} - \boldsymbol{\beta}_{b,c} \\ \mathbf{x}_{a,c}^*(\boldsymbol{\varepsilon}_c) - \boldsymbol{\beta}_{a,c} - \boldsymbol{\varepsilon}_{a,c}^x \end{array} \right] \right) - \mathbf{w}' \left[\begin{array}{c} \mathbf{x}_{b,c} \\ \mathbf{x}_{a,c}^*(\boldsymbol{\varepsilon}_c) \end{array} \right] \right] \right] \right\}.$$

Sa résolution à rebours en deux étapes est simple dès lors qu'il est supposé que la distribution de $\boldsymbol{\varepsilon}_c$ ne dépend pas de \mathbf{x}_c et que $\boldsymbol{\varepsilon}_{b,c} \equiv \mathbf{0}$, ce qui sera considéré ici. L'hypothèse d'indépendance est peu restrictive pour ce qui concerne l'espérance de $\boldsymbol{\varepsilon}_c$ car elle correspond alors à une simple normalisation (qui conduit essentiellement à redéfinir $g_c(\cdot)$). Elle est plus discutable pour ce qui concerne les moments de $\boldsymbol{\varepsilon}_c$ d'ordre supérieurs ou égal à 2. Elle simplifie cependant énormément la résolution du programme en raison de la non-linéarité de $g_c(\cdot)$ en $\boldsymbol{\varepsilon}_{a,c}^x$.

Les choix d'intrants de début de campagne peuvent par exemple modifier la gravité ou la fréquence de certaines maladies. L'hypothèse d'indépendance de $\boldsymbol{\varepsilon}_c$ par rapport à \mathbf{x}_c contraint

la prise en compte de tels effets, même si ces effets peuvent partiellement être représentés par la forme de $g_c(\cdot)$. Le modèle de choix séquentiel des intrants considéré par Antle (1983b) et Antle et Hatchett (1986) est légèrement moins restrictif de ce point de vue mais il est beaucoup plus rigide que celui considéré ici : il s'appuie sur un choix de forme fonctionnelle dont les spécificités (notamment propriétés de séparabilité) sont exploitées pour sa résolution. Cette hypothèse constitue un autre exemple des hypothèses associées à la « transformation » de la fonction de rendement générale (13) en fonction de forme spécifique (14).

La première étape de la résolution considère le choix de $\mathbf{x}_{a,c}$ conditionnellement aux choix de $\mathbf{x}_{b,c}$. Le choix optimal conditionnel $\mathbf{x}_{a,c}$ est défini, après simplifications, par :

$$(15a) \quad \text{Arg min}_{\mathbf{x}_{a,c}} \left\{ p_c g_c \left[\begin{bmatrix} \mathbf{x}_{b,c} - \boldsymbol{\beta}_{b,c} \\ \mathbf{x}_{a,c} - \boldsymbol{\beta}_{a,c} - \boldsymbol{\varepsilon}_{a,c}^x \end{bmatrix} \right] + \mathbf{w}'_a \mathbf{x}_{a,c} \right\}, \quad \forall (\boldsymbol{\varepsilon}_c; \mathbf{x}_{b,c}).$$

La forme de la fonction de rendement (13) et l'hypothèse d'indépendance de $\boldsymbol{\varepsilon}_c$ par rapport à \mathbf{x}_c permettent d'éliminer l'espérance de $\boldsymbol{\varepsilon}_c^y$. Les propriétés supposées de $g_c(\cdot)$ et de ses dérivées en \mathbf{x}_c impliquent que ce choix optimal a la forme suivante :

$$(15b) \quad \mathbf{x}_{a,c}(\mathbf{w}_{a,c}; \boldsymbol{\beta}_{a,c} + \boldsymbol{\varepsilon}_{a,c}, \mathbf{x}_{b,c} - \boldsymbol{\beta}_{b,c}) = \boldsymbol{\beta}_{a,c} + \boldsymbol{\varepsilon}_{a,c}^x - \boldsymbol{\gamma}_{a,c}(\mathbf{w}_{a,c}; \mathbf{x}_{b,c} - \boldsymbol{\beta}_{b,c}) \text{ avec } \mathbf{w}_{a,c} \equiv \mathbf{w}_a / p_c$$

et où $\boldsymbol{\gamma}_{a,c}(\cdot)$ est la fonction réciproque en $\mathbf{x}_{a,c} - \boldsymbol{\beta}_{a,c} - \boldsymbol{\varepsilon}_{a,c}^x$ de l'opposé du gradient de $g_c(\cdot)$ en $\mathbf{x}_{a,c}$. La seconde étape de résolution définit alors le choix optimal de $\mathbf{x}_{b,c}$ conditionnellement aux choix optimaux et adaptatifs de $\mathbf{x}_{a,c}$, après simplifications, par :

$$(15c) \quad \text{Arg min}_{\mathbf{x}_{b,c}} \left\{ p_c g_c \left[\begin{bmatrix} \mathbf{x}_{b,c} - \boldsymbol{\beta}_{b,c} \\ -\boldsymbol{\gamma}_{a,c}(\mathbf{w}_{a,c}; \mathbf{x}_{b,c} - \boldsymbol{\beta}_{b,c}) \end{bmatrix} \right] + \mathbf{w}'_b \mathbf{x}_{b,c} \right\}.$$

La forme de $\mathbf{x}_{a,c}(\mathbf{w}_{a,c}; \boldsymbol{\beta}_{a,c} + \boldsymbol{\varepsilon}_{a,c}, \mathbf{x}_{b,c} - \boldsymbol{\beta}_{b,c})$, en particulier son additivité en $\boldsymbol{\varepsilon}_{a,c}^x$, et l'hypothèse d'indépendance de $\boldsymbol{\varepsilon}_c$ par rapport à \mathbf{x}_c permettent d'éliminer le caractère aléatoire de cette fonction dans le programme d'optimisation de $\mathbf{x}_{b,c}$. Il est alors aisé de montrer que le choix optimal de $\mathbf{x}_{b,c}$ s'écrit sous la forme :

$$(15d) \quad \mathbf{x}_{b,c}^*(\mathbf{w}_c; \boldsymbol{\beta}_{b,c}) = \boldsymbol{\beta}_{b,c} - \mathbf{g}_{b,c}(\mathbf{w}_{b,c}; \mathbf{w}_{a,c}) \text{ avec } \mathbf{w}_{b,c} \equiv \mathbf{w}_b / p_c$$

et où $\mathbf{g}_{b,c}(\cdot; \mathbf{w}_{b,c})$ est la fonction réciproque en $\mathbf{x}_{b,c} - \boldsymbol{\beta}_{b,c}$ de l'opposé du gradient de :

$$(15e) \quad \mathbf{g}_c \left[\begin{array}{c} \mathbf{x}_{b,c} - \boldsymbol{\beta}_{b,c} \\ -\boldsymbol{\gamma}_{a,c}(\mathbf{w}_{a,c}; \mathbf{x}_{b,c} - \boldsymbol{\beta}_{b,c}) \end{array} \right]$$

en $\mathbf{x}_{b,c}$. La simplicité des résultats obtenus est déterminante pour l'utilisation empirique de ces modèles. Elle repose cependant sur des hypothèses assez restrictives. Il est en particulier intéressant de remarquer que ces résultats sont transposables, avec des modifications modestes, au cas où les agriculteurs sont supposés averses face au risque. Ce résultat surprenant tient à trois spécificités de ce modèle. 1) Les termes aléatoires $\boldsymbol{\varepsilon}_c$ sont « purement exogènes » vis-à-vis des quantités d'intrants. 2) L'information obtenue sur $\boldsymbol{\varepsilon}_{a,c}^x$ pour les choix de $\mathbf{x}_{a,c}$ est parfaite, les choix de $\mathbf{x}_{a,c}$ se font donc sans incertitude sauf, éventuellement, celle liée à $\boldsymbol{\varepsilon}_c^y$, les effets de ce terme étant entièrement subis par l'agriculteur. 3) Les choix de $\mathbf{x}_{a,c}$ éliminent entièrement les variations engendrées par $\boldsymbol{\varepsilon}_{a,c}^x$ au moment du choix de $\mathbf{x}_{b,c}$, ce qui est lié au fait que $\boldsymbol{\varepsilon}_{a,c}^x$ et $\mathbf{x}_{a,c}$ sont des substituts parfaits dans la fonction de rendement. De fait, la présence de $\boldsymbol{\varepsilon}_c^y$ modifie les choix d'agriculteurs averses face au risque, mais pas la structure ni l'enchaînement de ces choix.

De fait, la forme de la fonction de rendement (13) donne des conditions suffisantes sur la

forme de la technologie de production aléatoire permettant de dériver des fonctions de demande d'intrants et d'offre de production avec des termes aléatoires additifs.

Avec $\mathbf{g}_{a,c}(\mathbf{w}_c) \equiv \gamma_{a,c}(\mathbf{w}_{a,c}; -\gamma_{b,c}(\mathbf{w}_{b,c}; \mathbf{w}_{a,c}))$, $\mathbf{g}_c(\mathbf{w}_c) \equiv \begin{bmatrix} \mathbf{g}_{b,c}(\mathbf{w}_c) \\ \mathbf{g}_{a,c}(\mathbf{w}_c) \end{bmatrix}$ et $\gamma_c(\mathbf{w}_c) \equiv g_c[-\mathbf{g}_c(\mathbf{w}_c)]$, et

en réintroduisant les arguments (\mathbf{q}_c, s_c) on obtient finalement :

$$(16a) \quad \mathbf{x}_c^*(\mathbf{w}_c; \boldsymbol{\varepsilon}_c^x; \mathbf{e}_c^x, \mathbf{q}_c, s_c) = \mathbf{b}_c(\mathbf{q}_c, s_c) + \mathbf{e}_c^x - \mathbf{g}_c(\mathbf{w}_c; \mathbf{q}_c) + \boldsymbol{\varepsilon}_c^x,$$

$$(16b) \quad y_c^*(\mathbf{w}_c; \boldsymbol{\varepsilon}_c^y; e_c^y, \mathbf{q}_c, s_c) = a_c(\mathbf{q}_c, s_c) + e_c^y - \gamma_c(\mathbf{w}_c; \mathbf{q}_c) + \varepsilon_c^y$$

et :

$$(16c) \quad \pi_c^*(p_c, \mathbf{w}; \boldsymbol{\varepsilon}_c; \mathbf{e}_c, \mathbf{q}_c, s_c) = p_c \left[a_c(\mathbf{q}_c, s_c) + e_c^y - \gamma_c(\mathbf{w}_c; \mathbf{q}_c) + \varepsilon_c^y \right] \\ - \mathbf{w}' \left[\mathbf{b}_c(\mathbf{q}_c, s_c) + \mathbf{e}_c^x - \mathbf{g}_c(\mathbf{w}_c; \mathbf{q}_c) + \boldsymbol{\varepsilon}_c^x \right].$$

Ces fonctions s'avèrent très pratiques à utiliser d'un point de vue empirique, pour plusieurs raisons. Les fonctions de demande d'intrants et d'offre de rendement ont « chacune » leur effet d'hétérogénéité (e) et leur effet aléatoire (ε) même si ces effets sont statistiquement liés d'une fonction à l'autre. Ces termes aléatoires sont additifs. Enfin, en adoptant la normalisation $E[\boldsymbol{\varepsilon}_c] \equiv \mathbf{0}$, il est aisé de déterminer l'espérance de profit (restreint en s_c) par unité de surface de la culture c au moment des choix d'assolement. On a :

$$(16d) \quad \pi_c^e(p_c, \mathbf{w}; \mathbf{e}_c, \mathbf{q}_c, s_c) \equiv E_{\boldsymbol{\varepsilon}_c} \left[\pi_c^*(p_c, \mathbf{w}; \boldsymbol{\varepsilon}_c; \mathbf{e}_c, \mathbf{q}_c, s_c) \right] \\ = p_c \left[a_c(\mathbf{z}_c, s_c) + e_c^y - \gamma_c(\mathbf{w}_c; \mathbf{q}_c) \right] - \mathbf{w}' \left[\mathbf{b}_c(\mathbf{q}_c, s_c) + \mathbf{e}_c^x - \mathbf{g}_c(\mathbf{w}_c; \mathbf{q}_c) \right].$$

Cette espérance est déterminante ici car elle fonde les choix d'assolement des agriculteurs. Sa forme décrit en outre les relations qui existent entre les choix d'intrants et de rendements, notamment en ce qui concerne les termes d'hétérogénéité \mathbf{e} .

1.4 Les déterminants des choix d'assolement

Les choix étudiés jusqu'à présent sont les choix que les agriculteurs opèrent au niveau d'une culture. Il reste maintenant à définir un modèle pour leurs choix d'assolement, *i.e.* à définir formellement les raisons pour lesquelles ils décident de produire plusieurs cultures au cours d'une même campagne.

Cette partie comporte deux sections. La première présente, à travers une brève revue de la littérature, les principaux motifs de diversification des cultures et les principales modélisations associées. Cette discussion vise à présenter les intérêts et limites des modèles de choix d'assolement utilisés par les économistes agricoles et à montrer comment les modèles proposés ici s'insèrent dans cette littérature. La seconde section décrit la structure générale des modèles proposés pour les choix d'assolement. Cette structure repose sur les fonctions de production discutées dans la section précédente et un concept emprunté à la programmation mathématique positive : la fonction de coût implicite de gestion de l'assolement. Cette fonction de coût est présentée à travers les hypothèses qui la sous-tendent et ses limites.

1.4.1 *Choix des assolements et motifs de diversification*

On recense en économie agricole quatre raisons principales qui incitent les agriculteurs à diversifier leur assolement : la gestion du risque (de production et/ou de marché), la gestion des effets des rotations culturales, la qualité du sol et la gestion de l'utilisation des intrants fixes à court terme (main d'œuvre et matériel essentiellement). Le rôle des successions culturales est spécifique ici : la gestion de ces successions impose l'alternance des cultures sur une même parcelle, pas nécessairement la diversification de l'assolement une année donnée. Bien entendu, ces raisons expliquent vraisemblablement toutes les choix des agriculteurs mais à des degrés divers. Les poids relatifs de ces motifs de diversification dépendent de la spécialisation de l'exploitation, du site de production, de l'attitude face au risque du décideur, des quantités d'intrants fixes disponibles, du contexte économique et notamment des politiques

agricoles en vigueur ... L'identification de ces « poids relatifs » est très difficile empiriquement si on n'interroge pas directement les agriculteurs. La spécification d'un modèle intégrant l'ensemble des motifs de diversification cités ci-dessus est peut-être réalisable mais son utilisation empirique est plus que discutable. Outre la complexité des relations théoriques à définir, il est vraisemblable que les paramètres pertinents pour la révélation des « poids » des motifs soient quasiment impossibles à identifier empiriquement. Dans ce contexte, les économistes agricoles, qu'ils utilisent les techniques de la programmation mathématique ou celles de l'économétrie, se focalisent généralement sur un motif de diversification, voire deux.

Chavas et Holt (1990) ont été les premiers à proposer un modèle économétrique décrivant les choix d'assolement comme le choix d'un « portefeuille » de cultures. Les travaux initiaux de Chavas et Holt (1990) ont été poursuivis par, *e.g.*, Coyle (1992, 1999), Moro et Sckokai (2006) et Hennessy (2004, 2006). Un agriculteur averse face au risque aura tendance à choisir un assolement diversifié pour répartir les risques affectant son revenu. La gestion du risque de revenu est également considérée comme un motif de diversification des assolements dans certains modèles de programmation mathématique. La diversification des assolements n'est une stratégie efficace que si les risques de prix ou de production des différentes cultures sont indépendants ou, mieux encore, anti-corrélés. En outre, elle n'intéresse que les agriculteurs avec une aversion face au risque suffisamment importante.

Peu d'études économétriques ont abordé la question des effets des rotations, en termes d'amélioration du rendement et de diminution des quantités d'intrants à appliquer, sans doute parce que les solutions analytiques sont difficiles à obtenir dans les modèles d'optimisation dynamique (Eckstein, 1984 ; Ozarem et Miranowski, 1994). Cependant, des études plus récentes mais assez diverses tendent à montrer que ces questions suscitent un intérêt grandissant en économie agricole (Thomas, 2003 ; Arnberg et Hansen, 2006 ; Hennessy, 2006 ; Livingston, Roberts et Rust, 2008).

Les modèles de programmation mathématique intègrent ce qu'on nomme généralement des « contraintes agronomiques ». Certaines de ces contraintes permettent de rendre compte de ce que certaines successions culturales sont impossibles ou fortement déconseillées aux agriculteurs. Des successions culturales sont impossibles si les calendriers culturaux des cultures concernées se chevauchent, ce qui arrive si la récolte de la première culture a lieu après les semis de la seconde. Des successions culturales sont déconseillées pour deux raisons. La première est que les calendriers culturaux des cultures concernées ont une forte probabilité de se chevaucher en raison des aléas climatiques pouvant retarder la récolte de la première et/ou repousser à une date trop tardive les semis de la seconde. La seconde est liée aux problèmes phytosanitaires engendrés par certaines successions, notamment la monoculture d'une espèce, ou plus rarement à des questions liées à la structure du sol (notamment en production légumière). Par exemple, cultiver du colza deux années de suite sur la même parcelle est fortement déconseillé : le rendement du colza de la seconde année est généralement médiocre pour des raisons phytosanitaires. Le délai de retour sur une parcelle, conseillé pour le colza, est généralement de trois ans. Ceci permet de fixer une contrainte simple dans les modèles de programmation mathématique : la part du colza dans l'assolement ne peut dépasser $1/3$. Bien entendu, cette contrainte est discutable car elle suppose implicitement que le choix décrit est dans un régime stationnaire.

Les contraintes « agronomiques » utilisées dans les modèles de programmation mathématique sont donc de deux types. Les contraintes associées à des rotations impossibles imposent de réelles limites aux choix d'assolement. Mais les successions culturales déconseillées n'imposent pas de réelles contraintes. Les conseils sont liés au fait que, dans les gammes habituelle des prix des cultures, ces successions culturales ont des valeurs d'opportunités négatives : elles sont dominées en terme de rentabilité par d'autres successions culturales. Elles peuvent implicitement être considérées comme des contraintes dès lors que les niveaux de prix qui les rendraient rentables sont invraisemblables. Les successions colza/colza sont

dans ce cas. *A contrario* les successions blé/blé mettent en jeu un réel arbitrage vis-à-vis d'autres successions. Les pertes de rendement du second blé demeurent limitées dès lors que sa protection contre certaines maladies est renforcée. Le second blé est donc moins rentable que le premier mais peut être plus rentable qu'une autre culture si le prix du blé est élevé. Les successions blé/blé sont régulièrement pratiquées par les agriculteurs français. La quasi-monoculture de blé est une pratique très fréquente au Royaume-Uni.

L'hétérogénéité de la qualité des terres de l'exploitation est parfois avancée comme un motif de diversification des cultures (Howitt 1995 ; Moore et Negri, 1992 ; Guyomard *et al*, 1996). Cette hétérogénéité peut être liée à la nature des sols ou à la topographie des parcelles. L'hétérogénéité des conditions de production (pédologiques, topographiques et climatiques) est importante à l'échelle d'un pays comme la France et explique une large part des différences dans les assolements observés d'une région à une autre. L'hétérogénéité des conditions de production sur une exploitation est généralement assez limitée et ne peut justifier à elle seule la diversification des assolements, même si elle peut y contribuer.

De fait, l'hétérogénéité des conditions de production sur une exploitation est un argument utilisé par les économistes agricoles utilisant la programmation mathématique (Howitt, 1995) ou la modélisation économétrique (Moore et Negri, 1992 ; Guyomard *et al*, 1996) pour justifier les hypothèses de rendements marginaux de la terre décroissants. Cette hypothèse est à la base de la construction des modèles économétriques de choix d'assolement présentés dans la partie 1. Elle traduit l'idée selon laquelle le profit associé à une culture diminue en fonction de la sole qui lui est consacrée. Selon les agronomes que nous avons consultés, cet argument est discutable dans le cas des grandes cultures, cas considéré ici. Si une parcelle est « bonne » pour une culture, elle l'est généralement pour les autres. L'hétérogénéité de la qualité « absolue » des parcelles ne peut donc justifier la diversification des assolements au niveau des exploitations, sauf dans des cas extrêmes.

Cependant, et c'est l'autre argument avancé en faveur de cette hypothèse par les économistes agricoles, les rendements marginaux de la terre sont généralement décroissants pour une culture en raison des effets des rotations culturales. Si l'assolement de l'année précédente est diversifié, le schéma de rotation habituel doit également conduire à un assolement diversifié pour l'année à venir. Choisir d'accroître la surface d'une culture contraint l'agriculteur à implanter cette culture sur des parcelles avec des effets de précédents moins bénéfiques, voire négatifs. Il est utile de remarquer que cet argument n'est pas « purement » technique, il repose également sur l'hypothèse que les agriculteurs implantent en priorité leurs cultures sur les parcelles les plus appropriées. Dans le cas des effets des successions culturales, l'hétérogénéité de « la qualité des parcelles » est liée à la diversification de l'assolement précédent et les effets de cette hétérogénéité dépendent des espèces cultivées (selon les couples précédents/culture à venir). La qualité des parcelles est donc un concept relatif dans ce cas.

Le dernier motif de diversification des assolements est lié à l'existence d'autres facteurs fixes allouables que la terre. Le travail et le matériel non spécifique à une culture (tracteurs, pulvérisateurs, ...) sont en quantité limitée sur une exploitation. Les calendriers culturaux des plantes cultivées sur une exploitation peuvent donc être à l'origine d'arbitrages visant à éviter des pics trop importants de la charge de travail ou la sur-utilisation de certaines machines, voire éviter que certaines opérations soient à sacrifier au profit d'autres. Les agriculteurs ont deux solutions : choisir des assolements qui tiennent compte de l'étalement des charges de travail ou recourir à des travaux par tiers. Même si la seconde option semble se développer, la première est celle généralement retenue.

Dans les modèles des économistes utilisant les techniques de la programmation mathématique, la question de l'allocation du travail et du matériel est introduite selon deux approches. Les effets de l'organisation du travail et de l'utilisation des machines peuvent être

représentés par une série de contraintes sur les choix d'assolement (Heckelei, 1997) ou être représentés par une fonction de coût implicite, selon l'approche dite de la « programmation mathématique positive » (Howitt, 1995). Ces deux types de représentations sont en fait assez proches, leurs différences sont similaires à celles décrites lors de la comparaison des « contraintes de rotations » et des « effets de rotation ». La saturation des contraintes sur les choix d'assolement génère des coûts implicites mesurés par les multiplicateurs de Lagrange associés à ces contraintes. La fonction de coût implicite de la programmation mathématique positive représente implicitement directement ces coûts qui sont ceux associés aux surcharges de travail, de la main d'œuvre ou des machines (Howitt, 1995).

Dans les modèles spécifiés par les économistes utilisant des approches économétriques, les questions liés à l'allocation du travail et du matériel sont généralement gérées implicitement, à travers les notions de jointures technologiques. Les seuls auteurs abordant ce problème de manière explicite sont Arnberg et Hansen (2007). Ils construisent leurs modèles sur la base d'une fonction de coût implicite représentant « les effets de l'organisation des charges de travail » sur les choix d'assolement, mais leur fonction de coût diffère sensiblement de celle de la programmation mathématique positive parce qu'ils imposent certaines hypothèses de séparabilité. Leur approche n'est pas développée car elle est très proche, dans l'esprit, de celle que nous proposons dans cet article.

1.4.2 Rendements marginaux de la terre décroissants et coûts implicites des assolements

La discussion précédente sur les motifs de diversification des assolements met en évidence trois aspects fondamentaux liés aux choix d'assolement des agriculteurs. Tout d'abord, tout modèle de choix ne peut qu'être très approximatif étant donné la multiplicité des déterminants en jeu. Se focaliser sur un aspect de ces choix et exclure totalement les autres du modèle spécifié apparaît également discutable. Une solution « acceptable » semble donc consister en une solution permettant d'intégrer au mieux l'ensemble des motifs de diversification. Dans les modèles proposés, la question de la gestion du risque est cependant

exclue. La répartition des risques de production et de marché repose sur des considérations spécifiques liées à l'attitude face au risque des agriculteurs. L'intégration de ces considérations aux modèles présentés ici apparaît difficile car elle requiert l'utilisation d'un cadre d'analyse beaucoup plus « lourd » d'utilisation. En outre, aborder la question de la gestion du risque n'a d'intérêt que si les principaux aspects « techniques » des choix d'assolement sont considérés de manière satisfaisante.

Ensuite, la question de la gestion des successions culturales joue un rôle omniprésent dans les discussions des économistes et des agronomes sur les choix d'assolement des agriculteurs. En vertu de l'application du principe de rationalité de la micro-économie, une analyse du choix des assolements intégrant les effets des rotations culturales nécessite la spécification des effets dynamiques des successions culturales et l'utilisation de critères de choix inter-temporels. Même si quelques progrès ont récemment été réalisés dans cette direction, cette approche demeure très difficile, et ce d'autant plus que les données nécessaires, ou tout au moins facilitant la mise en œuvre, à de telles analyses (les successions culturales) ne sont quasiment jamais disponibles. Une solution à ce problème est nécessairement approximative. Pour être satisfaisante elle doit s'appuyer sur des hypothèses raisonnables (et explicites). Les modèles présentés ici s'appuient sur des hypothèses issues de nos discussions avec les agronomes que nous consultons. La principale d'entre elles est issue d'un fait stylisé : les schémas de successions culturales utilisés par les agriculteurs sont relativement stables et s'appuient sur des « règles agronomiques » bien établies, tout au moins à court terme. Cette hypothèse facilite la modélisation des choix de court terme des agriculteurs. Les choix d'assolement des agriculteurs sont certes flexibles mais dans certaines limites, celles imposées par les schémas de rotation qu'ils utilisent et les quantités de travail et le matériel à leur disposition. En tenir compte, au moins de manière implicite, permet de raisonner dans un cadre essentiellement statique. Dans ce cadre d'analyse les arbitrages intertemporels liés aux effets des successions culturales fixent les « limites » des choix d'assolement. Cette hypothèse est en fait

explicitement utilisée par Arnberg et Hansen (2007) et fonde plus ou moins implicitement de nombreux modèles de programmation mathématique « standards » (Heckeleï 1997) ou de programmation mathématique positive (Howitt 1995 ; Heckeleï et Wolff 2003). Ceci dit, et c'est là un point essentiel pour l'analyse empirique des choix de production des agriculteurs, si les schémas de rotation utilisés par les agriculteurs sont stables au niveau d'une exploitation, ils sont également hétérogènes au sein de la population des agriculteurs et ont pu évoluer dans le temps. Par exemple, si la sole de certaines cultures telles que le pois a récemment diminué en France en raisons de problèmes sanitaires, la sole d'autres cultures telles que le colza s'est développé grâce à des progrès techniques qui les ont rendues plus rentables.

Enfin, la question des intrants fixes allouables autres que la terre pose des problèmes spécifiques, notamment au niveau de la spécification des fonctions de production des cultures. Dans les modèles présentés dans la partie 1, la disponibilité des facteurs travail et capital est susceptible de jouer un rôle sur les choix d'assolement selon deux effets. Dans les modèles considérant que la technologie de production agricole est non-jointe dans les assolements, le seul effet présent est lié à l'utilisation des intrants variables. Schématiquement cet effet peut être décrit comme suit : la disponibilité des facteurs fixes peut modifier les choix d'intrants variables au niveau de chaque culture parce que ces intrants fixes sont nécessaires à l'application des intrants variables. De ce fait, les quantités d'intrants fixes ont un effet indirect sur les rentabilités des cultures, *via* les choix d'intrants variables, et donc sur les choix d'assolement. Cet effet est généralement exclu des modèles de programmation mathématique qui reposent sur des marges brutes « fixes », et donc sur des choix d'intrants variables invariants dans les assolements. Dans les modèles considérant que la technologie de production agricole est jointe dans les assolements, à l'effet précédent s'ajoute un effet qui est lui spécifique aux choix d'assolement eux-mêmes. Ces effets sont liés à la fonction $H(\bar{Z}, S)$ de l'équation (10). Cette fonction est en fait à rapprocher des contraintes sur les choix d'assolement et de la fonction de coût implicite de l'assolement utilisées dans les modèles de

programmation mathématique. Nous proposons ici une formalisation alternative de cette idée.

Le premier effet présenté ici implique que les agriculteurs sont susceptibles de modifier leurs pratiques culturales du fait des contraintes imposées par la disponibilité des facteurs fixes. Cet effet laisse septique la plupart des agronomes que nous avons consulté. Les pratiques culturales sont des ensembles cohérents de techniques de production qui peuvent certes être modifiées, mais relativement marginalement pour ne pas remettre en cause leur cohérence. Il est de fait moins coûteux pour un agriculteur d'adapter son assolement à sa dotation en facteurs fixes plutôt que de modifier certaines de ses pratiques culturales. Cette hypothèse est implicite dans la plupart des modèles de programmation mathématique. Elle est imposée et formalisée ici. L'analyse de cette question développée ici prolonge celle de Arnberg et Hansen (2007).

Les hypothèses décrites ci-dessus donnent les fondements théoriques des modèles proposés ici. Ces modèles sont dérivés d'une fonction de profit espéré de la forme suivante :

$$(17) \quad \bar{\pi}(\mathbf{p}, \mathbf{w}; \mathbf{s}; \mathbf{q}, \bar{\mathbf{z}}) \equiv \sum_{c=1}^C s_c \pi_c^e(p_c, \mathbf{w}; \mathbf{q}_c, s_c) - c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}}).$$

Les effets aléatoires $(\mathbf{e}_c, \varepsilon_c)$ sont omis ici pour simplifier les notations. Le profit est rapporté à la surface de l'exploitation, c'est donc une mesure de profit par unité de surface. Le vecteur \mathbf{q} est un vecteur comprenant les \mathbf{q}_c pour $c=1, \dots, C$ ainsi que, éventuellement, d'autres variables de contrôle définies à l'échelle de l'exploitation. Ce vecteur inclut la surface totale de l'exploitation, ce qui peut permettre de représenter des effets d'échelle. Le vecteur $\bar{\mathbf{z}}$ contient les quantités de facteurs fixes autres que la terre définies par unité de surface.

Cette fonction de profit est composée de deux parties : la somme pondérée par les parts d'assolement des profits par cultures (marges brutes), $\sum_{c=1}^C s_c \pi_c^e(p_c, \mathbf{w}; \mathbf{q}_c, s_c)$, et une fonction (de coût) $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ qui dépend de l'assolement \mathbf{s} et des quantités des intrants fixes $\bar{\mathbf{z}}$. D'un point de vue strictement « formel », cette fonction de profit contient deux motifs de diversification : les effets des s_c dans les fonctions de profit espéré des cultures $\pi_c^e(p_c, \mathbf{w}; \mathbf{q}_c, s_c)$ et les effets de \mathbf{s} dans la fonction $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$.

La fonction $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ est certainement ce qui distingue notre approche de celles utilisées habituellement pour la modélisation économétrique des choix d'assolement. Elle est ici désignée sous le terme de « fonction de coût implicite de gestion l'assolement ». Elle est supposée positive, strictement quasi-convexe et croissante en \mathbf{s} , et deux fois continûment différentiable en \mathbf{s} . Son rôle est de représenter, de manière approximative et selon des effets « lisses » plusieurs déterminants des choix d'assolements.

1) Elle traduit, en termes monétaires, les limites (à court terme) imposées sur les choix d'assolement par les schémas de rotation utilisés par les agriculteurs. De ce fait elle traduit également, en termes monétaires, l'effet des contraintes « agronomiques » utilisées dans les modèles de programmation mathématique. Les formes fonctionnelles présentées dans la partie suivante peuvent être normalisées, ou re-paramétrées, de telle sorte qu'il existe un assolement dit « de référence », $\mathbf{s}^r(\mathbf{q}, \bar{\mathbf{z}})$ tel que $\sum_{c=1}^C s_c^r(\mathbf{q}, \bar{\mathbf{z}}) = 1$, qui vérifie les conditions suivantes. Il existe une fonction $c^r(.)$ telle que :

$$18a) \quad c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}}) = c^r(\mathbf{s} - \mathbf{s}^r(\mathbf{q}, \bar{\mathbf{z}}); \mathbf{q}, \bar{\mathbf{z}})$$

et :

$$18b) \quad c^r(\mathbf{0}; \mathbf{q}, \bar{\mathbf{z}}) \leq c^r(\mathbf{s} - \mathbf{s}^r(\mathbf{q}, \bar{\mathbf{z}}); \mathbf{q}, \bar{\mathbf{z}})$$

pour tout \mathbf{s} tel que $\sum_{c=1}^C s_c = 1$. L'assolement $\mathbf{s}^r(\mathbf{q}, \bar{\mathbf{z}})$ est donc celui qui engendre les coûts de gestion minimum. La quasi-convexité de $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ (qui implique celle de $c^r(\mathbf{s} - \mathbf{s}^r(\mathbf{q}, \bar{\mathbf{z}}); \mathbf{q}, \bar{\mathbf{z}})$) entraîne que tout assolement \mathbf{s} qui s'écarte de $\mathbf{s}^r(\mathbf{q}, \bar{\mathbf{z}})$ accroît les coûts de gestion de l'assolement selon la métrique définie par $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}}) = c^r(\mathbf{s} - \mathbf{s}^r(\mathbf{q}, \bar{\mathbf{z}}); \mathbf{q}, \bar{\mathbf{z}})$. Cette fonction traduit donc l'idée que, au moins à court terme, les schémas de rotation utilisés par les agriculteurs exercent « une force de rappel » qui tend à rigidifier les choix d'assolement. Cette idée est analogue à celle qui fonde les coûts d'ajustement partiel des modèles dynamiques. Des mécanismes de ce type sont à l'œuvre pour décrire les seconds effets représentés par la fonction $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$.

2) Cette fonction de coût représente également les effets liés aux limites dans les choix d'assolement imposées par la disponibilité des facteurs fixes autres que la terre. Ceci explique pourquoi la forme de cette fonction dépend de la dotation en facteurs fixes de l'exploitation, *i.e.* de $\bar{\mathbf{z}}$. Les quantités limitées de main d'œuvre ou de matériel tendent à engendrer les mêmes effets que le respect de schémas de rotation. Dans ce cas, l'assolement de référence $\mathbf{s}^r(\mathbf{q}, \bar{\mathbf{z}})$ est celui qui correspond aux coûts de gestion minimaux étant donnée la dotation en facteurs fixes de l'exploitation. De fait, les schémas de rotations et les quantités de facteurs fixes sont des choix de long terme. Leur cohérence assure implicitement l'existence de l'assolement de référence $\mathbf{s}^r(\mathbf{q}, \bar{\mathbf{z}})$. Plus une exploitation est dotée en facteurs fixes, plus ses choix d'assolement sont flexibles ce qui doit se traduire formellement par le fait que $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ est d'autant plus « plate » en \mathbf{s} autour de $\mathbf{s}^r(\mathbf{q}, \bar{\mathbf{z}})$ et faible par rapport à $\sum_{c=1}^C s_c \pi_c^e(p_c, \mathbf{w}; \mathbf{q}_c, s_c)$.

3) Cette fonction de coût est implicite dans le sens où elle mesure des coûts non comptables comme les coûts des surcharges de travail de l'exploitant ou les effets de l'usure prématurée du matériel utilisé. En outre, elle représente les effets de contraintes qui n'ont pas réellement

de coût même si elle convertit l'effet de ces contraintes en unités monétaires. Enfin, cette fonction de coût comprend les coûts comptables des intrants variables non considérés dans les fonctions de production, notamment les charges de carburant.

Une fonction de coût telle que $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ a déjà été utilisée sans réelle justification autre que son rôle « technique » dans la dérivation des équations d'assolement par Oude Lansink et Stefanou (2001) ou Heckelei et Wolff (2003). La fonction $H(\bar{\mathbf{Z}}, \mathbf{S})$ de l'équation (10) du modèle « avec jointures technologiques » de Chambers et Just (1989) peut également être interprétée comme une contre-partie de $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$. Cependant, l'argument de l'utilisation de la fonction $H(\bar{\mathbf{Z}}, \mathbf{S})$ par Chambers et Just (1989) est essentiellement « technique » : cette fonction introduit de la « flexibilité » dans la forme de la technologie multi-produits spécifiée et donc dans les modèles de choix d'assolement qui en dérivent. De ce point de vue, la fonction $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ représente les jointures technologiques par les assolements dans les modèles proposés ici.

Arnberg et Hansen (2007) utilisent une fonction de cette forme, essentiellement pour représenter les effets des contraintes imposées par les facteurs fixes allouables. La même justification est fournie pour l'utilisation d'une fonction similaire à $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ dans la programmation mathématique positive (Howitt 1995). Cependant, la fonction de coût dans cette approche sert essentiellement à calibrer les modèles empiriques de programmation mathématique. Elle intègre généralement les coûts d'opportunité des choix d'assolement liés aux contraintes imposées sur ces choix. Ces coûts d'opportunité dépendent des prix des cultures et des intrants variables (\mathbf{p}, \mathbf{w}) , c'est d'ailleurs ce qui fonde la critique de Britz, Heckelei et Wolff (2003) et de certaines extensions de la programmation mathématique positive, notamment celle de Paris et Howitt (1998). La fonction de coût considérée est « indépendante » des choix opérés au niveau des cultures. Cette remarque renvoie à une autre hypothèse qui fonde la structure de la fonction de profit $\bar{\pi}(\mathbf{p}, \mathbf{w}; \mathbf{s}, \mathbf{q}, \bar{\mathbf{z}})$.

Dans la fonction de profit $\bar{\pi}(\mathbf{p}, \mathbf{w}; \mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$, la fonction de coût $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ ne dépend pas de (\mathbf{p}, \mathbf{w}) et les fonctions de profit espéré (de court terme) des cultures, les $\pi_c^e(p_c, \mathbf{w}; \mathbf{q}_c, s_c)$, ne dépendent pas de la dotation en facteurs fixes $\bar{\mathbf{z}}$. Cette caractéristique repose une hypothèse de séparabilité de la technologie de production entre les facteurs fixes et les intrants variables alloués aux cultures. Cette hypothèse impose que a) les quantités d'intrants variables utilisées pour les cultures ne dépendent pas des quantités de facteurs fixes disponibles et, par conséquent, b) l'allocation des facteurs fixes autres que la terre ne dépend que de l'assolement. Elle est également imposée par Arnberg et Hansen (2007) et dans la plupart des modèles de programmation mathématique mais elle n'est pas imposée dans les modèles économétriques présentés dans la section 1. Cette hypothèse ne peut être valide qu'en tant qu'approximation et à court terme. Elle repose sur l'idée que les agriculteurs utilisent des pratiques culturales relativement rigides et sont plus enclins à modifier leurs assolements que leurs pratiques culturales. Cette idée est considérée comme un fait stylisé ici.

Les ajustements d'utilisation d'intrants variables au niveau des cultures (en fonction des conditions de marché ou des effets des rotations) dans le cadre d'une pratique culturale donnée ne requièrent pas d'ajustement des quantités de facteurs fixes. L'hypothèse de séparabilité imposée ici est donc supposée valide « localement », *i.e.* pour des niveaux d'utilisations d'intrants variables proches de ceux observés. De manière plus fondamentale, les fonctions de rendement des cultures présentées dans les équations (14) sont définies comme des représentations formelles décrivant les effets des pratiques culturales utilisées. De fait, les relations existant entre y_c d'une part et \mathbf{x}_c d'autre part dépendent essentiellement des pratiques culturales utilisées, ce que résumant les fonctions $a_c(\cdot)$, $\mathbf{b}_c(\cdot)$ et $g_c(\cdot)$ et les termes d'hétérogénéité \mathbf{e}_c . Des changements importants dans les niveaux d'utilisations d'intrants \mathbf{x}_c ne peuvent se concevoir que dans une logique de changement de pratiques culturales, *i.e.* du passage d'un ensemble cohérent de techniques de production à un autre ensemble de techniques différent mais tout aussi cohérent. Aussi, des changements importants des niveaux

d'utilisations d'intrants \mathbf{x}_c impliquent un changement des fonctions de production sous-tendant ces choix. De ce point de vue, il est inutile de rendre flexibles des fonctions de production qui ne sont en fait valides que pour des gammes d'utilisation d'intrants limitées.

Ces remarques définissent les limites des modèles proposés ici. Leur domaine de validité est limité à un voisinage des niveaux d'utilisation d'intrants variables (et donc de rendements) et des choix d'assolement observés. Ils doivent donc être interprétés comme des « approximations » locales, *i.e.* des approximations valides autour des $(\mathbf{y}, \mathbf{x}, \mathbf{s})$ observés.

La dernière caractéristique importante de la fonction de profit $\bar{\pi}(\mathbf{p}, \mathbf{w}; \mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ réside dans la spécification des effets des s_c dans les $\pi_c^e(p_c, \mathbf{w}; \mathbf{q}_c, s_c)$. Ces effets déterminent entièrement les choix d'assolement de la plupart des modèles économétriques tels que les modèles « sans jointures technologiques » présentés dans la partie 1. De ce point de vue, ces modèles sont des cas particuliers des modèles proposés ici. Ces effets visent à approximer les effets des successions culturales, voire de l'hétérogénéité de la terre à l'échelle de l'exploitation, sur les profits des cultures. Ces effets sont supposés être négatifs (voire nuls) :

$$(19a) \quad \frac{\partial \pi_c^e(p_c, \mathbf{w}; \mathbf{q}_c, s_c)}{\partial s_c} \leq 0$$

pour $c = 1, \dots, C$. Ils traduisent donc l'idée des rendements marginaux de la terre décroissants, potentiellement au niveau de chaque culture. Cependant, ces effets concernent principalement les cultures dont la part dans l'assolement est importante. La décroissance de $\pi_c^e(p_c, \mathbf{w}; \mathbf{q}_c, s_c)$ en s_c provient en fait de deux types d'effets. Le premier type d'effets regroupe ceux liés à la structure du sol et à des problèmes sanitaires sans solutions chimiques pour les agriculteurs. Ces effets concernent directement le potentiel de rendement de la plante, ce qui se traduit dans les fonctions de rendement (14) par :

$$(19b) \quad \frac{\partial a_c(\mathbf{q}_c, s_c)}{\partial s_c} \leq 0.$$

Le second type d'effets regroupe ceux liés aux utilisations d'intrants. Les schémas de rotation recommandés visent à exploiter les effets de la culture précédente pour réduire les utilisations d'intrants. S'écarter de ces schémas tend à accroître les quantités d'intrants nécessaires au développement et à la protection de la culture, ce qui se traduit dans les fonctions de rendement (14) par :

$$(19c) \quad \frac{\partial \mathbf{b}_c(\mathbf{q}_c, s_c)}{\partial s_c} \geq \mathbf{0}.$$

Bien entendu, les effets représentés dans les équations (19b) et (19c) sont approximatifs. Dans les formes fonctionnelles généralement utilisées (les nôtres n'échappent pas à cette règle), ces effets sont représentés de manière continue alors qu'ils sont en fait définis par paliers, ces paliers dépendant de l'assolement précédent et des successions pratiquées.

Ces effets de \mathbf{s} dans les marges brutes des cultures complètent ceux de \mathbf{s} dans la fonction de coût implicite de gestion de l'assolement : ils pénalisent les écarts de l'assolement choisi par rapport au schéma de rotation employé par l'agriculteur. Ces effets mesurent les coûts d'opportunité de \mathbf{s} en termes de marges brutes espérées.

Les choix d'assolement optimaux peuvent être dérivés du modèle précédent par résolution du programme de maximisation de $\bar{\pi}(\mathbf{p}, \mathbf{w}; \mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ sous la contrainte de surface totale :

$$(20) \quad \mathbf{s}^*(\mathbf{p}, \mathbf{w}; \mathbf{s}; \mathbf{q}, \bar{\mathbf{z}}) \equiv \text{Arg max}_{\mathbf{s}} \left\{ \left[\sum_{c=1}^C s_c \pi_c^e(p_c, \mathbf{w}; \mathbf{q}_c, s_c) - c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}}) \right] \text{ s.c. } \sum_{c=1}^C s_c = 1 \right\}.$$

Il est ici supposé que la contrainte de surface totale est saturée et que $\mathbf{s}^*(\mathbf{p}, \mathbf{w}; \mathbf{s}; \mathbf{q}, \bar{\mathbf{z}}) > \mathbf{0}$. La résolution de ce programme est considérée dans la partie suivante.

Deux remarques s'imposent ici. Tout d'abord, il est tentant d'introduire les choix d'assolements des années précédentes (lorsque ces variables sont mesurées) dans le vecteur des variables de contrôle \mathbf{q} du modèle présenté ici. La prise en compte de ces effets de rotation pose un problème de cohérence vis-à-vis du cadre d'analyse des choix des agriculteurs considéré ici. Ce dernier est essentiellement statique. Il repose sur une hypothèse de stabilité des schémas de rotation utilisés par les agriculteurs. La fonction $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ et les effets de s_c introduits dans les fonctions de marge brute espérée $\pi_c^e(p_c, \mathbf{w}; \mathbf{q}_c, s_c)$ approximent déjà les arbitrages inter-temporels en jeu dans la gestion des rotations culturales.

Introduire les assolements précédents (voire des estimations des successions culturales) dans le modèle a deux implications. Tout d'abord, cela revient à relâcher l'hypothèse des schémas de rotation stables et donc à éliminer les contraintes liées au respect de ces schémas de rotation de la fonction $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$. Cette dernière ne contient alors plus que les effets liés aux facteurs fixes allouables et aux coûts variables non mesurés. Ensuite, considérer que les choix passés ont un impact sur les résultats présents impose d'entrer dans une logique d'optimisation inter-temporelle. En effet, si on considère que l'agriculteur sait que ses choix passés influencent ses revenus d'aujourd'hui, il est raisonnable de considérer également qu'il anticipera les effets de ses choix d'aujourd'hui sur ses revenus futurs. Ceci suppose une adaptation du modèle qui n'est pas triviale. Ce point est discuté dans l'Annexe A.

Ensuite, la fonction de coût implicite de gestion de l'assolement $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ joue un rôle prépondérant dans la structure des modèles de choix d'assolement proposés ici. Comme son nom l'indique, elle décrit des effets non observés. Elle doit donc être considérée comme la forme réduite résumant approximativement les effets de différentes contraintes et de coûts non mesurés sur les choix d'assolement des agriculteurs. Si la spécification des fonctions de profit des cultures relève d'une approche structurelle, celle de la fonction $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ relève de considérations beaucoup plus pragmatiques. Dans ce contexte, la re-paramétrisation de cette

fonction présentée dans les équations (18) s'avère importante : elle permet d'interpréter le « contenu » de la fonction $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ et donc de juger, au moins partiellement, les résultats empiriques obtenus.

1.5 Modèles micro-économétriques : formes fonctionnelles simples et mise en œuvre

Cette partie est essentiellement technique, elle présente des choix de forme fonctionnelle pour les fonctions de production et la fonction de coût implicite $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ qui permettent de déterminer des modèles de choix des agriculteurs sous forme explicite. Elle aborde également quelques difficultés liées à l'estimation de ces modèles. Enfin cette partie discute les principales généralisations des modèles présentés ici.

Cette partie s'articule en quatre sections. Les trois premières présentent des formes fonctionnelles spécifiques pour les fonctions de rendement ou les modèles de choix d'assolement. Chaque section présente les principales propriétés et défauts de ces formes, les fonctions qui lui sont associées et, s'il y a lieu, les généralisations proposées ou envisagées de ces fonctions. La première section présente la fonction de rendement « quadratique translatée ». Elle présente également les concepts duaux qui en dérivent : fonctions de demande d'intrants variables, fonction de d'offre (espérée) et fonction de profit indirecte. La seconde section présente les modèles de choix d'assolement de forme Multinomial Logit (MNL) et la fonction de coût implicite qui lui est associée. La troisième section présente les fonctions d'assolement associées à une fonction de coût de forme quadratique. La quatrième section présente les modèles économétriques de choix de production utilisant les formes fonctionnelles présentées ici. Elle met en évidence quelques difficultés associées à leur utilisation empirique, la principale étant l'absence de données sur les allocations des intrants variables au niveau des cultures.

1.5.1 La fonction de rendement « quadratique translatée »

Aucune forme fonctionnelle de fonction de production usuelle n'a la forme générale de la fonction de rendement définie par les équations (14). La plupart des fonctions de production utilisées empiriquement expriment les ln des niveaux de production en fonction des ln des quantités d'intrants. Les transformations logarithmiques ont l'avantage d'atténuer les variations observées des données et donc d'obtenir de meilleurs critères d'ajustement. Elles permettent également de travailler facilement avec des modèles à effets multiplicatifs. Les idées à l'origine de la forme de la fonction de rendement (14) peuvent être adaptées avec des effets multiplicatifs. Cette solution n'est pas adoptée ici pour deux raisons. Tout d'abord, les notions de rendement maximum et de niveaux d'intrants requis pour ce rendement maximum sont très pratiques d'utilisation. La fonction de profit est additive dans les marges brutes espérées, il est donc préférable d'utiliser des fonctions de rendement en niveau pour obtenir des fonctions de profit par culture avec une forme fonctionnelle simple. Ensuite, les généralisations des modèles proposés, que ce soit dans l'optique de considérer la dynamique des choix d'assolement ou dans l'optique de considérer les solutions en coin des choix d'assolement, sont beaucoup plus simples avec des fonctions de rendement en niveau. Enfin, les formes en niveau sont d'interprétation très aisée.

La forme fonctionnelle retenue ici est une fonction de rendement « quadratique translatée ». Elle est dérivée de la forme quadratique standard $y = \alpha_0 + \beta_0'x - .5 \times x' \Gamma^{-1} x$ par une simple translation des niveaux des intrants qui correspond en fait à une re-paramétrisation. En effet, avec $\alpha_0 = \alpha - .5 \times \beta' \Gamma^{-1} \beta$ et $\beta_0 = \Gamma^{-1} \beta$ on obtient une fonction de rendement de la forme :

$$(21) \quad y = f(x) = \alpha - .5 \times (x - \beta)' \Gamma^{-1} (x - \beta).$$

Une fonction de cette forme a été utilisée par Pope et Just (2003) dans un autre contexte. Ces auteurs ont utilisé cette re-paramétrisation de la fonction quadratique pour gérer des questions liées aux erreurs de mesure sur les variables explicatives et aux erreurs

d'optimisation des agriculteurs. Comme cela sera vu plus loin, l'avantage de cette forme fonctionnelle est de permettre une introduction « naturelle » des termes aléatoires des modèles.

Cette forme est utilisée ici pour cette raison mais également parce que l'interprétation de ses paramètres est très simple. Pour que la fonction de production soit strictement concave, la matrice Γ doit être définie positive (Γ^{-1} l'est alors également). Ceci assure que $(\mathbf{x} - \boldsymbol{\beta})' \Gamma^{-1} (\mathbf{x} - \boldsymbol{\beta}) > 0$ pour tout $\mathbf{x} \neq \boldsymbol{\beta}$. Aussi, α est le rendement maximum et $\boldsymbol{\beta}$ est le niveau d'intrant nécessaire pour obtenir ce rendement maximum. Ces interprétations ont un intérêt pratique évident. D'une part elles permettent de travailler directement avec des notions de rendement et de niveaux d'intrants. D'autre part, les notions portées par les paramètres α et $\boldsymbol{\beta}$ sont des notions utilisées par les agronomes et les conseillers des agriculteurs. L'utilisation des fonctions de rendement quadratique a également un autre intérêt : ces fonctions peuvent s'interpréter comme des approximations de Taylor au second ordre, en \mathbf{x} et autour de $\boldsymbol{\beta}$ pour la fonction « translaturée », des « véritables » fonctions de rendement.

La littérature économique agricole reconnaît l'importance de l'agronomie et son implication dans le choix de la forme fonctionnelle des fonctions de production et dans les méthodes d'estimations appropriées (Zhengfei et *al.*, 2006). Néanmoins, ces études se focalisent principalement sur la spécification des technologies en introduisant, par exemple, le rôle spécifique des pesticides dans le prolongement des travaux de Lichtenberg et Zilberman (1986). Leur intérêt est sujet à caution (Carpentier et Weaver, 1996). Les modèles proposés ici sont des extensions simples des modèles usuels qui offrent un cadre d'analyse flexible de différents effets liés à la spécificité de la production agricole.

Le principal défaut de $f(\mathbf{x})$ est qu'elle n'est pas nécessairement strictement croissante en \mathbf{x} . Néanmoins, la fonction $f(\mathbf{x})$ est croissante pour \mathbf{x} tel que $\Gamma^{-1} \boldsymbol{\beta} \geq \Gamma^{-1} \mathbf{x}$, ce qui assure en fait qu'elle est croissante dans les gammes d'utilisation des intrants qui assurent un profit

maximum dès lors que ces intrants ne sont pas gratuits. Le prix de la production étant p et le vecteur des prix des intrants étant \mathbf{w} , il est aisé de montrer que les fonctions de demandes d'intrants sont données par :

$$(22a) \quad \mathbf{x}^*(\mathbf{w}p^{-1}) = \boldsymbol{\beta} - \boldsymbol{\Gamma}\mathbf{w}p^{-1}$$

alors que la fonction d'offre de produit est donnée par :

$$(22b) \quad y^*(\mathbf{w}p^{-1}) = \alpha - .5 \times p^{-2} \mathbf{w}' \boldsymbol{\Gamma} \mathbf{w}.$$

L'utilisation optimale d'intrants $\mathbf{x}^*(\mathbf{w}p^{-1})$ est donc égale à la quantité d'intrants requise pour le rendement maximum $\boldsymbol{\beta}$ moins $\boldsymbol{\Gamma}\mathbf{w}p^{-1}$, terme qui tient compte de la « non-gratuité » des intrants. De même, le rendement objectif optimal $y^*(\mathbf{w}p^{-1})$ est égal au rendement maximum α moins $.5 \times p^{-2} \mathbf{w}' \boldsymbol{\Gamma} \mathbf{w}$, terme lié au coût des intrants.

Des équations (22a) et (22b), on déduit alors aisément que la fonction de profit indirecte associée à la culture est donnée par :

$$(22c) \quad \pi^*(p, \mathbf{w}) = py^*(\mathbf{w}p^{-1}) - \mathbf{w}' \mathbf{x}^*(\mathbf{w}p^{-1}) = p\alpha - \mathbf{w}' \boldsymbol{\beta} + .5 \times \mathbf{w}' \boldsymbol{\Gamma} \mathbf{w} p^{-1}.$$

Les fonctions duales associées à la fonction de production quadratique, translatée ou non, ont l'avantage d'être linéaires dans leurs paramètres.

Les fonctions de rendement quadratique translatées peuvent être utilisées empiriquement en tant que modèle de rendement des cultures. Pour $c = 1, \dots, C$, l'équation (14a) précédente devient :

$$(23a) \quad y_c = a_c(\mathbf{q}_c) - \delta_c^y s_c + e_c^y + \boldsymbol{\varepsilon}_c^y \\ - .5 \times \left[\mathbf{x}_c - \mathbf{b}_c(\mathbf{q}_c) - \delta_c^x s_c - \mathbf{e}_c^x - \boldsymbol{\varepsilon}_c^x \right]' \boldsymbol{\Gamma}_c^{-1}(\mathbf{q}_c) \left[\mathbf{x}_c - \mathbf{b}_c(\mathbf{q}_c) - \delta_c^x s_c - \mathbf{e}_c^x - \boldsymbol{\varepsilon}_c^x \right].$$

Les formes des termes a_c , \mathbf{b}_c et Γ_c en fonction des variables de contrôle ne sont pas spécifiées ici. Des formes paramétriques simples (linéaires dans leurs paramètres) sont généralement utilisées. Les effets de la sole de la culture c sont ici supposés linéaires dans le niveau de rendement maximum et les quantités d'intrants associées. Cette forme est imposée ici pour faciliter le calcul des choix d'assolement optimaux. Des paramètres δ_c^y et δ_c^x positifs assurent des rendements marginaux de la terre décroissants.

Les équations de choix d'intrants et de fonction de rendement associés sont données par :

$$(23b) \quad \mathbf{x}_c = \mathbf{b}_c(\mathbf{q}_c) + \delta_c^x s_c - \Gamma_c(\mathbf{q}_c) \mathbf{w}_c + \mathbf{e}_c^x + \boldsymbol{\varepsilon}_c^x$$

et :

$$(23c) \quad y_c = a_c(\mathbf{q}_c) - \delta_c^y s_c - .5 \times \mathbf{w}_c' \Gamma_c(\mathbf{q}_c) \mathbf{w}_c + e_c^y + \varepsilon_c^y.$$

L'avantage spécifique des fonctions translatées tient aux termes $a_c(\mathbf{q}_c) - \delta_c^y s_c + e_c^y + \varepsilon_c^y$ et $\mathbf{b}_c(\mathbf{q}_c) + \delta_c^x s_c + \mathbf{e}_c^x + \boldsymbol{\varepsilon}_c^x$. Ces termes sont à la fois les « paramètres constants » des fonctions d'offre de production et de demande d'intrants et, le niveau de rendement maximum et les quantités d'intrants requises associées. Leurs estimations sont donc faciles à interpréter, par les économistes et les non-économistes. Ces termes peuvent « porter » naturellement différents effets, aléatoires ou non. La spécification précédente peut être enrichie pour tenir compte d'effets de successions culturales, sur les niveaux de rendement ou les niveaux d'intrants.

Dans la plupart des modèles de fonctions de production ou des modèles duaux, ces effets doivent être introduits dans des paramètres représentant des productivités marginales voire des élasticités marginales de production, ce qui est plus délicat. De fait, les études empiriques des choix de production s'intéressent peu à la représentation des aléas de production agricole et pratiquement pas à l'hétérogénéité des conditions de production. Les termes d'erreur des

modèles empiriques sont simplement ajoutés de façon *ad hoc* aux modèles déterministes dérivées dans la partie théorique de l'étude, sans analyse de leur contenu (Just et Pope, 2001). Il sera vu dans la suite que le « contenu » des termes aléatoires des modèles empiriques s'avère crucial pour mener à bien l'analyse économétrique des choix d'assolement.

La fonction de profit espérée (indirecte et restreinte en s_c) est alors donnée par :

$$(23d) \quad \pi_c^e(s_c) = p_c [a_c(\mathbf{q}_c) - \delta_c^y s_c + e_c^y] - \mathbf{w}' [\mathbf{b}_c(\mathbf{q}_c) - \delta_c^x s_c - \mathbf{e}_c^x] + .5 \times \mathbf{w}' \Gamma_c(\mathbf{q}_c) \mathbf{w} p_c^{-1}.$$

$\pi_c^e(s_c)$ représente le profit moyen (pour une sole s_c) de la culture c espéré par l'agriculteur lorsqu'il choisit son assolement. Les termes aléatoires ε_c n'interviennent pas dans le calcul de $\pi_c^e(s_c)$. Ils ne sont pas encore observés au moment des choix d'assolement. Les termes ε_c étant additifs dans les fonctions d'offre et de demande et d'espérance nulle, ils ne jouent aucun rôle dans le calcul de $\pi_c^e(s_c)$. En revanche, l'agriculteur connaît les spécificités de son exploitation, il connaît donc \mathbf{e}_c , et plus généralement la forme de ses fonctions de rendement, ce qui lui permet de calculer les $\pi_c^e(s_c)$. Il sera utile par la suite d'écrire ces fonctions sous les formes alternatives :

$$(24a) \quad \pi_c^e(s_c) = \pi_c^e(0) - (p_c \delta_c^y + \mathbf{w}' \delta_c^x) s_c$$

ou encore :

$$(24b) \quad \pi_c^e(s_c) = \pi_c(0) - (p_c \delta_c^y + \mathbf{w}' \delta_c^x) s_c + (p_c e_c^y - \mathbf{w}' \mathbf{e}_c^x)$$

avec $\pi_c(0) \equiv p_c a_c(\mathbf{q}_c) - \mathbf{w}' \mathbf{b}_c(\mathbf{q}_c) + .5 \times \mathbf{w}' \Gamma_c(\mathbf{q}_c) \mathbf{w} p_c^{-1}$ et $\pi_c^e(0) \equiv \pi_c(0) + (p_c e_c^y - \mathbf{w}' \mathbf{e}_c^x)$.

1.5.2 Modèles de choix d'assolement MNL et fonction de coût « entropique »

Les hypothèses présentées dans la partie 3 permettent de décrire les choix de production des agriculteurs en deux étapes. La première consiste en un calcul de la profitabilité des

différentes cultures, via les « marges brutes espérées », i.e. les $\pi_c^e(s_c)$ pour $c = 1, \dots, C$. La seconde étape utilise les résultats de la première étape pour déterminer le choix d'assolement optimal, celui-ci étant le résultat du programme d'optimisation suivant :

$$(25) \quad \text{Max}_{\mathbf{s}} \sum_{c=1}^C s_c \pi_c^e(s_c) - c(\mathbf{s}) \quad \text{s.c.} \quad \sum_{c=1}^C s_c = 1.$$

Bien entendu, c'est la forme de la fonction $c(\mathbf{s})$ qui détermine la forme des choix d'assolement optimaux. Deux modèles sont proposés ici. Ils découlent de logiques différentes. Pour le modèle présenté dans cette section, les choix d'assolement de forme MNL sont associés aux fonctions de coût de forme entropique. La forme des choix d'assolement étant donnée, il reste à définir la forme de la fonction de coût $c(\mathbf{s})$ permettant d'obtenir ces modèles de choix d'assolement (chapitre 2). Le modèle présenté dans la section suivante est issu d'une logique différente. La forme de la fonction $c(\mathbf{s})$ étant a priori inconnue, une approximation quadratique en \mathbf{s} de cette fonction est un choix de forme fonctionnelle qui s'avère pertinent. Il suffit alors de déterminer la forme des choix d'assolement optimaux en utilisant cette forme fonctionnelle pour $c(\mathbf{s})$.

Il est aisé de montrer, en résolvant le programme (25), que si $c(\mathbf{s})$ a la forme suivante :

$$(26a) \quad c(\mathbf{s}) = A + \sum_{c=1}^C s_c g_c + \rho^{-1} \sum_{c=1}^C s_c \ln s_c$$

et si les rendements marginaux de la terre sont constants pour chaque culture, i.e. $\pi_c^e(s_c) = \pi_c^e$ ou encore $\delta_c^y = 0$ et $\delta_c^x = \mathbf{0}$, pour tout s_c où $c = 1, \dots, C$, alors les choix optimaux d'assolement sont donnés par :

$$(26b) \quad \mathbf{s}^*(\boldsymbol{\pi}^e; \mathbf{g}, \rho) = \frac{\exp[\rho(\pi_c^e - g_c)]}{\sum_{m=1}^C \exp[\rho(\pi_m^e - g_m)]}$$

où $\boldsymbol{\pi}^e \equiv (\pi_1^e, \dots, \pi_C^e)$ et $\mathbf{g} \equiv (g_1, \dots, g_C)$. La forme de $\mathbf{s}^*(\boldsymbol{\pi}^e; \mathbf{g}, \rho)$ est ici appelée forme MNL car elle caractérise les fonctions de probabilité du modèle Logit Multinominal (McFadden 1974) utilisé pour la modélisation de choix discrets. Cette forme a déjà été utilisée pour modéliser des choix d'assolement (Caswell et Zilberman, 1985 ; Lichtenberg, 1989 ; Wu et Segerson, 1995). Ce modèle de choix d'assolement a plusieurs avantages : les parts des surfaces prédites à partir de ce modèle seront comprises entre 0 et 1 ; il est parcimonieux en paramètres et leur utilisation empirique est aisée grâce à leur transformation log-linéaire :

$$(27) \quad \ln s_c^*(\boldsymbol{\pi}^e; \mathbf{g}, \rho) - \ln s_1^*(\boldsymbol{\pi}^e; \mathbf{g}, \rho) = \rho(\pi_c^e - \pi_1^e) - \rho(g_c - g_1) \text{ pour } c = 2, \dots, C.$$

Cependant, le modèle (26b) était jusqu'à présent utilisé comme un modèle d'utilisation pratique, sans justification théorique. Les fondements théoriques de l'utilisation de ces modèles, en particulier la fonction de coût décrite par l'équation (26b) sont définis dans le chapitre 2.

Dans cette fonction de coût, le terme A est un coût fixe qui ne peut être identifié mais qui peut être utilisé pour assurer que A est positif. Les termes g_c mesurent les coûts non mesurés dans les fonctions de marges brutes qui s'ajoutent aux coûts de production des cultures. La contrainte de normalisation $g_1 \equiv 0$ peut être utilisée puisque seules les différences des g_c (par rapport au g_c d'une culture de référence, la culture 1 ici) peuvent être identifiées. Le terme $\sum_{c=1}^C s_c \ln s_c$ est l'opposé de l'entropie de l'assolement \mathbf{s} . Il définit l'intérêt à la diversification. Le paramètre ρ est strictement positif. Il mesure les poids relatifs, dans la fonction de profit multi-produits, des marges « nettes » $\sum_{c=1}^C s_c (\pi_c^e(s_c) - g_c)$ d'une part et du coût de la gestion de l'assolement.

L'assolement de référence $s_c^r \equiv \exp(-\rho g_c) / \left[\sum_{m=1}^C \exp(-\rho g_m) \right]$ pour $c = 1, \dots, C$ permet de

définir une paramétrisation alternative de la fonction de coût décrite dans l'équation (26a) :

$$(28) \quad c(\mathbf{s}) = A^r + \rho^{-1} \sum_{c=1}^C s_c (\ln s_c - \ln s_c^r)$$

qui s'avère d'interprétation très facile (avec $A^r = A - \rho^{-1} \ln \sum_{m=1}^C \exp(-\rho g_m)$). Le terme $\sum_{c=1}^C s_c (\ln s_c - \ln s_c^r)$ est l'opposé de l'entropie croisée (par rapport à $\mathbf{s}^r \equiv (s_1^r, \dots, s_C^r)$) de l'assolement \mathbf{s} . Ce terme est minimal (et nul) si $\mathbf{s} = \mathbf{s}^r$. L'assolement de référence peut donc être défini comme l'assolement à coût de gestion minimal. Si ρ tend vers $+\infty$ alors $\mathbf{s}^*(\boldsymbol{\pi}^e; \mathbf{g}, \rho)$ tend vers un assolement totalement dédié à la culture qui procure la marge « nette » $\pi_c^e(s_c) - g_c$ maximale. Si ρ tend vers 0 alors $\mathbf{s}^*(\boldsymbol{\pi}^e; \mathbf{g}, \rho)$ tend vers \mathbf{s}^r , *i.e.* les coûts implicites associés au choix d'assolement sont tels que les marges brutes n'ont plus aucun rôle.

Le modèle de choix d'assolement MNL a quatre principaux défauts. Le premier est une conséquence de sa simplicité : il est relativement rigide. Ceci peut être illustré par les élasticités de $\mathbf{s}^*(\boldsymbol{\pi}^e; \mathbf{g}, \rho)$ par rapport aux $\boldsymbol{\pi}^e$:

$$(29a) \quad \frac{\partial \ln s_c^*(\boldsymbol{\pi}^e; \mathbf{g}, \rho)}{\partial \ln \pi_c^e} = \pi_c^e \rho [1 - s_c^*(\boldsymbol{\pi}^e; \mathbf{g}, \rho)] \quad \text{pour } c = 1, \dots, C$$

et:

$$(29b) \quad \frac{\partial \ln s_c^*(\boldsymbol{\pi}^e; \mathbf{g}, \rho)}{\partial \ln \pi_m^e} = -\rho \pi_c^e s_c^*(\boldsymbol{\pi}^e; \mathbf{g}, \rho) \quad \text{pour } c, m = 1, \dots, C \text{ et } c \neq m.$$

Les élasticités croisés de $s_c^*(\boldsymbol{\pi}^e; \mathbf{g}, \rho)$ par rapport à π_m^e sont toutes égales. Cette propriété est analogue à celle, bien connue du modèle MNL de choix discrets, dite de « l'indépendance des alternatives non pertinentes ». Dans le chapitre 2, nous proposons une généralisation des modèles de choix d'assolement MNL, les modèles de choix d'assolement MNL emboîtés, qui sont plus flexibles que leur contre-partie MNL « standard ».

Le second défaut de ce modèle est d'imposer des rendements marginaux de la terre constants culture par culture et donc de concentrer les motifs de diversification des choix d'assolement dans la seule fonction de coût.

Le troisième défaut de ce modèle, qui est dans certains cas un avantage, est que, par construction, il ne peut être utilisé pour analyser des choix d'assolement avec solutions en coin.

Enfin, les généralisations dynamiques de ce modèle semblent difficiles. Cependant, le modèle de choix d'assolement MNL peut être dérivé à partir d'une agrégation de choix discrets opérés parcelle par parcelle (Caswell et Zilberman, 1985). Nous montrons dans le chapitre 2 que cette interprétation des modèles de choix d'assolement MNL autorise des généralisations de ce modèle dans un contexte dynamique. Certaines de ces généralisations permettent d'ailleurs de lever l'hypothèse des rendements marginaux de la terre constants.

1.5.3 Modèles de choix d'assolement fondés sur des fonctions de coût quadratiques

Un autre modèle simple de choix d'assolement peut être dérivé du programme (25). Il est plus flexible que le modèle MNL, supporte l'hypothèse des rendements marginaux de la terre décroissants pour les cultures et peut être utilisé pour l'analyse de choix d'assolement avec solutions en coin (chapitre 4). Le généraliser dans un contexte de choix dynamique semble cependant difficile.

Puisque la forme de la fonction de coût $c(\mathbf{s})$ est inconnue, utiliser une forme quadratique en \mathbf{s} pour cette fonction permet de considérer cette forme fonctionnelle comme une approximation de Taylor au second ordre de $c(\mathbf{s})$:

$$(30) \quad c(\mathbf{s}) = A + \sum_{c=1}^C s_c g_c + .5 \times \sum_{m=1}^C \sum_{c=1}^C s_c s_m g_{cm} \equiv A + \mathbf{g}'\mathbf{s} + .5 \times \mathbf{s}'\mathbf{G}\mathbf{s}$$

avec $\mathbf{g} \equiv (g_1, \dots, g_C)$ et $\mathbf{G} \equiv [g_{cm} = g_{mc}]_{m,c=1,\dots,C}$. La matrice \mathbf{G} est symétrique et définie positive (ce qui assure que $c(\mathbf{s})$ est convexe en \mathbf{s} , même s'il suffit qu'elle soit simplement quasi-convexe). De telles fonctions de coût sont utilisées dans les modèles de choix d'assolement de la programmation mathématique positive (Howitt, 1995) et ont été utilisées dans le cadre de modélisations économétriques par Arnberg et Hansen (2007). L'assolement de référence est ici simplement défini par $\mathbf{s}^r \equiv -\mathbf{G}^{-1}\mathbf{g}$. On a alors :

$$(31) \quad c(\mathbf{s}) = A^r + .5 \times (\mathbf{s} - \mathbf{s}^r)' \mathbf{G} (\mathbf{s} - \mathbf{s}^r).$$

La matrice \mathbf{G} étant définie positive, la fonction $c(\mathbf{s})$ est minimale si $\mathbf{s} = \mathbf{s}^r$. Elle est alors égale à $A^r \equiv A - .5 \times \mathbf{s}^{r'} \mathbf{G} \mathbf{s}^r$, les termes A et A^r désignant des coûts fixes ne pouvant être identifiés. La fonction de coût quadratique est analogue à celles utilisées pour représenter les coûts d'ajustement de \mathbf{s} dans les modèles dynamiques à ajustement partiel. Dans ces modèles la matrice \mathbf{G} mesure la force des phénomènes de friction empêchant \mathbf{s} de trop s'écarter de \mathbf{s}^r qui est alors le dernier choix de \mathbf{s} . Le modèle considéré ici est statique, \mathbf{s}^r est fixe, mais les mécanismes à l'œuvre sont analogues.

En adaptant l'approche de Röhm et Dabbert (2003), il est possible de spécialiser la forme de $c(\mathbf{s})$ pour tenir compte des similarités des rôles des différentes cultures pour ce qui concerne les successions culturelles et l'organisation du travail. Si les cultures c et d sont jugées comme ayant des rôles similaires, il peut alors être utile d'imposer les contraintes $g_{cm} = g_{md}$ pour $m = 1, \dots, d$ afin de réduire le nombre de paramètres à estimer. Ce dernier croît très vite avec le nombre de cultures considérées. Ceci revient à spécifier $c(\mathbf{s})$ comme une fonction de $s_c + s_d$ et non plus de s_c et s_d séparément. Une logique similaire a conduit à la construction des groupes de cultures utilisés dans les modèles de choix d'assolement MNL emboîtés (chapitre 2).

Le profit espéré est alors donné par:

$$(32a) \quad \sum_{c=1}^C s_c \left(\pi_c^e(0) - (p_c \delta_c^y + \mathbf{w}' \delta_c^x) s_c \right) - \left[A + \sum_{c=1}^C s_c g_c + .5 \times \sum_{m=1}^C \sum_{c=1}^C s_c s_m g_{cm} \right]$$

ou encore, en regroupant les termes linéaires et quadratiques en s_c , par :

$$(32b) \quad \sum_{c=1}^C s_c \left(\pi_c^e(0) - g_c \right) - A - .5 \times \sum_{m=1}^C \sum_{c=1}^C s_c s_m \left(g_{cm} + 1_{c=m} (p_c \delta_c^y + \mathbf{w}' \delta_c^x) \right)$$

où $1_{c=m} = 1$ si $c = m$ et $1_{c=m} = 0$ sinon. Heckelei et Wolff (2003) ont proposé deux cas particuliers de la fonction de profit espéré (32). Le premier cas correspond au modèle où $g_c = g_{cm} = 0$ pour $c, m = 1, \dots, C$. Dans ce cas, la diversification de l'assolement n'est motivée que par les rendements marginaux décroissants de la terre. Ce modèle est similaire à celui de Moore et Negri (1992) mais s'en distingue par deux aspects. Moore et Negri n'imposent pas $g_c = 0$ (pour $c = 1, \dots, C$). En outre, leur modèle n'impose pas que la terre est un intrant essentiel pour la production, *i.e.* une culture peut générer des bénéfices ou des coûts sans que de la terre lui soit allouée. Cette propriété peut cependant être imposée dans leur modèle.

Le second cas considéré par Heckelei et Wolff (2003) correspond au modèle où $\delta_c^y = 0$ et $\delta_c^x = \mathbf{0}$, pour tout s_c où $c = 1, \dots, C$. Arnberg et Hansen (2007) considèrent également ce modèle. Dans ce cas, la diversification de l'assolement n'est due qu'aux effets décrits par la fonction de coût $c(\mathbf{s})$. Le modèle de choix d'assolement MNL est dérivé d'une fonction de profit espéré qui a la même structure.

La maximisation sur \mathbf{s} de la fonction de profit espéré (32) sous la contrainte de surface totale $\sum_{c=1}^C s_c = 1$ n'est pas particulièrement complexe (si les solutions sont toutes intérieures) mais nécessite l'introduction de nombreuses notations supplémentaires. Ces calculs sont développés dans l'Annexe B.

Ces calculs permettent de montrer que les conditions du premier (CO1) du programme de maximisation sont linéaires dans les assolements optimaux. Ils s'écrivent sous la forme :

$$(33a) \quad \pi_c^e(0) - \pi_1^e(0) - g_c - \lambda(p_1, \mathbf{w}, \delta_1) - \sum_{m=2}^C s_m^* q_{cm}(\gamma, \delta; \mathbf{p}, \mathbf{w}) = 0 \text{ pour } c = 2, \dots, C.$$

où $\delta_1 \equiv (\delta_1^y, \delta_1^x)$, $\lambda(p_1, \mathbf{w}, \delta_1) \equiv (p_1 \delta_1^y + \mathbf{w}' \delta_1^x)$ et $\delta \equiv (\delta_1^y, \delta_1^x, \dots, \delta_C^y, \delta_C^x)$. Le vecteur γ contient les éléments non-redondants de la matrice Γ . La matrice Γ est construite à partir des éléments de la matrice \mathbf{G} . Les paramètres g_c ($c = 2, \dots, C$) sont des fonctions d'éléments de \mathbf{d} et \mathbf{G} . Les paramètres $\mathbf{g} \equiv (g_2, \dots, g_C)$ et γ sont identifiables. La résolution en $\mathbf{s}^* \equiv (s_2^*, \dots, s_C^*)$ des CO1 permet de déterminer la forme des assolements optimaux. On a alors :

$$(33b) \quad s_c^* = \sum_{m=2}^C \kappa_{cm}(\gamma, \delta; \mathbf{p}, \mathbf{w}) \times (\pi_c^e(0) - \pi_1^e(0) - g_c - \lambda(p_1, \mathbf{w}, \delta_1)) \text{ pour } c = 2, \dots, C$$

où les termes $\kappa_{cm}(\gamma, \delta; \mathbf{p}, \mathbf{w})$ sont les éléments de la matrice inverse de $\mathbf{Q}(\gamma, \delta; \mathbf{p}, \mathbf{w}) \equiv [q_{cm}(\gamma, \delta; \mathbf{p}, \mathbf{w})]_{c,m=2, \dots, C}$. Les conditions \mathbf{G} est définie positive et δ est positif sont des conditions suffisantes pour l'inversibilité de $\mathbf{Q}(\gamma, \delta; \mathbf{p}, \mathbf{w})$. L'assolement de la culture de référence se déduit des autres à partir de la contrainte de surface :

$$(33c) \quad s_1^* = 1 - \sum_{m=2}^C s_m^*.$$

Il est intéressant de noter que l'hypothèse des rendements marginaux de la terre constants simplifie grandement les expressions (33a) et (33b). Cette hypothèse implique que $\delta = \mathbf{0}$. Dans ce cas les termes $q_{cm}(\gamma, \delta; \mathbf{p}, \mathbf{w})$ et $\kappa_{cm}(\gamma, \delta; \mathbf{p}, \mathbf{w})$ sont des paramètres qui ne dépendent que de γ , *i.e.* des paramètres constants, et le terme $\lambda(p_1, \mathbf{w}, \delta_1)$ est nul. On a alors :

$$(34a) \quad \pi_c^e(0) - \pi_1^e(0) - g_c - \sum_{m=2}^C s_m^* q_{cm}(\gamma) = 0 \text{ pour } c = 2, \dots, C$$

et :

$$(34b) \quad s_c^* = \sum_{m=2}^C \kappa_{cm}(\gamma) \times (\pi_c^e(0) - \pi_1^e(0) - g_c) \text{ pour } c = 2, \dots, C.$$

Supposer que la fonction de coût $c(\mathbf{s})$ ne contient pas de motifs de diversification, *i.e.* $\mathbf{G} = \mathbf{0}$, n'amène pas de simplification notable.

Les CO1 décrites dans l'équation (33a) permettent de définir des conditions estimantes des paramètres \mathbf{g} et γ plus faciles d'utilisation, dans le cas général, que le système des choix d'assolement optimaux (33b). Elles peuvent également être utilisées pour définir un système de conditions de Kuhn-Tucker permettant l'analyse des solutions en coin dans les choix d'assolement. En supposant que la sole de la culture 1 est toujours non nulle, on a alors :

$$(35a) \quad \pi_c^e(0) - \pi_1^e(0) - g_c - \lambda(p_1, \mathbf{w}, \delta_1) - \sum_{m=2}^C s_m^* q_{cm}(\gamma, \delta; \mathbf{p}, \mathbf{w}) = 0 \text{ si } s_c^* > 0$$

et :

$$(35b) \quad \pi_c^e(0) - \pi_1^e(0) - g_c - \lambda(p_1, \mathbf{w}, \delta_1) - \sum_{m=2}^C s_m^* q_{cm}(\gamma, \delta; \mathbf{p}, \mathbf{w}) \leq 0 \text{ si } s_c^* = 0.$$

Ces conditions de Kuhn-Tucker peuvent alors être utilisées pour l'analyse économétrique en utilisant l'approche de Wales et Woodland (1983). En effet, ces conditions sont linéaires dans les choix considérés. Elles sont analogues à celles qui sont dérivées par Wales et Woodland (1983) à partir d'un modèle de choix de consommation fondé sur la maximisation sous contrainte budgétaire d'une fonction d'utilité directe de forme quadratique (voir également Ransom (1987) pour l'intégration de la contrainte de budget). Bien entendu, le calcul des estimateurs des paramètres reste un défi dans le cas où le nombre de cultures considérées est supérieur à 4. Cependant, l'utilisation des techniques d'intégration des fonctions de vraisemblance par des méthodes de simulation semble offrir des perspectives intéressantes (Kao, Lee et Pitt, 2001).

1.5.4 Quelques aspects liés à l'utilisation empirique des modèles proposés

Dans le meilleur des cas, les utilisations d'intrants variables sont disponibles au niveau des cultures. Dans ce cas, l'estimation des paramètres d'un modèle de choix de production utilisant les choix d'assolement de forme MNL est très simple. En effet, le système d'équations considéré est de la forme :

$$(36) \quad \begin{cases} y_c = a_c(\mathbf{q}_c) - .5 \times \mathbf{w}'_c \mathbf{D}_c(\mathbf{q}_c) \mathbf{w}_c + e_c^y + \varepsilon_c^y & c = 1, \dots, C \\ \mathbf{x}_c = \mathbf{b}_c(\mathbf{q}_c) - \mathbf{D}_c(\mathbf{q}_c) \mathbf{w}_c + \mathbf{e}_c^x + \varepsilon_c^x & c = 1, \dots, C \\ \ln s_d - \ln s_1 = \rho(\mathbf{q}, \bar{\mathbf{z}}) (\pi_d^e(0) - \pi_1^e(0) - g_d(\mathbf{q}, \bar{\mathbf{z}}) - e_d^g) & d = 2, \dots, C \\ \pi_c^e(0) \equiv p_c a_c(\mathbf{q}_c) - \mathbf{w}' \mathbf{b}_c(\mathbf{q}_c) + .5 \times \mathbf{w}' \mathbf{D}_c(\mathbf{q}_c) \mathbf{w} p_c^{-1} + (p_c e_c^y - \mathbf{w}' \mathbf{e}_c^x) & c = 1, \dots, C. \end{cases}$$

Dans ce modèle, les paramètres à estimer sont les paramètres des fonctions $\mathbf{D}_c(\cdot)$, $a_c(\cdot)$, $\mathbf{b}_c(\cdot)$, $\rho(\cdot)$ et $g_d(\cdot)$ pour $c = 1, \dots, C$ et $d = 2, \dots, C$. Le terme ρ de la fonction de coût entropique est spécifié comme une fonction des variables de contrôle de l'hétérogénéité (\mathbf{q}) et des dotations en facteurs fixes ($\bar{\mathbf{z}}$). On s'attend en particulier à ce que $\rho(\mathbf{q}, \bar{\mathbf{z}})$ soit croissante en $\bar{\mathbf{z}}$ puisqu'une dotation importante en facteurs fixes tend à « flexibiliser » les choix d'assolement. Dans la même logique, ce terme doit être d'autant plus grand que l'exploitation considérée bénéficie de bonnes conditions de production. En particulier, de bonnes conditions pédoclimatiques offrent un large éventail de cultures et de pratiques culturales. Les termes de coût inobservés g_d sont spécifiés comme des fonctions de $(\mathbf{q}, \bar{\mathbf{z}})$ avec un terme d'erreur additif e_d^g :

$$(37a) \quad g_d \equiv g_d(\mathbf{q}, \bar{\mathbf{z}}) + e_d^g \text{ pour } d = 2, \dots, C.$$

L'ajout du terme d'erreur e_d^g est ici important car les termes g_d représentent des effets de contraintes, coûts et mécanismes inobservés. A ce propos, il est important de rappeler ici que les fonctions de coût utilisées pour construire les modèles de choix d'assolement sont, comme

leur nom l'indique, homogènes à des dépenses. Aussi, lorsque les données utilisées couvrent une période longue il convient de tenir compte de l'évolution éventuelle des « dépenses » implicites mesurées par ces fonctions de coût. L'utilisation de variables tendances semble indiquée ici, même si elle peut poser des problèmes d'identification. L'examen de l'évolution dans le temps des termes d'erreur des équations d'assolement peut révéler des questions liées à la « déflation » des fonctions de coût. Enfin, il est à noter que les termes d'erreurs des équations d'assolement sont composites en raison de la présence du terme $\rho(\mathbf{q}, \bar{\mathbf{z}})$ et de la définition des $\pi_c^e(0)$. Le terme d'erreur de la $d^{\text{ème}}$ équation d'assolement est donné par :

$$(37b) \quad -\rho(\mathbf{q}, \bar{\mathbf{z}})e_d^g + (p_d e_d^y - \mathbf{w}' \mathbf{e}_d^x) - (p_1 e_1^y - \mathbf{w}' \mathbf{e}_1^x).$$

En posant les hypothèses habituelles d'exogénéité des prix et des variables de contrôle :

$$(38) \quad E[\mathbf{e}_c / \mathbf{p}, \mathbf{w}, \bar{\mathbf{z}}, \mathbf{q}] = \mathbf{0} \text{ et } E[\boldsymbol{\varepsilon}_c / \mathbf{p}, \mathbf{w}, \bar{\mathbf{z}}, \mathbf{q}] = \mathbf{0} \text{ pour } c = 1, \dots, C$$

où $\mathbf{e}_c \equiv (e_c^y, \mathbf{e}_c^x, e_c^g)$, $\boldsymbol{\varepsilon}_c \equiv (\varepsilon_c^y, \boldsymbol{\varepsilon}_c^x)$ et $e_1^g \equiv 0$, on remarque aisément que le système d'équations (36) est un système de régressions empilées. L'estimation de ses paramètres peut donc être réalisée avec des estimateurs simples. La structure des termes d'erreur du modèle et la probable hétéroscédasticité des termes $\boldsymbol{\varepsilon}_c$ implique que le calcul de la distribution asymptotique des estimateurs utilisés doit tenir compte de l'hétéroscédasticité (et des corrélations inter-équations) des termes d'erreur du modèle. Il est à noter que dans ce cas la distinction entre les termes d'hétérogénéité \mathbf{e}_c et d'évènements aléatoires $\boldsymbol{\varepsilon}_c$ n'a aucune importance.

L'utilisation d'un modèle de choix de production basé sur les choix d'assolement dérivés de la fonction de coût quadratique est légèrement plus ardue. Dans le cas général, ces modèles sont de la forme :

$$(39) \quad \begin{cases} y_c = a_c(\mathbf{q}_c) - \delta_c^y s_c - .5 \times \mathbf{w}' \mathbf{D}_c(\mathbf{q}_c) \mathbf{w}_c + e_c^y + \varepsilon_c^y & c = 1, \dots, C \\ \mathbf{x}_c = \mathbf{b}_c(\mathbf{q}_c) + \delta_c^x s_c - \mathbf{D}_c(\mathbf{q}_c) \mathbf{w}_c + \mathbf{e}_c^x + \varepsilon_c^x & c = 1, \dots, C \\ \pi_d^e(0) - \pi_1^e(0) - g_d(\mathbf{q}, \bar{\mathbf{z}}) - \lambda(p_1, \mathbf{w}, \delta_1) - \sum_{m=2}^C s_m q_{dm}(\boldsymbol{\gamma}, \boldsymbol{\delta}; \mathbf{p}, \mathbf{w}) = -e_d^g & d = 2, \dots, C \\ \pi_c^e(0) \equiv p_c a_c(\mathbf{q}_c) - \mathbf{w}' \mathbf{b}_c(\mathbf{q}_c) + .5 \times \mathbf{w}' \mathbf{D}_c(\mathbf{q}_c) \mathbf{w} p_c^{-1} + (p_c e_c^y - \mathbf{w}' \mathbf{e}_c^x) & c = 1, \dots, C. \end{cases}$$

L'utilisation des CO1 (33a) est plus aisée que celle des choix d'assolements. Les équations dérivées des choix d'assolement (33b) sont très non linéaires dans leurs paramètres en raison de l'inversion de la matrice $\mathbf{Q}(\boldsymbol{\gamma}, \boldsymbol{\delta}; \mathbf{p}, \mathbf{w})$. La forme des modèles de choix d'assolement montre que les \mathbf{s} sont des fonctions des termes d'hétérogénéité \mathbf{e}_c . Les assolements s_c sont donc des variables endogènes dans toutes les équations du système : dans les équations dérivées des CO1 des choix d'assolement évidemment, mais également dans les équations d'offre et de demande d'intrants. L'utilisation de techniques à variables instrumentales est donc requise. Ces variables instrumentales doivent être construites à partir des variables exogènes du système, *i.e.* $(\mathbf{p}, \mathbf{w}, \bar{\mathbf{z}}, \mathbf{q})$. Les travaux de Chamberlain (1987) sur la forme des instruments optimaux suggèrent une procédure d'estimation par étapes. La première étape consiste à calculer des estimateurs, même potentiellement biaisés, des paramètres des modèles de choix d'assolement. La seconde étape utilise les estimations ainsi calculées pour faire des prédictions des s_c en tant qu'instruments de ces variables. Cette seconde étape consiste alors à construire un estimateur de la Méthode des Moments Généralisée des paramètres du modèle, en utilisant comme conditions estimantes les conditions d'orthogonalité construites à partir des produits croisés des termes d'erreur des équations du modèle avec des variables explicatives du modèle, les prédictions des s_c jouant le rôle de variables instrumentales pour les s_c .

Dans la plupart des cas, les données utilisées par les économistes agricoles sont issues de bases de données de comptabilités d'exploitations telles que le « Réseau d'Information Comptable Agricole » (RICA). Le principal défaut de ces données est qu'elles donnent une mesure des dépenses des intrants variables au niveau de l'exploitation, mais ne fournissent pas

d'indications sur la répartition de ces dépenses par culture. Dans ce cas, il n'est plus possible d'utiliser le sous-système des équations des demandes d'intrants par culture des systèmes (36) et (39). Ce sous-système est alors remplacé par un (sous)-système d'équations de répartition des intrants aux cultures. En notant $\bar{\mathbf{x}}$ le vecteur des quantités d'intrants variables utilisées sur l'exploitation (par unité de surface), ces équations s'écrivent :

$$(40) \quad \bar{\mathbf{x}} = \sum_{c=1}^C s_c (\mathbf{b}_c(\mathbf{q}_c) + \delta_c^x s_c - \mathbf{D}_c(\mathbf{q}_c) \mathbf{w}_c) + \sum_{c=1}^C s_c (\mathbf{e}_c^x + \boldsymbol{\varepsilon}_c^x).$$

L'utilisation de ces équations pose problème en raison de la présence du terme $\sum_{c=1}^C s_c \mathbf{e}_c^x$. Ce terme aléatoire n'est pas d'espérance nulle (conditionnellement à $(\mathbf{p}, \mathbf{w}, \bar{\mathbf{z}}, \mathbf{q})$). Cela provient simplement de ce que \mathbf{s} est une fonction des \mathbf{e}_c . Cette équation est un exemple de modèle à coefficients aléatoires corrélés selon la terminologie de Heckman et Vitlacil (1998). Ce problème ne peut être résolu par l'utilisation de techniques à variables instrumentales « standard ». En s'inspirant des travaux de Wooldridge (2008), Carpentier (2009) a défini des procédures d'estimation permettant de résoudre ce problème. Ces techniques reposent sur l'utilisation de « fonctions de contrôle » ou de « fonctions de correction » selon la terminologie de Wooldridge (2008). Elles requièrent également des hypothèses sur la distribution des \mathbf{e}_c . La seule hypothèse nécessaire sur les $\boldsymbol{\varepsilon}_c$ est l'indépendance de leurs moments du second ordre vis-à-vis des \mathbf{e}_c (les espérances des $\boldsymbol{\varepsilon}_c$ ne dépendent pas des \mathbf{e}_c par construction). La spécification de modèles cohérents dans leurs termes aléatoires et la distinction des termes \mathbf{e}_c et $\boldsymbol{\varepsilon}_c$ apparaissent comme des propriétés cruciales des approches de modélisation employées ici. Les \mathbf{e}_c sont les variables non observées qui lient les choix d'assolement des agriculteurs à leurs choix d'intrants et de rendements objectifs. Le fait de connaître leur rôle dans la structure des termes d'erreur des modèles utilisés peut alors être exploitée pour résoudre des problèmes d'identification des paramètres des fonctions de demande d'intrants variables. L'utilisation de cette approche par les fonctions de contrôle est décrite et illustrée dans les chapitres 3 et 4.

Conclusion

Les modèles de choix de production construits ici ont été conçus pour être employés dans le cadre de recherches impliquant des économistes et des agronomes. Ils se distinguent des modèles généralement utilisés pour décrire les choix de production des agriculteurs principalement sur trois points.

Tout d'abord, ces modèles reposent sur une spécification de la technologie agricole moins « flexible » que celles considérées dans les modèles économétriques fondés sur la théorie duale du producteur. Les modèles proposés sont cependant plus « transparents » dans leurs hypothèses relatives à la technologie et ont une structure simple. Cette simplicité permet leur utilisation comme modèles de base dans des structures plus complexes. En outre leurs paramètres sont directement interprétables, ce qui permet de les utiliser comme outils de collaboration avec des non-spécialistes de la micro-économie ou de l'économétrie. Ils ont d'ailleurs été conçus principalement dans cette optique.

Ensuite, ces modèles empruntent aux économistes agricoles utilisant les techniques de la programmation mathématique le concept de la fonction de coût implicite de gestion des assolements. Dans nos modèles, cette fonction de coût implicite traduit, en unité monétaire, les effets de « contraintes agronomiques » et de contraintes de disponibilité des facteurs fixes (autres que la terre), sur les choix d'assolement des agriculteurs. Elle est analogue aux fonctions de coût d'ajustement utilisées dans les modèles dynamiques à ajustement partiel. Cependant, les modèles présentés ici sont essentiellement statiques. Cette fonction de coût est certainement ce qui distingue le plus, notre approche de celles habituellement utilisées pour la modélisation économétrique des choix d'assolement.

Enfin, ces modèles contiennent des termes aléatoires qui font partie intégrante de la spécification choisie. Dans la plupart des modèles de choix de production agricole, les termes d'erreurs sont simplement « ajoutés » et interprétés comme des erreurs d'approximation. Dans

nos modèles, les termes aléatoires représentent soit l'hétérogénéité des conditions de production des exploitations, soit les aléas inhérents à la production agricole. Cette cohérence de nos modèles dans leurs parties aléatoires permet de résoudre certains problèmes d'inférence statistique. Par exemple, les quantités d'intrants variables sont généralement mesurées au niveau de l'exploitation. Estimer leurs utilisations, culture par culture, est intéressant en soi. Cela permet également de vérifier ou de juger de la pertinence des résultats issus des modèles utilisés. Cependant, l'estimation de ces utilisations par culture est complexe car, si les assolements sont les clés de répartition des intrants aux cultures, ils dépendent également des quantités d'intrants utilisées sur chaque culture. L'information apportée par la structure des termes aléatoires permet l'identification de l'allocation des intrants variables aux cultures, *i.e.* de résoudre le problème lié à la simultanéité des utilisations d'intrants et des choix d'assolement.

L'ensemble des modèles proposés offre des perspectives intéressantes de développements futurs, particulièrement dans deux domaines : la prise en compte des solutions en coin dans les choix d'assolement et la prise en compte des effets de rotation dans l'optique d'une modélisation des choix « dynamiques » des agriculteurs. Le principal avantage des modèles proposés ici par rapport aux modèles directement issus de la théorie de la dualité est que leur construction repose principalement sur des fonctions de production et une spécification, certes très fruste, mais explicite des autres déterminants des choix d'assolement des agriculteurs. La transparence de la structure de ces modèles est un avantage dans l'optique de leur généralisation. En effet, préciser ou introduire certaines spécificités de la production agricole ou des motivations des agriculteurs peut se faire en modifiant certaines parties de ces modèles. La structure des modèles directement issus de la théorie duale du producteur est élégante. Mais elle est généralement trop opaque pour permettre des généralisations satisfaisantes.

Par exemple, le modèle de choix d'assolement fondé sur des fonctions de coût quadratiques peut être utilisé pour l'analyse des solutions en coin. Les développements récents des techniques d'inférence statistique assistée par des méthodes de simulation offre des perspectives intéressantes dans ce contexte.

La question de la gestion des successions culturales joue un rôle omniprésent dans les discussions entre économistes et agronomes. Les modèles proposés ici peuvent offrir une base à des recherches menées dans ce sens. En particulier les modèles de fonctions de production présentés peuvent facilement intégrer les effets de successions culturales. Ceci-dit considérer les effets des rotations culturales selon les principes d'optimisation dynamique en toute rigueur s'avère très complexe. Des solutions approchées, ou tout au moins exploratoires, sont toutefois envisageables.

Références

- Antle J.M. 1987. "Econometric Estimation of Producer's Risk Attitudes." *American Journal of Agricultural Economics* 69(3), 509-522.
- Antle J.M. 1983a. "Incorporating Risk in Production Analysis." *American Journal of Agricultural Economics* 65(5), 1099-1113.
- Antle J.M. 1983b. "Sequential Decision Making in Production Models." *American Journal of Agricultural Economics* 65(2), 282-290.
- Antle J.M. et S.A. Hatchett. 1986. "Dynamic Input Decisions in Econometric Production Models." *American Journal of Agricultural Economics* 68(4), 939-949.
- Arnberg, S. et L.G. Hansen. 2007. "The Dynamics of Farm Land Allocation – Short and Long Run Reactions in a Long Micro Panel." American Agricultural Economics Association Annual Meeting, Portland, July 29 - August 1, 2007.
- Aunka et Shumway. 1996. "Allocatable Fixed Inputs and Jointness in Agricultural production: More implications." *Agricultural and Resource Economics Review* 25:143-48.
- Blundell, R. and S. Bond. 2000. "GMM Estimation with Persistent Panel Data: An Application to Production Functions," *Econometric Reviews* 19, 321-340.
- Buyse, J., Fernagut, B., Harmignie, O., Henry de Frahan, B., Lauwers, L., Polomé, P., Van Huylenbroeck, G. et J. Van Meesel. 2004. "Modelling the Impact of Sugar Reform on Belgian Agriculture." Selected Paper presented at the International Conference on Policy Modelling, Paris, 30 June – 2 July, 2004.
- Buyse, J., Van Huylenbroeck, G. et L. Lauwers. 2007. "Normative, Positive and Econometric Mathematical Programming as Tools for Incorporation of Multifunctionality in Agricultural Policy Modelling." *Agriculture, Ecosystems and Environment* 120:70-81.
- Carpentier, A. 2009. "Allocation of input uses to activities with correlated random coefficient models: an agricultural production example." Working Papers.
- Carpentier, A. et R.D. Weaver. 1997. "Damage Control Productivity: Why specification matters." *American Journal of Agricultural Economics* 52:11-22.

- Carpentier, A. et R.D. Weaver. 1996. "Intertemporal and Interfirm Heterogeneity: Implications for Pesticide Productivity." *Canadian Journal of Agricultural Economics* 44:219-236.
- Caswell, M., and D. Zilberman. 1985. "The Choice of Irrigation Technologies in California." *American Journal of Agricultural Economics* 67(2):224-234.
- Chambers, R.G., 1988. "Applied Production Analysis: A Dual Approach." Cambridge University Press, New York.
- Chambers, R. et R. Just. 1989. "Estimating Multioutput Technologies." *American Journal of Agricultural Economics* 71:980-95.
- Chavas, J.-P. and T. Holt. 1990. "Acreage Decisions under Risk: The Case of Corn and Soybeans." *American Journal of Agricultural Economics* 72(3):529-538.
- Coyle, B. T. 1992. "Risk Aversion and Price Risk in Duality Models of Production: A Linear Mean-Variance Approach." *American Journal of Agricultural Economics* 74:849-859.
- Coyle, B. T. 1999. "Risk Aversion and Price Risk in Duality Models of Production: A Mean-Variance Approach." *American Journal of Agricultural Economics* 81:553-567.
- Eckstein, Z. 1984. "A Rational Expectations Model of Agricultural Supply." *The Journal of Political Economy* 92(1):1-19.
- Feder G. 1979. "Pesticides, Information and Pest Management under Uncertainty." *American Journal of Agricultural Economics* 61(1), 97-103.
- Griliches, Z. et J. Mairesse. 1995 "Production Functions: the Search for Identification" National Bureau of Economic Research, Cambridge, MA, NBER Working Paper No. 5067.
- Guyomard, H., Baudry, M. and A. Carpentier. 1996. "Estimating Crop Supply Response in the Presence of Farm Programmes: Application to the CAP." *European Review of Agricultural Economics* 23:401-420.
- Heckeley T. 1997. "Positive Mathematical Programming: Review of the Standard Approach." CAPRI Working paper 97-03. Bonn : University of Bonn.

- Heckelei, T. et H. Wolff. 2003. "Estimation of constrained optimisation models for agricultural supply analysis based on generalised maximum entropy." *European Review of Agricultural Economics* 30:27-50.
- Hennessy, D.A. 2006. "On Monoculture and the Structure of Crop Rotations" *American Journal of Agricultural Economics* 88:900-914.
- Hennessy, D.A. 2004. "On Monoculture and the Structure of Crop Rotations" Working Paper - Center for Agricultural and Rural Development, Iowa State University.
- Henry de Frahan, B., Buysse, J., Polome', P., Fernagut, B., Harmignie, O., Lauwers, L., Van Huylenbroeck, G., Van Meensel, J. 2007. "Positive mathematical programming for agriculture and environmental policy analysis: review and practice." In: Weintraub, A., Bjorndal, T., Epstein, R., Romero, C. (Eds.), *Management of Natural Resources: A Handbook of Operations Research Models, Algorithms, and Implementations*. Kluwer Academic Publishers, Kluwer's International Series in Operations Research and Management Science (Series Editor: Hillier, F.S.), in press.
- Howitt, E. 1995. "Positive Mathematical Programming." *American Journal of Agricultural Economics* 77:329-342.
- Just, R.E. et R.D. Pope. 2001. "The Agricultural Producer: Theory and Statistical Measurement." In B.L. Gardner and G.C. Rausser, ed. *Handbook of Agricultural Economics*, Vol. 1A. Amsterdam, Elsevier.
- Just R.E. et R.D. Pope 1979. "Production Function Estimation and Related Risk Considerations." *American Journal of Agricultural Economics* 61(2), 276-284.
- Just, R.E., Zilberman, D., et E. Hochman. 1983. "Estimation of Multicrop Production Functions." *American Journal of Agricultural Economics* 65:770-780.
- Kao, C., L.-F. Lee and M.M. Pitt, 2001. "Simulated Maximum Likelihood Estimation of the Linear Expenditure System with Binding Non-negativity Constraints." *Annals of Economics and Finance* 2:215-55.

- Leathers, H. 1991. "Allocable Fixed Inputs as a Cause of Joint Production: A Cost Function Approach." *American Journal of Agricultural Economics* 73:1080-1090.
- Lichtenberg, E. 1989. "Land Quality, Irrigation Development and Cropping Pattern in the Northern High Plains." *American Journal of Agricultural Economics* 71(1):187-94.
- Livingston, M., M.J. Roberts, and J. Rust. 2008. "Optimal Corn and Soybean Rotations." Paper presented at AAEA annual meeting, Orlando, Florida, 27-29 July.
- Lynne, G.D. 1988. "Allocable Inputs fixed and Jointness in Agricultural Production: Implications for Economic Modelling: Comment." *American Journal of Agricultural Economics* 70:947-949.
- McElroy, M.B. 1987. "Additive General Error Models for Production, Cost, and Derived Demand or Share systems." *Journal of Political Economy* 95:737-757.
- McFadden, D. 1974. "Conditional Logit Analysis of Quantitative Choice Behavior." In P. Zarembka, ed. *Frontiers in Econometrics*. New-York: Academic Press.
- Moore, M.R. and A. Dinar. 1995. "Water and Land as Quantity-Rationed Inputs in California Agriculture: Empirical Tests and Water Policy Implications." *Land Economics* 74(4):445-461.
- Moore, M.R., Gollehon, N.R. and M.B. Carey. 1994. "Multicrop Production Decisions in Western Irrigation Agriculture: The Role of Water Price." *American Journal of Agricultural Economics* 11:143-158.
- Moore, M. R. et D. H. Negri. 1992. "A Multicrop Production Model of Irrigated Agriculture, Applied to Water Allocation Policy of the Bureau of Reclamation." *Journal of Agricultural and Resource Economics* 17:29-43.
- Moro, D. and P. Sckokai. 1999. "Modelling the CAP Arable Crop Regime in Italy: Degree of Decoupling and Impact of Agenda 2000." *Cahiers d'économie et sociologie rurales* 53:50-73.

- Moschini, G. and D. Hennessy. 2001. "Uncertainty, Risk Aversion and Risk Management for Agriculture." Chapter 2 in B.L. Gardner and G. Rausser, Eds., *Handbook of Agricultural Economics* Volume 1A, Chapter 1. Amsterdam, The Netherlands.
- Oude Lansink, A. and J. Peerlings. 1996. "Farm-specific Impacts of Quantitative Restrictions of N-Fertiliser Use in Dutch Arable Farming." *Journal of Agricultural Economics* 52:38-52.
- Oude Lansink, A. et S. Stefanou. 2001. "Dynamic Area Allocation and Economies of Scale and Scope." *Journal of Agricultural Economics* 52:38-52.
- Ozarem, P.F. and J.A. Miranowski. 1994. "A Dynamic Model of Acreage Allocation with General and Crop-Specific Soil Capital." *American Journal of Agricultural Economics* 76:385-395.
- Paris, Q. 1989. "A Sure Bet on Symmetry." *American Journal of Agricultural Economics* 71:344-351.
- Paris, Q. et R.E. Howitt. 1998. "An Analysis of Ill-posed Production Problems using Maximum Entropy." *American Journal of Agricultural Economics* 80:124-138.
- Pope, R. D. and R. E. Just. 2003. "Distinguishing Errors in Measurement from Errors in Optimization." *American Journal of Agricultural Economics* 85:348-358.
- Ransom, M., 1987. "A Coherency on Demand Systems with Binding Non-negativity Constraints." *Journal of Econometrics* 34:355-359.
- Shumway, R. 1995. "Recent Duality Contributions in Production Economics." *Journal of Agricultural and Resource Economics* 20:178-194.
- Shumway, C.R., Pope, R.D. and E.K. Nash. 1984. "Allocatable Fixed Inputs and Jointness in Agricultural Production: Implications for Economic Modeling". *American Journal of Agricultural Economics* 66:72-78.
- Thomas, A. 2003. "A Dynamic Model of On-Farm Integrated Nitrogen Management." *European Review of Agricultural Economics* 30(4):439-460.
- Train, K. 2003. *Discrete Choice Methods with Simulation*. Cambridge University Press.

- Wales, T.J., A.D. Woodland, 1983. "Estimation of Consumer Demand Systems with Binding Non-Negativity Constraints." *Journal of Econometrics* 21:263-285.
- Weaver, R. D. 1983. "Multiple Input, Multiple Output Production Choices and Technology in the U.S. Wheat Region." *American Journal of Agricultural Economics* 65:45-56.
- Wooldridge, J. M., 2008. "Instrumental Variables Estimation of the Average Treatment Effect in Correlated Random Coefficient Models," *Mimeo, Michigan State University*.
- Wooldridge, J.M. 2003. "Cluster-Sample Methods in Applied Econometrics," *American Economic Review* 93, 133-138.
- Wu, J., and K. Segerson. 1995. "The Impact of Policies and Land Characteristics on Potential Groundwater Pollution in Wisconsin." *American Journal of Agricultural Economics* 77(4):1033-47.
- Zellner, A., J. Kmenta et J. Dreze. 1966. "Specification and Estimation of Cobb-Douglas Production Models," *Econometrica*. 34:784-95.
- Zhengfei, G., Oude Lansink, A., Van Ittersum M. and Ada Wossink. 2006. "The Integration of Agronomic Principles in Production Function Estimation: Dichotomy of Growth Inputs and Facilitating Inputs." *American Journal of Agricultural Economics* 88: 203-214.

Annexe A. Choix intertemporels : une approche exploratoire

Dans les modèles présentés ici ces effets sont simplement résumés dans la forme de $c(\mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ et dans les effets des s_c sur les $\pi_c^e(p_c, \mathbf{w}; \mathbf{q}_c, s_c)$.

Considérer les effets des assolements passés revient à spécifier les marges brutes espérées comme des fonctions de $\mathbf{s}_{(-1)}$ où l'indice (-1) indique l'assolement de (ou des) l'année(s) précédentes. On a alors : $\pi_c^e(p_c, \mathbf{w}; \mathbf{q}_c, s_c, \mathbf{s}_{(-1)})$. Il convient ensuite de spécifier comment les agriculteurs mesurent les effets de leurs choix actuels sur leurs résultats futurs, quel poids ces effets ont sur leurs décisions d'aujourd'hui, à quel horizon de temps ils raisonnent et enfin leur schéma d'anticipation de ce qui est susceptible d'évoluer (contexte économique, conditions climatiques, ...). La seule solution rigoureuse consiste en la spécification d'un modèle d'optimisation intertemporel « complet ». Une solution approximative consiste à choisir un horizon de temps donné, *e.g.* un an, un taux d'actualisation des profits futurs (δ) et à définir une mesure des effets bénéfiques du choix d'assolement actuel sur le profit de l'année d'après. Pour ce qui concerne les reliquats d'engrais, cette mesure peut être simplement définie (Thomas, 2003): il suffit de mesurer l'effet du choix de \mathbf{s} sur les reliquats futurs et de l'introduire en tant que dépenses (actualisées) d'engrais économisées pour l'année d'après. Si l'engrais est le $k^{\text{ième}}$ intrant variable, les reliquats laissés par la culture c sont mesurés par les termes : $\partial b_{kd}(\mathbf{q}_d, s_d, \mathbf{s}_{(-1)}) / \partial s_{c(-1)}$ qui sont en principe égaux pour $d = 1, \dots, C$. La mesure des autres bénéfices futurs des effets de rotation est plus délicate car elle dépend des couples culture actuelle/culture à venir. Or la culture à venir n'est en principe pas encore choisie et les prix de l'année à venir ne sont pas connus. Une approximation possible consiste à choisir comme culture à venir la culture la plus probable dans les schémas de rotation usuels et à utiliser comme anticipation des prix futurs les prix actuels. Ceci revient à remplacer les fonctions de marge brute de la fonction de profit $\bar{\pi}(\mathbf{p}, \mathbf{w}; \mathbf{s}; \mathbf{q}, \bar{\mathbf{z}})$ par :

$$\pi_c^e(p_c, \mathbf{w}; \mathbf{q}_c, s_c, \mathbf{s}_{(-1)}) + \delta \frac{\partial \pi_{v(c)}^e(p_{v(c)}, \mathbf{w}; \mathbf{q}_{v(c)}, s_{v(c)}, \mathbf{s})}{\partial s_{c(-1)}} \text{ pour } c = 1, \dots, C$$

où $v(c)$ désigne la culture qui succède la culture c dans schémas de rotation usuels. Cette approche peut être utilisée pour explorer les effets des successions culturales.

Annexe B : Assolements optimaux dans le cas de la fonction de coût quadratique

La maximisation sur \mathbf{s} de la fonction de profit espéré (31b) :

$$\sum_{c=1}^C s_c (\pi_c^e(0) - d_c) - A - .5 \times \sum_{m=1}^C \sum_{c=1}^C s_c s_m (g_{cm} + 1_{c=m} (p_c \delta_c^y + \mathbf{w}' \delta_c^x))$$

sous la contrainte de surface totale $\sum_{c=1}^C s_c = 1$ est plus aisée sous forme matricielle et après intégration de la contrainte de surface. Après intégration de cette contrainte, la fonction de profit espéré (31b) peut être écrite sous la forme :

$$(B2a) \quad (\Delta' \boldsymbol{\pi}^e(0) - \mathbf{g} - \mathbf{j} \times (p_1 \delta_1^y + \mathbf{w}' \delta_1^x))' \mathbf{s}^- - .5 \times \mathbf{s}' \mathbf{Q}(\boldsymbol{\gamma}, \boldsymbol{\delta}; \mathbf{p}, \mathbf{w}) \mathbf{s}^- + f$$

où $\mathbf{s}^- \equiv (s_2, \dots, s_C)$, f est un terme ne dépendant pas de \mathbf{s}^- , $\boldsymbol{\pi}^e(0) \equiv (\pi_1^e(0), \dots, \pi_C^e(0))$, \mathbf{j} est un vecteur unitaire de dimension $C-1$, \mathbf{I} la matrice identité de dimension $C-1$, $\boldsymbol{\delta} \equiv (\delta_1^y, \delta_1^x, \dots, \delta_C^y, \delta_C^x)$, $\boldsymbol{\gamma}$ est le vecteur des éléments non redondants de la matrice $\boldsymbol{\Gamma}$:

$$(B2b) \quad \Delta' \equiv [\mathbf{I} \quad -\mathbf{j}], \quad \mathbf{G} \equiv \begin{bmatrix} g_{11} & \mathbf{g}'^- \\ \mathbf{g}^- & \mathbf{G}^- \end{bmatrix}, \quad \mathbf{g} \equiv \Delta' \mathbf{d} + \mathbf{g}^- ,$$

$$(B2c) \quad \boldsymbol{\Gamma} \equiv \mathbf{G}^- - \mathbf{j} \mathbf{g}'^- - \mathbf{g}^- \mathbf{j}' + \mathbf{j} \mathbf{j}' g_{11} = \Delta' \mathbf{G} \Delta \equiv [\gamma_{cm} = \gamma_{mc}]_{m,c=2,\dots,C} ,$$

$$(B2d) \quad \mathbf{H}(\boldsymbol{\delta}; \mathbf{p}, \mathbf{w}) \equiv 2 \times \begin{bmatrix} p_2 \delta_2^y + \mathbf{w}' \delta_2^x & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & p_C \delta_C^y + \mathbf{w}' \delta_C^x \end{bmatrix} + 2 \times \mathbf{j} \mathbf{j}' \times (p_1 \delta_1^y + \mathbf{w}' \delta_1^x)$$

et :

$$(B2e) \quad \mathbf{Q}(\boldsymbol{\gamma}, \boldsymbol{\delta}; \mathbf{p}, \mathbf{w}) = \boldsymbol{\Gamma} + \mathbf{H}(\boldsymbol{\delta}; \mathbf{p}, \mathbf{w}) \text{ (ouf !)}.$$

Les conditions du premier ordre du programme de maximisation de (B2a) en \mathbf{s}^- s'écrivent alors :

$$(B3a) \quad \Delta' \boldsymbol{\pi}^e(0) - \mathbf{g} - \mathbf{j} \times (p_1 \delta_1^y + \mathbf{w}' \delta_1^x) - \mathbf{Q}(\boldsymbol{\gamma}, \boldsymbol{\delta}; \mathbf{p}, \mathbf{w}) \mathbf{s}^{-*}(\boldsymbol{\pi}^e(0), \mathbf{p}, \mathbf{w}; \boldsymbol{\gamma}, \boldsymbol{\delta}) = \mathbf{0}$$

ce qui donne les assolements optimaux sous la forme :

$$(B3b) \quad \mathbf{s}^*(\boldsymbol{\pi}^e(0), \mathbf{p}, \mathbf{w}; \mathbf{g}; \boldsymbol{\gamma}, \boldsymbol{\delta}) = \mathbf{Q}(\boldsymbol{\gamma}, \boldsymbol{\delta}; \mathbf{p}, \mathbf{w})^{-1} \left[\boldsymbol{\Delta}' \boldsymbol{\pi}^e(0) - \mathbf{g} - \mathbf{j} \times (p_1 \boldsymbol{\delta}_1^y + \mathbf{w}' \boldsymbol{\delta}_1^x) \right].$$

Les formes des conditions du premier ordre du programme de maximisation et des assolements optimaux sont particulièrement simples dans le cas où $\boldsymbol{\delta} = \mathbf{0}$, i.e. dans le cas où les rendements de la terre sont constants. On a alors :

$$(B4a) \quad \boldsymbol{\Delta}' \boldsymbol{\pi}^e(0) - \mathbf{g} - \boldsymbol{\Gamma} \mathbf{s}^*(\boldsymbol{\pi}^e(0); \mathbf{g}; \boldsymbol{\gamma}) = \mathbf{0}$$

ce qui donne les assolements optimaux sous la forme :

$$(B4b) \quad \mathbf{s}^*(\boldsymbol{\pi}^e(0); \mathbf{g}; \boldsymbol{\gamma}) = \boldsymbol{\Gamma}^{-1} (\boldsymbol{\Delta}' \boldsymbol{\pi}^e(0) - \mathbf{g}).$$

Les conditions du premier ordre (B3a) forment un système de $C - 1$ équations de la forme :

$$(B5a) \quad \pi_c^e(0) - \pi_1^e(0) - g_c - \lambda(p_1, \mathbf{w}, \boldsymbol{\delta}_1) - \sum_{m=2}^C s_m^* q_{cm}(\boldsymbol{\gamma}, \boldsymbol{\delta}; \mathbf{p}, \mathbf{w}) = 0 \text{ pour } c = 2, \dots, C,$$

où $\boldsymbol{\delta}_1 \equiv (\boldsymbol{\delta}_1^y, \boldsymbol{\delta}_1^x)$, $\lambda(p_1, \mathbf{w}, \boldsymbol{\delta}_1) \equiv (p_1 \boldsymbol{\delta}_1^y + \mathbf{w}' \boldsymbol{\delta}_1^x)$ et $\mathbf{Q}(\boldsymbol{\gamma}, \boldsymbol{\delta}; \mathbf{p}, \mathbf{w}) \equiv [q_{cm}(\boldsymbol{\gamma}, \boldsymbol{\delta}; \mathbf{p}, \mathbf{w})]_{c,m=2,\dots,C}$. Les $C - 1$ assolements qui en dérivent peuvent alors être écrits sous la forme

$$(B5b) \quad s_c^* = \sum_{m=2}^C \kappa_{cm}(\boldsymbol{\gamma}, \boldsymbol{\delta}; \mathbf{p}, \mathbf{w}) \times [\pi_c^e(0) - \pi_1^e(0) - g_c - \lambda(p_1, \mathbf{w}, \boldsymbol{\delta}_1)] \text{ pour } c = 2, \dots, C,$$

où $\mathbf{Q}(\boldsymbol{\gamma}, \boldsymbol{\delta}; \mathbf{p}, \mathbf{w})^{-1} \equiv [\kappa_{cm}(\boldsymbol{\gamma}, \boldsymbol{\delta}; \mathbf{p}, \mathbf{w})]_{c,m=2,\dots,C}$. Bien entendu, l'assolement de la culture 1 se déduit des autres à partir de la contrainte de surface:

$$(B5c) \quad s_1^* = 1 - \sum_{m=2}^C s_m^*.$$

Chapitre II

On Modelling Acreage Decisions

within the Multinomial Logit Framework²

Résumé

La forme Logit Multinomiale a déjà été utilisée en économie de la production agricole pour spécifier des équations de parts de surface, mais il n'a jamais été démontré que cette forme particulière pouvait être dérivée d'un programme de maximisation du profit. Cet article étend les travaux de Caswell et Zilberman (1985) et de Wu et Segerson (1995), en développant le cadre théorique qui permet de dériver des équations de parts de surface de forme Logit Multinomiale. Deux approches sont considérées : « l'approche par la fonction de coût » et « l'approche par les choix discrets ». Il est ensuite montré que les parts de surfaces de forme Logit Multinomiale peuvent être utilisées pour définir un modèle économétrique multi-produit qui considère la terre comme un input fixe allouable. Enfin, plusieurs généralisations de ce modèle sont proposées. En particulier ce type de modèle peut facilement être étendu pour tenir compte des effets des rotations culturales dans les choix d'assolements. Une application empirique est réalisée à partir de données françaises sur des exploitants spécialisés en céréales de 1989 à 2006.

² This chapter is based on a paper jointly written with Alain Carpentier.

2.1 Introduction

The Multinomial Logit (MNL) framework is mainly used for modeling discrete choices or market shares (McFadden 1974), but its application has also been considered for consumer budget shares (Theil 1969) or cost shares (Considine and Mount 1984). Since the publication of McFadden's (1974) seminal article numerous generalizations of the standard multinomial logit model have then been introduced in the literature (Train 2003; Ackerberg et al. 2007), increasing the supply of models and inference procedures available for applied work.

The MNL framework has been used in the agricultural production economics literature to model acreage share choices (*e.g.*, Caswell and Zilberman 1985; Bewley *et al.* 1987; Lichtenberg 1989; Wu and Segerson 1995), crop decisions (Livingston *et al.* 2008) or land use decisions (*e.g.*, Lubowski, Plantinga and Stavins 2006). The acreage share models built within the MNL framework are mainly used for three reasons: (i) they ensure that the predicted share functions (strictly) lie in the interior of the zero-one interval, (ii) they are parsimonious in parameters and (iii) they are empirically tractable thanks to the so-called log-linear transformation. Nevertheless, the MNL framework was mainly employed for modeling plot level discrete decisions, the work of Wu and Segerson (1995) being a notable exception in this respect (see also Bewley *et al.* 1987). Furthermore, the MNL acreage share models were not integrated into economic production choice models and, as a result, were mainly used as convenient empirical functional forms for modeling acreage share choices.

The first objective of this article is to extend the pioneering works of Caswell and Zilberman (1985) and of Wu and Segerson (1995) by developing further the theoretical background of the MNL acreage share models. The (standard) MNL acreage share functional form considered in this article is defined as:

$$(1) \quad s_c(\boldsymbol{\pi}, \mathbf{c}, a) = \frac{\exp[a(\pi_c - c_c)]}{\sum_{m=1}^C \exp[a(\pi_m - c_m)]}, \quad k = 1, \dots, K.$$

In past studies using MNL acreage share models the arguments of the exponential functions in (1) were defined as linear functions of crop choices determinants. It is shown that these reduced form functions can be replaced by the $a(\pi_c - c_c)$ terms, where $\pi_c - c_c$ is a well defined measure of crop c profitability and where a is a parameter with a simple interpretation. Wu and Segerson (1995) defined model (1) by considering profit maximization at the farm level with land as an allocable fixed input and used MNL functional forms for their empirical acreage share equations. But they did not provide the link between their theoretical and empirical models. It is shown in this article that the MNL acreage share model can be derived from a farm level profit maximization program where the (restricted) profit function is defined as the weighted sum of the crop gross margins (the weights being the acreage shares) minus an “implicit cost function” of the chosen acreage. This approach is hereafter called the “cost function approach”. Caswell and Zilberman (1985) derived model (1) by aggregation at the farm level of crop (discrete) choices at the plot level. This approach is hereafter called the “discrete choice approach”. This approach is first extended in a simple dynamic setting with acreage adjustment costs in order to address a problem not considered before: although crop decisions are made at the plot level, these decisions are not independent the one from each other.

The second objective of this article is to show that MNL acreage share models can be used to define simple multi-crop econometric models with land as an allocable fixed input. These models are systems composed of yield supply, variable input demand and acreage share demand equations. The econometric models derived along these lines are fairly easy to

implement in practice. Their simple structure is particularly useful in research projects involving economists and non-economists.³

The third purpose of this article is to suggest generalizations of the standard MNL acreage share model (1). It is shown that Nested MNL acreage share functions can be derived following the “cost function approach”. The other proposed generalization of the standard MNL acreage share model extends the “discrete choice approach” of Caswell and Zilberman (1985) in a dynamic setting by exploiting the flexibility of the MNL discrete choice models. Using Rust’s (1987) framework, it is possible to build dynamic acreage choice models accounting for crop rotation effects, one of the major motives for crop diversification. Livingston *et al* (2008) also use the dynamic MNL discrete choice framework, but their perspective is normative and defined at the plot level.

The article is organized as follows. The assumptions necessary to build MNL acreage share models are discussed in the first section. The second and the third sections present the “cost function based approach” and the “discrete choice based approach” for defining standard MNL acreage share models. In the fourth section an application of the proposed multi-crop econometric models integrating the “cost function” and “discrete choice” (standard) MNL acreage share models is presented. It aims at illustrating the empirical relevance of the proposed models, at comparing them and at presenting their limits. In the fifth section two generalizations of the standard MNL acreage share models are presented to illustrate the potential of this framework as a basis for modeling acreage decisions. The last section provides concluding remarks and proposes directions for further research.

³ These models are developed by the authors for that purpose.

2.2 Main assumptions on the multi-crop production technology

The modeling frameworks used by agricultural economists to model farmers' acreage decisions differ by their focus on one or two of the main motives for crop diversification: decreasing marginal return to crop acreages (or more generally scale and scope economies), (production or/and price) risk spreading, constraints associated to allocated quasi-fixed factors (other than land) or crop rotation effects. Multi-crop econometric models considering land as fixed but allocable mostly focus on decreasing marginal returns to crop acreage (see, *e.g.*, Just *et al.* 1983; Chambers and Just 1989; Moore and Negri 1992;) and on risk spreading (see, *e.g.*, Chavas and Holt 1990; Coyle 1992) as the motives for crop diversification. Crop rotation effects are more rarely considered in multi-crop econometric models, probably due to the complexity of dynamic choice modeling (see, *e.g.*, Ozarem and Miranowski 1994; Thomas 2003). Although they also consider other motives for crop diversification the distinctive feature of the (positive) mathematical programming ((P)MP) models is that they allow to consider constraints on acreage choices faced by farmers (Howitt 1995).

The MNL models are mainly built by considering the constraints on acreage choices as the farmers' motive for crop diversification. These constraints are agronomic constraints (impossible or "forbidden" rotations) and/or constraints associated to limiting quantities of quasi-fixed inputs (labor, machinery...). These constraints are represented by the acreage management cost function in the "cost function approach" and by adjustment costs in the "discrete choice approach". In this respect, this article contributes to the growing literature linking programming models and duality-based models (see, *e.g.*, Howitt 1995 or Heckelevi and Wolff 2003).

The MNL acreage share models share several assumptions. The first is the farmers' risk neutrality. This assumption is innocuous where risk issues can be neglected. Although it appears restrictive, it is imposed in models not considering risk spreading motives for crop

diversification. Besides this assumption on farmers' attitude toward risk, the derivation of MNL acreage share models is based on two main assumptions related to the production technology:

H1) marginal short run returns to land are constant in acreages ;

H2) variable input uses do not depend on quasi-fixed input quantities, at least "locally".

These assumptions are uncommon and deserve comments.

Constant short run returns to acreages is used as a simplifying assumption in multi-crop econometric models considering risk spreading as the motive for crop diversification or in ((P)MP) models. This assumption is also imposed by Wu and Segerson (1995). The MNL models considered here can not be extended to accommodate marginal gross margins decreasing in acreages, at least as far as the usual representation of the scale effects is considered. However the "dynamic discrete choice" MNL acreage share model presented in the last section accounts for the main source of decreasing short run returns to land, *i.e.* crop rotation effects. As the acreage allocated to a given crop increases, farmers need to allocate land with less favorable crop rotations. In this respect the "dynamic discrete choice" MNL acreage share model account for decreasing marginal gross margins according to step effects. Crop returns are constant for a given crop rotation. They decrease as the crops are produced on plots with less suitable crop rotations.

The second assumption is admittedly restrictive as it implies that the variable input uses depend on the available quantities of labor and machinery only through the acreage choices. The available quantities of quasi-fixed factors determine the shape of the implicit cost function of the "cost function approach" and the adjustment costs of the "discrete choice approach". The MNL framework is well suited if it is more profitable for farmers to adapt their land allocation choices to their available quasi-fixed input quantities rather than to

adapt their variable input uses at the crop level. The agricultural scientists and the extension agents consulted by the authors usually assert that farmers are more reluctant to change their cropping practices than their land allocation, at least in the short run and within standard rotation patterns. The independence assumption of the quasi-fixed input requirements with respect to variable input uses can only hold “locally” (and approximately), *i.e.* for variable input uses in the neighborhood of the current use levels. This limits the applicability of MNL acreage share models to short run decisions, *i.e.* decisions only involving moderate changes in variable input use levels or decisions not involving technological changes. Under these conditions the MNL acreage choice models can be interpreted as “local” approximations of the “true” choice process of the farmers.

According to the usual structural interpretation of the multicrop technology, the MNL framework also imposes non-jointness restrictions of the multicrop technology in variable inputs, in outputs and in acreages. Non-jointness in variable inputs and in outputs is commonly assumed while non-jointness in acreages is more debated (see, *e.g.*, Just, Zilberman and Hochman 1983; Just and Chambers 1988; Asunka and Shumway 1996). However discussing these assumptions especially makes sense if the considered models are employed for investigating the properties of multicrop technology or for investigating drastic changes of the economic environment. The MNL acreage share models can not be used for these purposes. Much more flexible representations of the production technology are required and the agricultural production economics literature provides numerous examples of models much more relevant in this context (see, *e.g.*, Just and Pope 2001).

2.3 Acreage decisions within MNL framework: the cost function approach

In this section, it is shown that MNL acreage shares can be derived from a profit maximization program defined at the farm level. The presented framework considers a riskless environment but it is easily extended to cases with price and/or production risk as far as farmers are assumed to be risk neutral.

The typical short run problem faced by a farmer is to allocate his land to C different crops according to the acreage shares s_c for $c=1,\dots,C$ with $\sum_{b=1}^C s_c = 1$. Crop c output is sold at price p_c . Variable inputs are bought at prices \mathbf{w} . These prices are assumed to be known for simplicity. Each crop production technology is represented by a yield function of the form:

$$(2) \quad y_c = f_c(\mathbf{x}_c)$$

where y_c is the yield of crop c , \mathbf{x}_c is the quantity vector of variable input uses per unit of land of crop c and $f_c(\cdot)$ is assumed to be nondecreasing and concave in \mathbf{x}_c . As is discussed below, \mathbf{x}_c only include fertilizers, pesticides and seeds. The yield functions only depend on variable inputs and thus mostly represent the biological crop production process. Land and variable inputs such as fertilizers, seeds and pesticides are directly involved in the crop growth and development processes. The other quasi-fixed inputs (mainly machinery and labour) and the other variable inputs (mainly energy) are used for the variable input applications, for harvesting or for the soil preparation. Also the availability of quasi-fixed inputs mostly plays an indirect role in the biological crop production process. The main benefit of this framework is that the yield functions $f_c(\cdot)$ are similar to the ones considered by agricultural scientists.

Farmers' short run profit or gross margin function per hectare of any crop c is defined by:

$$(3) \quad \pi_c(p_c, \mathbf{w}) = \text{Max}_{y_c, \mathbf{x}_c} [p_c y_c - \mathbf{x}'_c \mathbf{w}] \quad \text{s.t.} \quad y_c = f_c(\mathbf{x}_c), y_c \geq 0, \mathbf{x}_c \geq \mathbf{0}$$

where \mathbf{s} is the vector of acreage shares. This model describes short run production choices and considers that the variable input choices vary within the range of variable input choices defined by the cropping practices used at the time the decisions are made. Moderate changes in cropping practices, *e.g.* in fertilizer use or pesticide use levels, do not change the short run production technologies, *i.e.* the $f_c(\cdot)$ functions, and only slightly modify the requirements for quasi-fixed inputs service. Whereas drastic changes in cropping practices involve long term choices: adoption of new cropping practices involving changes in the yield functions $f_c(\cdot)$, or/and adaptation of the quasi-fixed input quantities.

The farmers' restricted profit function explicitly defines a trade-off between the crop gross margins $\pi_c(p_c, \mathbf{w})$ of the different crops on the one hand and the "implicit management cost" of the chosen allocation $C(\mathbf{s})$ on the other hand:

$$(4) \quad \Pi(\mathbf{s}; \mathbf{p}, \mathbf{w}) = \sum_{c=1}^C s_c \pi_c(p_c, \mathbf{w}) - C(\mathbf{s})$$

where \mathbf{p} is the output price vector. The cost function $C(\mathbf{s})$, which defines the motive for crop diversification "concentrates" the non-linear effects of \mathbf{s} in the farmers' restricted profit function. It can be interpreted as a reduced form function smoothly approximating i) the unobserved variable costs associated with a given acreage (energy costs, ...) and ii) the effects of binding constraints on acreage choices, *e.g.* agronomic constraints or constraints associated to limiting quantities of quasi-fixed inputs. Quasi-fixed inputs such as labor and machinery are limiting in the sense that their cost per unit of land devoted to a given crop is likely to increase due to work peak loads or due to machinery overuse, whether machinery is specific

or not. Some crop rotations are impossible due to inconsistencies in planting and harvesting dates. Crop rotations may also be strongly unwarranted due to dramatic expected pest damages. These crop rotations are thus almost “forbidden” because their opportunity cost is very large in standard price ranges. These impossible and “forbidden” crop rotations determine the bounds imposed to acreage choices in (P)MP models. The implicit cost function $C(\mathbf{s})$ is assumed to be nondecreasing and quasi-convex in \mathbf{s} to reflect the constraints due to the limiting quantities of quasi-fixed factors (other than land) and due to the implicit bounds imposed on the acreage choices due to impossible or “forbidden” crop rotations. It is also assumed to be decreasing in the available quantities of quasi-fixed inputs (other than land).

Restricted profit functions, similar to the one defined in (4), are used in the PMP literature (Howitt 1995; Paris and Howitt 1998).⁴ Heckelevi and Wolff (2003) also propose to use this form of restricted profit function to define multi-crop econometric models with land as an allocable fixed factor. The main differences between the cost function used here, $C(\mathbf{s})$, and the ones used in the PMP literature are that i) $C(\mathbf{s})$ includes the effects of all binding constraints on acreage choices and ii) $C(\mathbf{s})$ is defined with (cross-entropy) measures of \mathbf{s} whereas the PMP implicit cost functions are usually quadratic in \mathbf{s} . Such implicit cost functions are also considered in dynamic models to account for adjustment costs, see *e.g.* Oude Lansink and Stefanou (2001).

The above discussion highlights the main features and limitations of the basic assumption set required to derive the standard MNL acreage share model from a restricted profit function. The main feature of this required assumption set is that the farmers short run production choices can be defined as the results of two optimization programs. First, farmers choose the

⁴ In most PMP applications, the cost function is designed to measure the total variable cost of \mathbf{s} . In this article, the observed variable input costs are part of the gross margins.

optimal objective yield and input uses for each crop by solving the programs in (3). Second, they choose the optimal acreages by maximizing $\Pi(\mathbf{s}; \mathbf{p}, \mathbf{w})$ in \mathbf{s} subject to the land use constraint $\sum_{c=1}^C s_c = 1$.

Building on the work of Anderson and *al.* (1992) it can be shown that the maximization of $\Pi(\mathbf{s}; \mathbf{p}, \mathbf{w})$ in \mathbf{s} subject to the land use constraint $\sum_{c=1}^C s_c = 1$ leads to acreage share functions with the MNL functional form defined by (1) if the cost function $C(\mathbf{s})$ has the following form:⁵

$$(5a) \quad C(\mathbf{s}) = A + \sum_{c=1}^C s_c c_c + a^{-1} \sum_{c=1}^C s_c \ln s_c .$$

The term A is an unidentifiable fixed cost. The c_c parameters are fixed costs (in the short run) per unit of land devoted to crop c . The term $\sum_{c=1}^C s_c \ln s_c$ is the opposite of the entropy function of the acreage share vector \mathbf{s} . Given that the acreage shares strictly lie between 0 and 1 and sum to 1 this term is negative. It is also strictly increasing and strictly convex in \mathbf{s} . The term $\sum_{c=1}^C s_c \ln s_c$ is minimal at $s_c = C^{-1}$ for $c=1, \dots, C$ implying that A can be chosen to ensure that the cost function is positive. This implicit management cost function has a fairly simple interpretation with an alternative but equivalent specification. Let define the parameters $d_c = \exp(-ac_c) / \left[\sum_{m=1}^C \exp(-ac_m) \right]$ and $A^d = A - a^{-1} \ln \left[\sum_{m=1}^C \exp(-ac_m) \right]$ where $c=1, \dots, C$.

The implicit cost function can then be defined in the alternative functional form:

$$(5b) \quad C(\mathbf{s}) = A^d + a^{-1} \sum_{c=1}^C s_c (\ln s_c - \ln d_c) .$$

⁵ The formal proof is provided in Appendix A.

The term $\sum_{c=1}^C s_c (\ln s_c - \ln d_c)$ is the opposite of the cross-entropy function of the acreage share vector. Its minimum is achieved at $s_c = d_c$ for $c = 1, \dots, C$. The vector of the parameters d_c , \mathbf{d} , defines the acreage share vector for which the implicit management cost is minimum, *i.e.* the most suitable acreage for the farm according to the implicit management costs. These costs increase with the difference between \mathbf{s} and \mathbf{d} according to the distance defined by the opposite of the cross-entropy function. In the implicit cost function, the a parameter is assumed to be strictly positive. It defines the relative “weight” of the management costs in the restricted profit function. It can be shown that the farmer only grows the most profitable crop if a goes to infinity and he chooses the minimum cost acreage \mathbf{d} if a goes to 0. The parameter vector \mathbf{d} (or \mathbf{c}) and the parameter a depend on the available quasi-fixed factors quantities and the previous acreage choices.

Along with this simple interpretation, the MNL acreage share models have two other interesting properties. First, the congruent indirect profit function and Lagrange multiplier associated to the land constraint have simple closed forms which depend on the well-known log-sum term:

$$(6) \quad \Pi^*(\mathbf{p}, \mathbf{w}; a, \mathbf{c}) \equiv a^{-1} \ln \left[\sum_{c=1}^C \exp \left[a (\pi_c(p_c, \mathbf{w}) - c_c) \right] \right] - A \text{ and } \lambda^*(\mathbf{p}, \mathbf{w}; a, \mathbf{c}) \equiv a \left[\Pi^*(\mathbf{p}, \mathbf{w}; a, \mathbf{c}) + A \right].$$

Second, thanks to the so-called log-linear transformation, the MNL acreage share models can be defined as:⁶

$$(7a) \quad \ln s_c - \ln s_C = a \left[(\pi_c(p_c, \mathbf{w}) - \pi_C(p_C, \mathbf{w})) - (c_c - c_C) \right] + u_c$$

or:

⁶ Crop K is chosen as the reference crop without any loss of generality.

$$(7b) \quad \ln s_c - \ln s_C = a(\pi_c(p_c, \mathbf{w}) - \pi_C(p_C, \mathbf{w})) + \ln(d_c/d_C) + u_c$$

for $c=1, \dots, C-1$ and where the u_c terms are error terms. This specification of the acreage choice equation system uses crop C as the reference crop. It is linear in the crop profit function and, as a result, can easily be incorporated into multi-crop econometric models including yield supply and input demand functions as shown by the illustrative applications.

2.4 Acreage decisions within the MNL framework: the discrete choice approach

MNL acreage shares can also be defined as the result of plot by plot discrete decisions. The MNL acreage share model presented in this section is based on two main points: the aggregation of choices made at the plot level, along the lines of Caswell and Zilberman (1985), and the logic of partial adjustment of acreage choices.

The farmer is assumed to own P plots of equal size ($n=1, \dots, P$) and to decide which of the C crops to grow on each plot. It is assumed at this point that farmers' decisions depend only on the expected gross margins $\pi_{cn}(p_c, \mathbf{w})$ and on the implicit "management" costs c_{cn} for crop c on plot n . Under the farmers' risk neutrality assumption, these expected gross margins can formally be defined as in the preceding section. The farmer's expected short run profit of growing crop c on plot n is given by:

$$(8) \quad \pi_{cn}(p_c, \mathbf{w}) - c_{cn} = \pi_c(p_c, \mathbf{w}) - c_c + e_{cn}$$

for $n=1, \dots, P$ and $c=1, \dots, C$. The term e_{cn} is known to the farmer but is random from the econometrician's point of view. Its expectation is normalized to be null. The plots are assumed to be sufficiently homogenous for considering the expected profit of growing crop c

to be constant across plots, it is then given by $\pi_c(p_c, \mathbf{w}) - c_c$, and for assuming that e_{cn} terms are identically and independently distributed across plots and crops. The probability (as it is perceived by the econometrician) that the farmer chooses crop c for plot n has a standard multinomial logit form:

$$(9) \quad P_{cn}(\mathbf{p}, \mathbf{w}; \mathbf{c}, \sigma) = \frac{\exp[\sigma^{-1}(\pi_c(p_c, \mathbf{w}) - c_c)]}{\sum_{m=1}^C \exp[\sigma^{-1}(\pi_m(p_m, \mathbf{w}) - c_m)]} = P_c(\mathbf{p}, \mathbf{w}; \mathbf{c}, \sigma),$$

if the $\sigma^{-1}e_{cn} + e$ terms have a standard Extreme Value distribution, where e is the Euler constant, and σ is a scale parameter of the variance of the e_{cn} terms. The assumptions stating that σ does not depend on (\mathbf{p}, \mathbf{w}) and that the e_{cn} terms are identically and independently distributed across plots and crops is consistent with the assumption that e_{cn} is part of c_{cn} , *i.e.* of the implicit cost of growing crop c on plot n . According to this interpretation the e_{cn} terms represent the effects of the plots' characteristics (topography, spatial distribution, ...). The homoskedasticity assumption related to the e_{cn} error terms can be relaxed, *e.g.*, to account for heterogeneity across farms of the expected gross margins $\pi_{cn}(p_c, \mathbf{w})$ due to differences in expected input uses or in expected yields. This would however result in more involved econometric models, at least in some cases.

If farmer's choices of crops were independent across plots, the expected (as it is perceived by the econometrician) share of plots allocated to crop k would be given by $P_c(\mathbf{p}, \mathbf{w}; \mathbf{c}, \sigma)$. Indeed the $P_c(\mathbf{p}, \mathbf{w}; \mathbf{c}, \sigma)$ terms define the "ideal" choice of the farmer, *i.e.* the acreage shares the farmer would choose if he was not constrained in his acreage choices. In this sense the $P_c(\mathbf{p}, \mathbf{w}; \mathbf{c}, \sigma)$ terms describe a long term (*i.e.* with optimal fixed factors quantities) choice of acreage shares. According to this logic and assuming that the farm is close to an equilibrium path, the farmer's dynamic optimal choice of acreage shares can be approximated by a simple

partial adjustment model (Treadway 1971; Considine and Mount 1984). A similar logic was used by Oude Lansink and Stefanou (2001). Denoting by s_c the share of land devoted to crop c and by $s_{c,(-1)}$ its counterpart for the preceding year, the resulting partial adjustment model is given by:

$$(10) \quad \ln s_c - \ln s_{c,(-1)} = r \left(\ln P_c(\mathbf{p}, \mathbf{w}; \mathbf{c}, \boldsymbol{\sigma}) - \ln s_{c,(-1)} \right) + \varepsilon_c, \text{ for } k = 1, \dots, K$$

where ε_{ct} is an error term including the approximation error due to the use of the simple adjustment model as well as the error term due to the use of $P_c(\mathbf{p}, \mathbf{w}; \mathbf{c}, \boldsymbol{\sigma})$ in place of the true long term choice of the farmer. Parameter r is the coefficient measuring the “friction” effects due to adjustment costs. It lies between 0 and 1 and can be defined in the empirical model as a function of available quasi-fixed factor quantities.

Differentiation of equation (9) for crop c and the reference crop C leads to the following equation:

$$(11) \quad \ln(s_c s_C^{-1}) = r \boldsymbol{\sigma}^{-1} (\pi_c(p_c, \mathbf{w}) - \pi_C(p_C, \mathbf{w})) - r \boldsymbol{\sigma}^{-1} (c_c - c_C) + (1-r) \ln(s_{c,(-1)} s_{C,(-1)}^{-1}) + \varepsilon_c - \varepsilon_C$$

for $c = 1, \dots, C-1$. Equation (11) is close to equation (7b) as could be expected: the “cost function approach” and the “discrete choice approach” with partial adjustment both rely on implicit management costs of the acreages. According to the assumption that farms are close to an equilibrium path, the lagged acreage shares in equation (11) are acreage shares with low management costs. Parameter r defines the weight of the “target” or ideal acreage choices relative to the adjustment costs. The partial adjustment MNL acreage share model can also be used within a production choice equations system. An empirical illustration is presented in the next section.

2.5 Empirical illustration

This section presents two simple applications of the modeling frameworks presented in the second and third sections, *i.e.* the “cost function” MNL acreage share model and the partial adjustment “discrete choice” MNL acreage share model.

2.5.1 Data

The data base is a rotating panel data sample (3 years *per* farm on average) of 5986 observations of French grain crop producers over the years 1989 to 2006, obtained from the Farm Accountancy Data Network (FADN). It provides detailed information on crop production for each farm i and year t ($t = B_i, \dots, T_i$): acreage s_{kit} , yield y_{kit} and price at the farm gate p_{kit} for each crop k . They only provide aggregate data on variable input (pesticides, fertilizers, seeds) expenditures whereas input price indices are made available at the regional level. Variable input quantities are aggregated into a single variable input for simplicity. X_{it} denotes the per hectare quantity of input purchased by farm i during year t and w_{it} denotes the corresponding price index. Total land area is used to control for scale effects in the presented empirical models. Acreage choices of three crops are considered: wheat, other cereals (mainly barley and corn) and, oilseeds (mainly rapeseed) and protein crops (mainly peas). Root crops (sugar beets and potatoes) acreages were considered as exogenous due to the sugar beet quota system implemented in the UE and because most of the potato acreages are defined by contracts. Fodder crop acreage (mainly silage corn) was also considered as exogenous due to feeding constraints.

2.5.2 Multi-crop econometric models

The quadratic functional form is chosen for the yield functions for three reasons. First its congruent dual functions have simple functional forms. Second, the quadratic production function can be parameterized in a form which is fairly easy to interpret by agricultural scientists or extension agents. Third, the resulting yield supply, input demand and (indirect)

gross margin functions can be generalized to account for farms and farmers unobserved heterogeneity and for production stochastic events in a “natural” way, *i.e.* by introducing additive random terms with simple interpretations. Pope and Just (2003) used this parameterization of the quadratic production function for this reason, albeit in a different context. Yield functions are defined as:

$$(12a) \quad y_{cit} = \alpha_{cit}(\mathbf{a}_c^\alpha) - \frac{1}{2} \gamma_{cit}^{-1}(\mathbf{a}_c^\gamma) [\beta_{cit}(\mathbf{a}_c^\beta) - x_{cit}]^2$$

$$(12b) \quad \text{with: } \alpha_{cit}(\mathbf{a}_c^\alpha) \equiv \alpha(\mathbf{z}_{it}; \mathbf{a}_c^\alpha) + \varepsilon_{cit}^y, \quad \beta_{cit}(\mathbf{a}_c^\beta) \equiv \beta(\mathbf{z}_{it}; \mathbf{a}_c^\beta) + \varepsilon_{cit}^x \quad \text{and} \quad \gamma_{cit}(\mathbf{a}_c^\gamma) \equiv \gamma(\mathbf{z}_{it}; \mathbf{a}_c^\gamma)$$

for $c=1, \dots, C$ where x_{cit} is the quantity of variable input used per hectare devoted to crop c by farmer i at t .⁷ The terms $\mathbf{a}_c \equiv (\mathbf{a}_c^\alpha, \mathbf{a}_c^\beta, \mathbf{a}_c^\gamma)$ are parameter vectors to be estimated, the vector \mathbf{z}_{it} contain variables used to control for farm heterogeneity, variations in production levels over time and technological changes, and $\alpha(\cdot)$, $\beta(\cdot)$ and $\gamma(\cdot)$ are known functions. The ε_{cit}^y and ε_{cit}^x terms are random terms representing farms unobserved heterogeneity and the effects on production of stochastic events such as climatic conditions or pest infestations. In this primal framework, the $\alpha_{cit}(\mathbf{a}_c^\alpha)$ and $\beta_{cit}(\mathbf{a}_c^\beta)$ terms have direct interpretations: $\beta_{cit}(\mathbf{a}_c^\beta)$ is the variable input quantity required to achieve the maximum yield $\alpha_{cit}(\mathbf{a}_c^\alpha)$. Both terms need to be positive. The $\gamma_{cit}(\mathbf{a}_c^\gamma)$ term determines the curvature of the yield function and, as a result, determines the magnitude of the price effects. They need to be positive for the yield function to be concave. These parameters have direct “agronomic” interpretations allowing the results to be “checked” with agricultural scientists and extension agents.

⁷ Generalization of the yield function to the multiple input case is straightforward.

Farmers' acreage choices are based on the expected crop gross margins. Prices are assumed to be known by farmers in this illustrative application. Maximization of the expected gross margin for each crop leads to gross margin functions of the form:

$$(13) \quad \pi_{cit}(\mathbf{a}_c) + e_{cit}^{\pi} \equiv \left[p_{cit} \alpha_{cit}(\mathbf{a}_c^y) - w_{it} \beta_{cit}(\mathbf{a}_c^{\beta}) + p_{cit} \frac{1}{2} \gamma_{cit}(\mathbf{a}_c^{\gamma}) (w_{it}^2 / p_{cit}) \right] + e_{cit}^{\pi}$$

where $e_{cit}^{\pi} \equiv p_{cit}(e_{cit}^y - e_{Cit}^y) - w_{it}(e_{cit}^x - e_{Cit}^x)$. e_{cit}^y and e_{cit}^x are farmers' expectations of the ε_{cit}^y and ε_{cit}^x terms at the time they choose their acreages.⁸ The nice feature of the quadratic yield functions in this framework is that their congruent yield supply, input demand and gross margin functions have additive error terms with simple interpretations. The e_{cit}^y and e_{cit}^x error terms account for the unobserved (not controlled by \mathbf{z}_{it}) heterogeneity across farms and time in the yield functions. These terms are known to the farmers when they choose their acreages. Their mean is normalized to 0. The $\varepsilon_{cit}^y - e_{cit}^y$ and $\varepsilon_{cit}^x - e_{cit}^x$ terms account for the stochastic events affecting production levels. These terms are unknown to the farmers when they choose their acreages, but there are known to them when they decide their variable input quantities. These error terms are unknown to the econometrician and their expectation is normalized to 0.

The considered multi-crop econometric models are defined by equation systems composed of C yield supply functions:

$$(14) \quad y_{cit} = \alpha_{cit}(\mathbf{a}_c^y) - \frac{1}{2} \gamma_{cit}(\mathbf{a}_c^{\gamma}) (w_{it} / p_{cit})^2 + \varepsilon_{cit}^y, \quad c = 1, \dots, C,$$

⁸ The (per hectare of grain crop) compensatory payments provided by the CAP are added to this gross margin functions in the estimated multi-crop econometric models.

an aggregate input demand equation:⁹

$$(15) \quad X_{it} = \sum_{c=1}^{C+L} s_{cit} x_{cit}^e \left(\beta_{cit}(\mathbf{a}_c^\beta) - \gamma_{cit}(\mathbf{a}_c^\gamma) (w_{it}/p_{cit}) \right) + \left(\sum_{c=1}^C s_{cit} \varepsilon_{cit}^x + \varepsilon_{it}^X \right)$$

and $C-1$ acreage equations. These equations are defined, for the “cost function” MNL acreage share model, by:

$$(16) \quad \ln(s_{cit} s_{Cit}^{-1}) = a_{it}(\mathbf{b}) \left(\pi_{cit}(\mathbf{a}_c) - \pi_{Cit}(\mathbf{a}_C) - c_{cit}(\mathbf{g}_c) \right) + \left(e_{cit}^\pi + e_{cit}^c \right), \quad c = 1, \dots, C-1,$$

and, for the “discrete choice” MNL acreage share model with partial adjustment, by:

$$(17) \quad \ln(s_{cit} s_{Cit}^{-1}) = r \sigma^{-1} \left(\pi_{cit}(\mathbf{a}_c) - \pi_{Cit}(\mathbf{a}_C) - c_c \right) + (1-r) \ln(s_{ci,t-1} s_{Ci,t-1}^{-1}) + \left(e_{cit}^\pi + e_{cit}^c \right), \quad c = 1, \dots, C-1.$$

The terms \mathbf{b} , \mathbf{g}_c , c_c and r are parameters to be estimated ($c = 1, \dots, C-1$). The fixed cost per hectare of crop c is defined by $c_{cit}(\mathbf{g}_c) \equiv c(\mathbf{z}_{it}; \mathbf{g}_c) + e_{cit}^c$ in equations (16) and by $c_c + e_{cit}^c$ in equations (17), with the normalization constraints $c_{Cit} \equiv 0$. The error terms e_{cit}^c are known to the farmers but unknown to the econometrician. Their expectation is normalized to 0. The error term added to the input demand equation (14), ε_{it}^X , account for measurement errors due to stock variations.

The \mathbf{z}_{it} variable vector contains control variables. Quadratic time trends were introduced in the yield functions to account for (disembodied) technical changes. A “production potential index”, $q_{it} \equiv (y_{li,t-1} - y_{li,t-1}^{Med}) / (y_{li,t-1}^{Max} - y_{li,t-1}^{Min})$, is created to control for farm heterogeneity. $y_{li,t-1}^{Med}$,

⁹ In the application the input uses for root crops and fodder crops are added in the input use equations ($k = K+1, \dots, L$). The corresponding input uses are defined as linear functions of the \mathbf{z}_{it} variable vector defined below (cubic time trends, quadratic effects of the production potential index ...).

$y_{li,t-1}^{Max}$ and $y_{li,t-1}^{Min}$ denote, respectively, the median, 99% quantile and 1% quantile of the yield of wheat in the sample in year $t-1$. It is based on wheat yields due to the specialization of the sampled farms, and it is defined on a year by year basis to control for year specific conditions. Quadratic effects of q_{it} are introduced in the parameters of the yield functions. The specified effects of q_{it} can be interpreted as control functions of the farms' heterogeneity. This control function approach is analogous to Chamberlain's (1982) Π matrix approach and to Mundlak's (1978) device for controlling for the so-called individual fixed effects in panel data econometrics. While this index mostly accounts for persistent production conditions, farmers' choices and yields also depend on crop rotation effects. The lagged acreage shares of root crops are introduced in the cereal yield functions to account for the beneficial effects of the induced crop rotations.

2.5.3 Estimation issues

The control variable vector \mathbf{z}_{it} and the price variables are exogenous with respect to the error terms of the econometric models. The $\alpha_{cit}(\mathbf{a}_c^y)$, $\gamma_{cit}(\mathbf{a}_c^y)$, $\beta_{cit}(\mathbf{a}_c^\beta)$ and $c_{cit}(\mathbf{g}_c)$ terms are defined as linear functions of \mathbf{z}_{it} and are linear in their respective parameters. The $a_{it}(\mathbf{b})$ term is defined as the exponential of a linear function of \mathbf{z}_{it} and \mathbf{b} . The acreage shares are potentially endogenous in the input demand equation. The ε_{cit}^x error terms contain the heterogeneity effects e_{cit}^x which partly determine the acreage choices. In this illustration it is assumed that the heterogeneity control ensured by q_{it} in the crop input demand functions is sufficient to neglect the effects of the e_{cit}^x terms (the q_{it} index is defined for that purpose). Albeit it is standard (see, *e.g.*, Hornbaker *et al* 1989), this assumption is admittedly restrictive. Note however that this estimating equation is only needed for identifying \mathbf{a}_C^β . The yield supply and acreage share equations identify the whole set of parameters, \mathbf{a}^β excluded. The acreage share equations also identify $\mathbf{a}_c^\beta - \mathbf{a}_C^\beta$ for $c = 1, \dots, C-1$.

The parameter estimators are constructed within the Generalized Method of Moment (GMM) framework.¹⁰ They are based on the orthogonality conditions defined by the vector of the cross product of the (composite) error term of each equation with each of their explanatory variables. The resulting GMM estimator is robust to heteroskedasticity of unknown form and does not exclude correlation of the error terms across equations.

2.5.4 Main results

Table 1 presents the estimates of yield supply, input demand and acreage shares function parameters for the “cost function” and “discrete choice” models. Table 2 presents the average price elasticities of the crop supply, input demand and acreage share functions.

Both models yield similar results with respect to the input demand and yield supply function parameters. The fit of the models to these micro-level data is correct. The R^2 criteria lie between .32 and .43. Estimates of the maximum yield and input requirements for maximum yield, *i.e.* the $\alpha_{cit}(\mathbf{a}_c^y)$ and $\beta_{cit}(\mathbf{a}_c^\beta)$ terms, are in the ranges expected by the agricultural scientists and extension agents the authors have consulted. As expected, the production potential index has positive effects on the the $\alpha_{cit}(\mathbf{a}_c^y)$ and $\beta_{cit}(\mathbf{a}_c^\beta)$ terms. Past acreages of root crops have a positive effect on wheat yield and a negative effect on the demand of wheat variable inputs. These effects are consistent with the known beneficial effects of root crops at the beginning of the crop rotation sequence. Estimated average price elasticities of the yield supply and crop input demands are reported in table 2. They lie in standard ranges, albeit the price responsiveness of the “other cereals” functions is surprisingly low. These results demonstrate that both multi-crop models provide satisfactory econometric modeling

¹⁰ The econometric models are not standard Seemingly Unrelated Regression systems despite that they are composed of regression equations only. The acreage shares are the dependent variables of the acreage equations whereas they are independent variables in the input demand equation.

frameworks: they yield sensible estimated price effects and expected heterogeneity control variable effects.

Estimates of the acreage equations lead to more contrasted conclusions according to the comparison of the “cost function” and “discrete choice” MNL acreage share models.

As expected, the “cost function” acreage share model show that farms with large root crop acreages devote also more land to cereals rather than to oilseeds and protein crops. The estimated $a_{it}(\mathbf{b})$ terms imply that the (expected) crop gross margin variations account for about 14% of variations of the differences in log acreage shares. As expected the estimated $a_{it}(\mathbf{b})$ terms are decreasing in past root crop acreages, indicating that these crops offer much flexibility for subsequent crop choices. The own price elasticities of the crop acreage shares presented in table 2 range from .57 to .70. These estimated average elasticities are close to each other because their values mostly depend on the estimated values of the single term $a_{it}(\mathbf{b})$. The Nested MNL acreage share model which is presented in the next section offers more flexibility in this respect. These results globally indicate that the “cost function” MNL acreage share model provides sensible results with respect to price effects and heterogeneity control variables effects. Nevertheless, the low R^2 criteria (.11 and .19) for these acreage equations call for improvement of the econometric “cost function” MNL acreage share model with respect to, *e.g.* the use of extra variables to better control for heterogeneity. This lack of fit may also be due to the CAP instruments implemented in the period covered by our data. Price supports to grain crops sharply declined in the EU during the nineties. But this decrease in price support has been compensated by direct payments (which are incorporated in the empirical models). These direct payments were defined for each grain crop for compensating producers’ gross margins at the *département* level (France is divided into 95 *départements*). As a result grain crop acreages have been “frozen” due to the implied negative correlation between the direct payments and crop gross margins.

The “discrete choice” acreage share model with partial adjustment model has a much better fit to the data. This is not surprising for a model basically predicting acreages in year t by acreages in year $t-1$. However, the estimated value of the “friction” parameter r is equal to .25 showing that acreages respond to short run economic incentives despite significant adjustment costs. This model allows to compute price elasticities of the crop acreage shares in the long run, *i.e.* without adjustment constraints, and in the short run, *i.e.* with limited adjustment possibilities. The own price average long run elasticities of the crop acreage shares are close to 1. The corresponding short run average elasticities are close to .25. As expected, the estimated price elasticities derived from the “cost function” model lie between the short run and long run elasticities derived from the “discrete choice” partial adjustment model.

It is also interesting to note that the *per* hectare “fixed costs” terms of both acreage share models tend to show that these models underestimate the oilseeds and protein crops acreages. A modified version of the model incorporating a crude measure of the beneficial effects of the oilseeds/protein crops-wheat rotation on future wheat gross margins provides a “correction” for this underestimation problem. This suggests that dynamic generalizations of the MNL acreage share models may provide significant improvements for acreage choice modeling. The next section briefly presents the basic framework for building “dynamic discrete choice” MNL acreage share models.

2.6 Generalized MNL acreage share models

The main aim of this section is to present a brief overview of the possible generalizations of the MNL framework for acreage choice modeling. Other generalizations are possible. For example, the Nested MNL acreage share model derived using the “cost function approach” can also be derived by using the “discrete choice approach”. To present these generalizations also allows to point out some drawbacks of the “standard” MNL acreage share models. Two generalizations of the standard multinomial logit acreage share models are presented.

2.6.1 "Cost function approach": the Nested multinomial logit model

The simplicity of the log-linear transformation used in equations (6) is mainly due to a specific feature of the MNL acreage shares. The ratio of the acreage shares of two different crops only depends on the payoffs of these crops. This "independence of the irrelevant crops" property also is a potential drawback of this simple acreage share model. The acreage share elasticities with respect to crop (expected) gross margins mainly depend on the single a parameter, *i.e.* the relative "weight" of the acreage management cost in the farmers' objective function. All crops are equivalently considered in terms of management costs in the MNL acreage share model. The "cost function" MNL acreage share framework can be generalized to account for similarities and differences in the management of the different crops, in the spirit of the PMP framework developed by Röhm and Dabbert (2003).

If the C crops can be allocated to Q mutually exclusive nests according to their management costs, it is possible to define the corresponding "Nested MNL acreage share models" and their corresponding indirect profit and indirect restricted profit functions. The set of crops belonging to nest ℓ ($\ell = 1, \dots, Q$) is denoted by $B(\ell)$, the share of land allocated to the crops of nest ℓ is denoted by $\bar{s}_\ell = \sum_{c \in B(\ell)} s_c$ and the share of crop c within its nest ℓ is denoted by $s_{c/\ell} = s_c / \bar{s}_\ell$. The price arguments are omitted to simply notations and obvious vector notations are used without definitions. Building on the work of Verboven (1996), it can be shown that the maximization in \mathbf{s} of the restricted indirect profit function:

$$(18) \quad \Pi(\mathbf{s}; \boldsymbol{\pi}, \mathbf{c}; a, \boldsymbol{\rho}) = \sum_{\ell=1}^Q \sum_{m \in B(\ell)} s_m (\pi_m - c_m) - a^{-1} \left[\sum_{\ell=1}^Q \bar{s}_\ell \ln \bar{s}_\ell + \sum_{\ell=1}^Q a \rho_\ell^{-1} \bar{s}_\ell \sum_{m \in B(\ell)} s_{m/\ell} \ln s_{m/\ell} \right]$$

subject to the total land allocation constraint leads to Nested MNL acreage share functions.^{11,12}

If crop c belongs to nest q , we have:

$$(19) \quad s_c(\boldsymbol{\pi}, \mathbf{c}; a, \boldsymbol{\rho}) = \frac{\exp[\rho_q(\boldsymbol{\pi}_c - c_c)]}{\sum_{m \in B(q)} \exp[\rho_q(\boldsymbol{\pi}_m - c_m)]} \frac{\left[\sum_{m \in B(q)} \exp[\rho_q(\boldsymbol{\pi}_m - c_m)] \right]^{a\rho_q^{-1}}}{\sum_{\ell=1}^Q \left[\sum_{m \in B(\ell)} \exp[\rho_\ell(\boldsymbol{\pi}_m - c_m)] \right]^{a\rho_\ell^{-1}}}$$

In the restricted profit function a^{-1} is the weight parameter of the management cost function for the different nests while ρ_ℓ^{-1} is the weight parameter of the management cost function for the crops of nest ℓ . Note that the MNL restricted indirect profit functions and acreage share functions are special cases of their Nested MNL counterparts. The former is obtained from the latter with $\rho_\ell = a$ for $\ell = 1, \dots, Q$. The first right hand side term of equation (18) defines the share function of crop c within its nest, $s_{c/q}(\boldsymbol{\pi}_q, \mathbf{c}_q; \rho_q)$. The second right hand side term of equation (18) defines the share of total land allocated to nest q , $\bar{s}_q(\boldsymbol{\pi}, \mathbf{c}; a, \boldsymbol{\rho})$. Note that $\bar{s}_q(\boldsymbol{\pi}, \mathbf{c}; a, \boldsymbol{\rho})$ can also be written as:

$$(20) \quad \bar{s}_q(\boldsymbol{\pi}, \mathbf{c}; a, \boldsymbol{\rho}) = \frac{\exp[a\Pi_q^*(\boldsymbol{\pi}_q, \boldsymbol{\rho}_q, \mathbf{c}_q)]}{\sum_{\ell=1}^Q \exp[a\Pi_\ell^*(\boldsymbol{\pi}_\ell, \boldsymbol{\rho}_\ell, \mathbf{c}_\ell)]} \text{ with } \Pi_\ell^*(\boldsymbol{\pi}_\ell, \boldsymbol{\rho}_\ell, \mathbf{c}_\ell) \equiv \rho_\ell^{-1} \ln \left[\sum_{m \in B(\ell)} \exp[\rho_\ell(\boldsymbol{\pi}_m - c_m)] \right].$$

The $\bar{s}_q(\boldsymbol{\pi}, \mathbf{c}; a, \boldsymbol{\rho})$ acreage share function is defined in a standard MNL form by using the indirect profit functions or inclusive value functions, $\Pi_\ell^*(\boldsymbol{\pi}_\ell, \boldsymbol{\rho}_\ell, \mathbf{c}_\ell)$, associated to the crops the different nests ℓ . The Nested MNL framework is less tractable than the standard multinomial logit framework since there is no simple counterpart to the log-transformation

¹¹ The formal proof is provided in Appendix B.

¹² The indirect restricted profit function (as well as its congruent functions) defined with the ‘‘reference acreage share vector’’ is not given here but can be readily be derived.

used with standard MNL models. However, in the particular cases where there is a single specific crop (an “outside” crop), the technique developed by Berry (1994) can be used to define empirically tractable estimating equations.¹³

2.6.2 “Discrete choice approach”: MNL acreage share models accounting for crop rotation effects

The “discrete choice” MNL acreage share model can be generalized to account for the fact that farmers consider the expected crop rotation effects of their acreage choices. It is assumed for simplicity that production dynamics is of order 1 and that farmers only consider anticipations with respect to the next year with a discount factor d . The results presented in this section heavily rely on Rust’s (1987) framework for discrete choice modeling. Let denote the profit of growing crop c on plot n where crop m was grown during the preceding year by:

$$(21) \quad \pi_{cn/m}(p_{ct}, \mathbf{w}_t) - c_{cn/m} = \pi_{c/m}(p_{ct}, \mathbf{w}_t) - c_{c/m} + e_{cnt}$$

where the functional form of the $\pi_{cn/m}(\cdot)$ functions is known and the e_{cnt} terms are identically and independently distributed across crops, plots and time, and have the distribution defined in section 3. In year t the e_{cnt} terms are known to the farmer but the $e_{cn,t+1}$ terms are not. It is however assumed that the farmer’s perceived distribution of the $e_{cn,t+1}$ terms is also the distribution described in section 3. According to this model, if the prices $(\mathbf{p}_{t+1}, \mathbf{w}_{t+1})$ are known to the farmer (and to the econometrician) then he knows that if he chooses crop c for plot n in t , the probability of his choosing crop ℓ in year $t+1$ on the same plot is given by:

¹³ The Nested MNL version of the econometric model provides interesting results which are similar to those presented in section 4.

$$(22) \quad P_{\ell/c}(\mathbf{p}_{t+1}, \mathbf{w}_{t+1}) = \frac{\exp\left[\sigma^{-1}\left(\pi_{\ell/c}(p_{\ell,t+1}, \mathbf{w}_{t+1}) - c_{\ell/c}\right)\right]}{\sum_{m=1}^C \exp\left[\sigma^{-1}\left(\pi_{m/c}(p_{m,t+1}, \mathbf{w}_{t+1}) - c_{m/c}\right)\right]}.$$

As a result his expected pay-off on plot n in year $t+1$ (as perceived in year t) is given by:

$$(23) \quad E\left[\pi_{\ell/c}(\mathbf{p}_{t+1}, \mathbf{w}_{t+1}) - c_{\ell/c}\right] = \sigma \ln \left[\sum_{\ell=1}^C \exp\left[\sigma^{-1}\left(\pi_{\ell/c}(p_{\ell,t+1}, \mathbf{w}_{t+1}) - c_{\ell/c}\right)\right] \right]$$

i.e. the expected profit has the well-known log-sum form. Thus, in year t the (risk neutral) farmer (who has a perfect foresight on $t+1$ prices) chooses the crop on plot n according to the expected pay-offs given by:

$$(24) \quad \pi_{c/m}(p_{ct}, \mathbf{w}_t) - c_{c/m} + e_{cnt} + dE\left[\pi_{\ell/c}(\mathbf{p}_{t+1}, \mathbf{w}_{t+1}) - c_{\ell/c}\right]$$

From the econometrician point of view, the probability that the farmer chooses c has the standard MNL functional form:

$$(25) \quad P_{c/m}(\mathbf{p}_t, \mathbf{w}_t; \mathbf{p}_{t+1}, \mathbf{w}_{t+1}) = \frac{\exp\left[\sigma^{-1}\left(\pi_{c/m}(p_{ct}, \mathbf{w}_t) - c_{c/m} + dE\left[\pi_{\ell/c}(\mathbf{p}_{t+1}, \mathbf{w}_{t+1}) - c_{\ell/c}\right]\right)\right]}{\sum_{q=1}^C \exp\left[\sigma^{-1}\left(\pi_{q/m}(p_{qt}, \mathbf{w}_t) + dE\left[\pi_{\ell/q}(\mathbf{p}_{t+1}, \mathbf{w}_{t+1}) - c_{\ell/q}\right]\right)\right]}.$$

The closed form of the expected pay-off in year $t+1$ permits further generalizations. For example, uncertainty about prices in $t+1$ can be handled using integration of the expectation of the probability function (24) according to the assumed distribution of prices. Simulation methods that are now widely used can be employed for that purpose (see, *e.g.*, Train, 2003). These probability functions can also be used in the partial adjustment framework defined in the third section. Note however that the resulting empirical models remain close to the ones presented in this article only in the case where the crop rotations are observed. The resulting

empirical models are more complicated where only acreages are observed. In this case the probability of choosing crop c at t is given by $P_c(\mathbf{p}_t, \mathbf{w}_t; \mathbf{p}_{t+1}, \mathbf{w}_{t+1}) = \sum_{m=1}^C s_{m,t-1} P_{c/m}(\mathbf{p}_t, \mathbf{w}_t; \mathbf{p}_{t+1}, \mathbf{w}_{t+1})$ and the difference in the log-acreage shares does not provide any simplification. Nevertheless these crop choice probabilities may be simplified thanks to similarities of the rotation effects of certain crop sequences.

The agronomic constraints considered in the “cost function approach” differ from the crop rotation effects considered here. In the present framework, crop rotation effects generate inter-temporal trade-offs while the agronomic constraints considered in the “cost function approach” restrict acreage choices. In this respect, the “cost function approach” is only suitable where farmers’ can be assumed to use restrictive rotation patterns whereas the “discrete choice approach” can be used as a modelling framework in broader situations.

Concluding remarks

Two approaches are presented that provide theoretical backgrounds for using MNL acreage share models: the “cost function” and the “discrete choice” approaches. The “discrete choice” based approach remains mainly empirical. This approach focuses on some of the farmers’ decision parameters but it either ignores or only uses reduced form effects for the other farmers’ decision parameters.¹⁴ It exploits the flexibility of the MNL framework to focus on some determinants of the acreage choices, *e.g.* crop gross margins, acreage management costs or crop rotation effects. It ignores other determinants, *e.g.* risk spreading. The relevance of this choice depends on the context and needs to be empirically evaluated.

¹⁴ Even though these reduced form effects can also be theoretically grounded.

The “cost function” based approach appears to be more “structural” in the sense that it is based on profit functions. However, as it is the case for any simple theoretical model of production choices, the MNL acreage share models are to be used with caution. Just and Pope (2001) convincingly argue that any econometric model of farmers’ choices necessarily contains reduced form effects because, among others, of the complexity of agricultural production processes, of the limitations of the usual data sets, of the complexity of the farmers’ objective functions, ... The MNL acreage share models can be interpreted in two ways: either as a structural model relying on restrictive assumptions with respect to the underlying technology, or as a model approximating the “true” model. This second interpretation is preferred in this article. Introduction of appropriate control variables in empirical MNL acreage share models allows to define simple econometric models to be interpreted as local approximations of the “true” models and to be used to investigate the effects of moderate changes in the production context.

Both approaches accommodate generalizations of the standard MNL acreage share model. But the “discrete choice” MNL framework seems more flexible than its “cost function” counterpart. Accounting for crop rotation effects and for dynamic optimization by farmers appears to be a promising direction for further research as shown by the applications presented in this article as well as the results obtained by Livingston *et al* (2008). These generalizations can benefit from the rapidly expanding literature on dynamic discrete choice econometric models.

Despite their limitations but thanks to their simple structure, the MNL acreage share models appear to be useful tools for investigating farmers’ short run production decisions. They can be used to produce simple comparative statics results. They can also be used to build simple and reliable multi-crop econometric models as shown by the illustration presented in this article. Economists involved in multi-disciplinary research projects may also find it useful for

defining production choice models which are likely to be preferred to the standard multi-crop dual models by non-economists thanks to the immediate interpretation of their parameters. The MNL acreage models also share another advantage with MP models: thanks to their simple structure they can easily be used for investigating the effects of new cropping practices on land allocation. Finally, these models can also be used by researchers as simple acreage choice models in more elaborated econometric models of production choice models.¹⁵ One of the main drawbacks of the MNL framework is that it rules out corner solutions in acreage shares. However, this certainly calls for original approaches for corner solution modeling.

¹⁵ The MNL framework is the workhorse of the recent empirical industrial organization literature because it allows to define empirically tractable econometric demand functions which can be employed in various market equilibrium models (Ackerberg *et al.* 2007).

References

- Akerberg, D., L. Benkard, S. Berry and A. Pakes. 2007. "Econometric Tools for Analyzing Market Outcomes." In *Handbook of Econometrics 6A*, Heckman J. J. and E. E. Leamer (Eds). Amsterdam, Elsevier.
- Anderson, S.P., A. de Palma, and J. Thisse. 1992. *Discrete Choice Theory of Product Differentiation*. Cambridge, Mass. MIT Press.
- Asunka, S., and C. R. Shumway. 1996. "Allocatable Fixed Inputs and Jointness in Agricultural Production: More Implications." *Agricultural and Resource Economics Review* 25:143-48.
- Berry, S. T. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *The RAND Journal of Economics* 25(2):242-262.
- Bewley, R., T. Young, and D. Colman. 1987. "A System Approach to Modelling Supply Equations in Agriculture." *Journal of Agricultural Economics* 38(2):151-166.
- Caswell, M., and D. Zilberman. 1985. "The Choice of Irrigation Technologies in California." *American Journal of Agricultural Economics* 67(2):224-234.
- Chamberlain, G. 1982. "Multivariate Regression Models for Panel Data." *Journal of Econometrics* 18:5-46.
- Chambers, R., and R. Just. 1989. "Estimating Multioutput Technologies." *American Journal of Agricultural Economics* 71(4):980-95.
- Chavas, J.-P. and T. Holt. 1990. "Acreage Decisions under Risk: The Case of Corn and Soybeans." *American Journal of Agricultural Economics* 72(3):529-538.
- Considine, T.J., and T.D. Mount. 1984. "The Use of Linear Logit Models for Dynamic Input Demand Systems." *Review of Economics and Statistics* 66(3):434-443.
- Coyle, B. T. 1992. "Risk Aversion and Price Risk in Duality Models of Production: A Linear Mean-Variance Approach." *American Journal of Agricultural Economics* 74(4):849-859.
- Heckelei, T. and H. Wolff. 2003. "Estimation of Constrained Optimisation Models for Agricultural Supply Analysis Based on Generalised Maximum Entropy." *European Review of Agricultural Economics* 30(1):27-50.

- Hornbaker, R. H., B. L. Dixon and S. T. Sonka. 1989. "Estimating Production Activity Costs for Multioutput Firms with a Random Coefficient Regression Model." *American Journal of Agricultural Economics* 71(1):167-177.
- Howitt, E. 1995. "Positive Mathematical Programming." *American Journal of Agricultural Economics* 77(2):329-342.
- Just, R.E., and R.D. Pope. 2001. "The Agricultural Producer: Theory and Statistical Measurement." In B.L. Gardner and G.C. Rausser, ed. *Handbook of Agricultural Economics*, Vol. 1A. Amsterdam, Elsevier.
- Just, R.E., Zilberman, D., and E. Hochman. 1983. "Estimation of Multicrop Production Functions." *American Journal of Agricultural Economics* 65(4):770-780.
- Lichtenberg, E. 1989. "Land Quality, Irrigation Development and Cropping Pattern in the Northern High Plains." *American Journal of Agricultural Economics* 71(1):187-94.
- Livingston, M., M.J. Roberts, and J. Rust. 2008. "Optimal Corn and Soybean Rotations." Paper presented at AAEA annual meeting, Orlando, Florida, 27-29 July.
- Lubowski, R.N, A.J. Plantinga, R.N. Stavins. 2006. "Land-Use and Carbon Sinks: Econometric Estimation of the Carbon Sequestration Supply Function." *Journal of Environmental Economics and Management* 51(2):135-152.
- McFadden, D. 1974. "Conditional Logit Analysis of Quantitative Choice Behavior." In P. Zarembka, ed. *Frontiers in Econometrics*. New-York: Academic Press.
- Moore, M. R., et D. H. Negri. 1992. "A Multicrop Production Model of Irrigated Agriculture, Applied to Water Allocation Policy of the Bureau of Reclamation." *Journal of Agricultural and Resource Economics* 17:29-43.
- Mundlak, Y. 1978. "On the Pooling of Time Series and Cross Section Data." *Econometrica* 46(1):69-85.
- Oude Lansink, A. and J. Peerlings. 1996. "Farm-specific Impacts of Quantitative Restrictions of N-Fertiliser Use in Dutch Arable Farming." *Journal of Agricultural Economics* 52(3):38-52.

- Oude Lansink, A. and S. Stefanou. 2001. "Dynamic Area Allocation and Economies of Scale and Scope." *Journal of Agricultural Economics* 52(3):38-52.
- Ozarem, P.F., and J.A. Miranowski. 1994. "A Dynamic Model of Acreage Allocation with General and Crop-Specific Soil Capital." *American Journal of Agricultural Economics* 76(3):385-395.
- Paris, Q. and R.E. Howitt. 1998. "An Analysis of Ill-posed Production Problems using Maximum Entropy." *American Journal of Agricultural Economics* 80(1):124-138.
- Pope, R. D. and R. E. Just. 2003. "Distinguishing Errors in Measurement from Errors in Optimization." *American Journal of Agricultural Economics* 85(2):348-358.
- Röhm, O., and S. Dabbert. 2003. Integrating Agri-Environmental Programs into Regional Production Models: An Extension of Positive Mathematical Programming, *American Journal of Agricultural Economics* 85(1):254-265.
- Rust, J. 1987. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica* 55(5):999-1033.
- Theil, H. 1969. "A Multinomial Extension of the Linear Logit Model." *International Economic Review* 10(3): 251-259.
- Thomas, A. 2003. "A Dynamic Model of On-Farm Integrated Nitrogen Management." *European Review of Agricultural Economics* 30(4):439-460.
- Train, K. 2003. *Discrete Choice Methods with Simulation*. Cambridge University Press.
- Treadway, A.B. 1971. "On the Rational Multivariate Flexible Accelerator." *Econometrica* 39(5):845-855.
- Verboven, F. 1996. "The Nested Logit Model and Representative Consumer Theory." *Economics Letters* 50(1):57-63.
- Wu, J., and K. Segerson. 1995. "The Impact of Policies and Land Characteristics on Potential Groundwater Pollution in Wisconsin." *American Journal of Agricultural Economics* 77(4):1033-47.

Table 1. Estimates of the Yield, Input Demand and Acreage Shares Equations, 1989-2006.

Explanatory Variable	<i>“Cost function” model</i>			<i>“Discrete choice” model</i>		
	Wheat	Other cereals	Oilseeds protein crops	Wheat	Other cereals	Oilseeds protein crops
Yield supply						
Price effects (γ)	1.89***	1.34***	1.71***	1.56***	0.75***	1.83***
Production index	0.48*	0.19	3.06***	0.57*	-0.64**	2.83***
Average potential yield (α)	8.69***	8.28***	6.74***	8.56***	8.02***	6.79***
Constant	7.89***	8.36***	5.86***	7.98***	8.08***	6.02***
Trend	0.12***	0.04**	0.11***	0.09***	0.10***	0.09***
Trend square	-3 10 ⁻³ ***	2 10 ⁻³ **	-5 10 ⁻³ ***	-2 10 ⁻³ ***	-2 10 ⁻³ ***	-4 10 ⁻³ ***
Production index	2.68***	2.14***	2.75***	2.77***	1.94***	2.58***
Root crop acreage	2.42***	-	-	2.22***	-	-
<i>R-square</i>	<i>0.42</i>	<i>0.34</i>	<i>0.32</i>	<i>0.42</i>	<i>0.34</i>	<i>0.32</i>
Input demand						
Average optimal input use (β)	5.18***	5.51***	5.40***	4.80***	5.01***	5.25***
Constant	5.53***	5.67***	6.13***	5.63***	4.98***	5.71***
Production index	1.84***	-0.46	3.29***	1.34***	0.10	3.16***
Root crops		10.13***			11.61***	
Fodder crops		1.44			1.75	
<i>R-square</i>		<i>0.42</i>			<i>0.44</i>	
Acreage shares						
Fixed costs (c)	-	2.90***	1.50***	-	1.17***	1.28***
Production index	-	2.88***	-0.58	-	-	-
Root crops	-	23.16**	40.05**	-	-	-
Fodder crops	-	3.84***	5.98***			
Cost weight (a)			0.14***			-
Root crops			-2.85***			-
Friction parameter (r)			-			0.25***
Scale parameter (σ)			-			0.19***
<i>R-square</i>	-	<i>0.11</i>	<i>0.19</i>	-	<i>0.59</i>	<i>0.51</i>

Note: (*), (**) and (***) denote parameter estimates statistically different from 0 at, respectively, 10%, 5% and $\leq 1\%$ confidence levels.

Table 2. Estimated average price elasticities of yield supply, input demand and acreage shares

	<i>“Cost function” econometric model</i>				<i>“Discrete choice” econometric model</i>			
	Price				Price			
	Wheat	Other cereals	Oilseeds protein crops	Input	Wheat	Other cereals	Oilseeds protein crops	Input
Yield supply functions								
Wheat	0.178	-	-	-0.178	0.143	-	-	-0.143
Other cereals	-	0.156	-	-0.156	-	0.090	-	-0.090
Oilseeds, protein crops	-	-	0.234	-0.234	-	-	0.248	-0.248
Input demand functions								
Wheat	0.427	-	-	-0.427	0.353	-	-	-0.353
Other cereals	-	0.275	-	-0.275	-	0.158	-	-0.158
Oilseeds, protein crops	-	-	0.419	-0.419	-	-	0.460	-0.460
Acreage share functions, short-run								
Wheat	0.569	-0.269	-0.228	-	0.228	-0.107	-0.090	-
Other cereals	-0.479	0.702	-0.228	-	-0.199	0.285	-0.090	-
Oilseeds, protein crops	-0.479	-0.269	0.617	-	-0.199	-0.107	0.253	-
Acreage share functions, long-run								
Wheat	-	-	-	-	0.978	-0.461	-0.388	-
Other cereals	-	-	-	-	-0.856	1.221	-0.388	-
Oilseeds, protein crops	-	-	-	-	-0.856	-0.461	1.087	-

Appendix A. Acreage Shares in the Standard MNL Model

$$\Pi = \max_{s_c} \left[\sum_{x=1}^C \pi_x s_x - \left(A + \sum_{x=1}^C c_x s_x + a^{-1} \sum_{c=1}^C s_c \ln s_c \right) \right] \text{ subject to } \sum_{c=1}^C s_c = 1$$

The Lagrangian is defined by:

$$L(s, \lambda) = \sum_{c=1}^C \pi_c s_c - \left(A + \sum_{c=1}^C c_c s_c + a^{-1} \sum_{c=1}^C s_c \ln s_c \right) - \lambda \left(\sum_{c=1}^C s_c - 1 \right)$$

The first order conditions are then:

$$(1) \quad \frac{\partial L}{\partial s_c} = \pi_c - c_c - a^{-1} (\ln s_c + 1) - \lambda = 0$$

$$(2) \quad \frac{\partial L}{\partial \lambda} = \sum_{m=1}^C s_m - 1 = 0$$

Equation (1) leads to:

$$\pi_c - c_c - a^{-1} \ln s_c - a^{-1} = \lambda$$

$$s_c = \exp[a(\pi_c - c_c)] \exp[-(a\lambda + 1)]$$

Integration equation (2) implies:

$$\sum_{m=1}^C s_m = \left[\sum_{k=1}^C \exp[a(\pi_m - c_m)] \right] \exp[-(a\lambda + 1)] = 1$$

$$\exp[-(a\lambda + 1)] = \left[\sum_{m=1}^C \exp[a(\pi_m - c_m)] \right]^{-1}$$

$$s_c = \exp[a(\pi_c - c_c)] \left[\sum_{m=1}^C \exp[a(\pi_m - c_m)] \right]^{-1}$$

$$s_c = \frac{\exp[a(\pi_c - c_c)]}{\sum_{m=1}^C \exp[a(\pi_m - c_m)]}$$

Appendix B. Acreage Shares in the Nested MNL Model

$$\Pi = \max_s \sum_{l=1}^Q \sum_{m \in B(l)} s_m (\pi_m - c_m) - a^{-1} \left(\sum_{l=1}^Q \bar{s}_l \ln \bar{s}_l + \sum_{l=1}^Q a \alpha_l^{-1} \bar{s}_l \sum_{m \in B(l)} s_{m/l} \ln s_{m/l} \right) \quad \text{subject to}$$

$$\sum_{k=1}^K s_m = 1$$

Note $\bar{s}_l = \sum_{m \in B(l)} s_m$ and $s_{m/l} = s_m / \bar{s}_l$. The Lagrangian is defined by:

$$L(s, \lambda) = \sum_{l=1}^Q \sum_{m \in B(l)} s_m (\pi_m - c_m) - a^{-1} \left(\sum_{l=1}^Q \bar{s}_l \ln \bar{s}_l + \sum_{l=1}^Q a \alpha_l^{-1} \bar{s}_l \sum_{m \in B(l)} s_{m/l} \ln s_{m/l} \right) - \lambda \left(\sum_{m=1}^Q s_m - 1 \right)$$

The first order conditions for the crop c into the nest q are:

$$(1) \quad \frac{\partial L}{\partial s_c} = \pi_c - c_c - a^{-1} (\ln \bar{s}_q + 1) - \alpha_q^{-1} \ln s_{c/q} - \lambda = 0$$

$$(2) \quad \frac{\partial L}{\partial \lambda} = \sum_{c=1}^C s_c - 1 = 0$$

Equation (1) leads to:

$$(3) \quad \alpha_q^{-1} \ln s_{c/q} + a^{-1} \ln \bar{s}_q = \pi_c - c_c - a^{-1} - \lambda$$

First stage:

$$\alpha_q^{-1} \ln s_c - \alpha_q^{-1} \ln \bar{s}_q + a^{-1} \ln \bar{s}_q = \pi_c - c_c - a^{-1} - \lambda$$

$$\alpha_q^{-1} \ln s_c = \pi_c - c_c - a^{-1} - \lambda - (a^{-1} - \alpha_q^{-1}) \ln \bar{s}_q$$

$$s_c = \exp \left[\alpha_q (\pi_c - c_c) \right] \exp \left[-\alpha_q (a^{-1} + \lambda) \right] \exp \left[(1 - \alpha_q a^{-1}) \ln \bar{s}_q \right]$$

Integration of equation $\bar{s}_l = \sum_{m \in B(l)} s_m$ leads to:

$$\sum_{m \in B(q)} s_m = \exp \left[-\alpha_q (a^{-1} + \lambda) \right] \exp \left[(1 - \alpha_q a^{-1}) \ln \bar{s}_q \right] \sum_{m \in B(q)} \exp \left[\alpha_q (\pi_m - c_m) \right] = \bar{s}_q$$

$$\exp\left[-\alpha_q (a^{-1} + \lambda)\right] \exp\left[+(1 - \alpha_q a^{-1}) \ln \bar{s}_q\right] = \bar{s}_q \left[\sum_{m \in B(q)} \exp\left[\alpha_q (\pi_m - c_m)\right] \right]^{-1}$$

$$s_c = \frac{\exp\left[\alpha_q (\pi_c - c_c)\right]}{\sum_{m \in B(q)} \exp\left[\alpha_q (\pi_m - c_m)\right]} \bar{s}_q$$

$$s_c / \bar{s}_q = s_c = \frac{\exp\left[\alpha_q (\pi_c - c_c)\right]}{\sum_{m \in B(q)} \exp\left[\alpha_q (\pi_m - c_m)\right]}$$

Second stage:

$$\alpha_q^{-1} \ln s_{c/q} + a^{-1} \ln \bar{s}_q = \pi_c - c_c - a^{-1} - \lambda$$

$$\ln \bar{s}_q = a(\pi_c - c_c) - 1 - a\lambda - a\alpha_q^{-1} \ln s_{c/q}$$

$$\bar{s}_q = \exp\left[a(\pi_c - c_c)\right] \exp[-1 - a\lambda] \exp\left[-a\alpha_q^{-1} \ln s_{c/q}\right]$$

Integration of equation $\sum_l \bar{s}_l = 1$ leads to:

$$\sum_l \bar{s}_l = \exp\left[a(\pi_c - c_c)\right] \exp[-1 - a\lambda] \sum_l \left[\exp\left[-a\alpha_l^{-1} \ln s_{c/l}\right] \right] = 1$$

$$\exp\left[a(\pi_c - c_c)\right] \exp[-1 - a\lambda] = \left[\sum_l \left[\exp\left[-a\alpha_l^{-1} \ln s_{c/l}\right] \right] \right]^{-1}$$

$$\bar{s}_q = \frac{\exp\left[-a\alpha_q^{-1} \ln s_{c/q}\right]}{\sum_l \left[\exp\left[-a\alpha_l^{-1} \ln s_{c/l}\right] \right]}$$

$$\bar{s}_q = \frac{\left(s_c / \bar{s}_q\right)^{(-a\alpha_q^{-1})}}{\left[\sum_l \left(s_c / s_l\right)^{(-a\alpha_l^{-1})} \right]}$$

Integration of the equation $s_c / \bar{s}_q = \frac{\exp[\alpha_q (\pi_c - c_c)]}{\sum_{m \in B(q)} \exp[\alpha_q (\pi_m - c_m)]}$ leads to:

$$\bar{s}_q = \frac{\left[\sum_{m \in B(q)} \exp[\alpha_q (\pi_m - c_m)] \right]^{(a\alpha_q^{-1})}}{\left[\exp[\alpha_q (\pi_c - c_c)] \right]^{(a\alpha_q^{-1})}} \sum_l \left[\frac{\left[\exp[\alpha_l (\pi_c - c_c)] \right]^{(a\alpha_l^{-1})}}{\left[\sum_{m \in B(l)} \exp[\alpha_l (\pi_m - c_m)] \right]^{(a\alpha_l^{-1})}} \right]$$

Because $\sum_l \left[\frac{\left[\exp[\alpha_l (\pi_c - c_c)] \right]^{a\alpha_l^{-1}}}{\left[\sum_{m \in B(l)} \exp[\alpha_l (\pi_m - c_m)] \right]^{a\alpha_l^{-1}}} \right] = \frac{\left[\exp[\alpha_q (\pi_c - c_c)] \right]^{a\alpha_q^{-1}}}{\sum_l \left[\sum_{m \in B(l)} \exp[\alpha_l (\pi_m - c_m)] \right]^{a\alpha_l^{-1}}}$ then:

$$\bar{s}_q = \frac{\left[\sum_{m \in B(q)} \exp[\alpha_q (\pi_m - c_m)] \right]^{a\alpha_q^{-1}}}{\sum_l \left[\sum_{m \in B(l)} \exp[\alpha_l (\pi_m - c_m)] \right]^{a\alpha_l^{-1}}}$$

The final expression of the acreage shares for the crop c into the nest q is:

$$s_c = \frac{\exp[\alpha_q (\pi_c - c_c)]}{\sum_{m \in B(q)} \exp[\alpha_q (\pi_m - c_m)]} \frac{\left[\sum_{m \in B(q)} \exp[\alpha_q (\pi_m - c_m)] \right]^{a\alpha_q^{-1}}}{\sum_l \left[\sum_{m \in B(l)} \exp[\alpha_l (\pi_m - c_m)] \right]^{a\alpha_l^{-1}}}$$

Appendix C. Acreage Share Price Elasticities in the Standard MNL Model

$$\begin{aligned}\frac{\partial s_c(\boldsymbol{\pi}; a, \mathbf{c})}{\partial p_c} &= a \frac{\partial \pi_c}{\partial p_c} s_c(\boldsymbol{\pi}; a, \mathbf{c}) [1 - s_c(\boldsymbol{\pi}; a, \mathbf{c})] > 0 \\ \frac{\partial s_c(\boldsymbol{\pi}; a, \mathbf{c})}{\partial p_\ell} &= -a \frac{\partial \pi_\ell}{\partial p_\ell} s_c(\boldsymbol{\pi}; a, \mathbf{c}) s_\ell(\boldsymbol{\pi}; a, \mathbf{c}) < 0 \\ \frac{\partial s_c(\boldsymbol{\pi}; a, \mathbf{c})}{\partial \mathbf{w}} &= a s_c(\boldsymbol{\pi}; a, \mathbf{c}) \sum_{m=1, \neq c}^C \left(\frac{\partial \pi_c}{\partial \mathbf{w}} - \frac{\partial \pi_m}{\partial \mathbf{w}} \right) s_m(\boldsymbol{\pi}; a, \mathbf{c})\end{aligned}$$

because:

$$\begin{aligned}s_c(\boldsymbol{\pi}; a, \mathbf{c}) &= \frac{\exp[a(\pi_c - c_c)]}{\sum_{m=1}^C \exp[a(\pi_m - c_m)]} \\ \frac{\partial s_c(\boldsymbol{\pi}; a, \mathbf{c})}{\partial p_c} &= a \frac{\partial \pi_c}{\partial p_c} \frac{\exp[a(\pi_c - c_c)] \sum_{m=1}^C \exp[a(\pi_m - c_m)] - \exp[a(\pi_c - c_c)]^2}{\left(\sum_{m=1}^C \exp[a(\pi_m - c_m)] \right)^2} \\ &= a \frac{\partial \pi_c}{\partial p_c} s_c(\boldsymbol{\pi}; a, \mathbf{c}) [1 - s_c(\boldsymbol{\pi}; a, \mathbf{c})] \\ \frac{\partial s_c(\boldsymbol{\pi}; a, \mathbf{c})}{\partial p_\ell} &= -a \frac{\partial \pi_\ell}{\partial p_\ell} \frac{\exp[a(\pi_c - c_c)] \exp[a(\pi_\ell - c_\ell)]}{\left(\sum_{m=1}^C \exp[a(\pi_m - c_m)] \right)^2} \\ &= -a \frac{\partial \pi_\ell}{\partial p_\ell} s_c(\boldsymbol{\pi}; a, \mathbf{c}) s_\ell(\boldsymbol{\pi}; a, \mathbf{c}) \\ \frac{\partial s_c(\boldsymbol{\pi}; a, \mathbf{c})}{\partial \mathbf{w}} &= a \exp[a(\pi_c - c_c)] \frac{\sum_{m=1}^C \frac{\partial \pi_c}{\partial \mathbf{w}} \exp[a(\pi_m - c_m)] - \sum_{m=1}^C \frac{\partial \pi_m}{\partial \mathbf{w}} \exp[a(\pi_m - c_m)]}{\left(\sum_{m=1}^C \exp[a(\pi_m - c_m)] \right)^2} \\ &= a \exp[a(\pi_c - c_c)] \frac{\sum_{m=1, \neq c}^C \left(\frac{\partial \pi_c}{\partial \mathbf{w}} - \frac{\partial \pi_m}{\partial \mathbf{w}} \right) \exp[a(\pi_m - c_m)]}{\left(\sum_{m=1}^C \exp[a(\pi_m - c_m)] \right)^2} \\ &= a s_c(\boldsymbol{\pi}; a, \mathbf{c}) \sum_{m=1, \neq c}^C \left(\frac{\partial \pi_c}{\partial \mathbf{w}} - \frac{\partial \pi_m}{\partial \mathbf{w}} \right) s_m(\boldsymbol{\pi}; a, \mathbf{c})\end{aligned}$$

Chapitre III

Land and Variable Input Allocation:

Why Heterogeneity Matters ? ¹⁶

Résumé

L'allocation des intrants entre les variables est un problème récurrent en économie agricole à cause du manque de données disponibles sur les dépenses d'intrants par culture. Les méthodes utilisées pour répartir les intrants entre les cultures s'appuient généralement sur des méthodes de régression standards. Nous montrons que ces méthodes sont vraisemblablement biaisées du fait de la corrélation des surfaces avec le terme d'erreur dans l'équation de répartition des intrants. Cette corrélation est due à l'hétérogénéité non-observée des exploitations qui influence en même temps les décisions d'utilisation d'intrants et les choix d'assolements. Nous construisons un modèle des choix de production « complet » qui tient compte explicitement de l'hétérogénéité des exploitations, à travers des variables (hétérogénéité observée) et d'un terme d'erreur (hétérogénéité non-observée). Une approche par les fonctions de contrôle est alors utilisée pour déterminer explicitement la corrélation entre les surfaces et le terme d'erreur liés à l'hétérogénéité et contrôler ainsi l'endogénéité des surfaces. L'application de cette méthode est réalisée à partir d'une procédure en trois étapes sur des données d'exploitations françaises de 1989 à 2006.

¹⁶ This chapter is based on a paper jointly written with Alain Carpentier and Céline Nauges.

3.1 Introduction

The allocation of variable inputs among crops is a common problem in applied studies using farm accountancy data. Standard farm-accounting information is typically restricted to aggregate or whole-farm input expenditures, with no details on how these expenditures are split among crops. Most of studies employing multi-crop econometric models with land as an allocable fixed input considered generally variable input uses at the farm level (Moore and Negri, 1992). However allocation of variable inputs among crops appears to be useful for several reasons: to analyze the evolution of the gross margins at the crop level, to investigate the empirical validity of the multi-crop econometric model or to provide important information for extension agents or farmers' advisor.

A large number of authors have been working on this topic, either to provide solutions for allocating input costs between crops or activities (Just et al., 1983; Chambers and Just, 1989), or to compute input-output coefficients (Dixon and al., 1984; Hornbaker, Dixon and Sonka, 1989; Peeters and Surry, 1993); or because this was a necessary step of their analysis (for example the evaluation of agro-environmental policies on input use in Lence and Miller, 1998). The most widely used methods to allocate variable input uses to crops are based on regression models or production function models with constraints on variable input total uses (Dixon and *al.*, 1984; Hornbaker and *al.*, 1989; Just and *al.*, 1990). However allocation of variable inputs among crops depends on how the farmers allocate land among crops, a decision which itself takes into account input uses by crop. Crop input use decisions and acreage choices are partially simultaneous. The underlying idea is that variable input allocation requires the specification of a complete production model, *i.e.* describing land allocation, use of variable inputs and crop yields in order to take into account the link between the acreage and the input use choices.

The contribution of this article is threefold. First, it shows that the standard regression based approaches for allocating variable input uses to crops are likely to be biased due to the partial simultaneity of the (expected) crop variable input and acreage choices. Second,

it proposes a structural econometric multi-crop model for determining the origin of these biases. The structure of the model relies on the timing of the farmers' choices. The specified model distinguishes two sorts of error terms: the terms accounting for farms' heterogeneity and the terms accounting for the stochastic events affecting crop production. It provides explicit functional forms of the links between the error terms of the yield supply, input demand allocation and acreage equations. Third, it proposes a method based on control functions to eliminate the bias associated with the standard regression based methods. It builds on previous result obtained for the estimation of the so-called « correlated random coefficient models » (see, *e.g.*, Imbens and Wooldridge, 2007; Wooldridge, 2008) and « average treatment effects » (see, *e.g.*, Heckman and al., 2003). The empirical implementation of the proposed methods is described in three stages and an application is presented on French farm-level data.

This paper proceeds as follows. The next section presents a review of the literature about input allocation method and presents briefly the endogeneity problems in these standard approaches and the solution adopted in this paper, *i.e.* the control function based approach. It requires an econometric multi-crop (for acreage, yield and input choices) model which is described in the second section. The third section presents the control functions approach used to take into account the links between the acreage and the input use choices in the variable input allocation equation. In the fourth section, a general three-stage procedure for implementing the approach and an application on French farm-level data are proposed. The last section of this paper provides some concluding remarks.

3.2 Literature review

The most common farm data on crop production consist in acreages, yields and prices at the crop level, and variable input uses and quasi-fixed factor quantities (measures of labour and capital) at the farm level. Input price indices are generally made available by the national departments of agriculture at the regional level. Farmer i ($i=1, \dots, N$) produces C crops ($c=1, \dots, C$) to which they allocate their S units of land.

In what follows, we suppose one single variable input. X_i denotes the quantity of variable input use at the farm level for farm i , w_i is the input price for farm i , x_{ci} denotes the quantity of variable input uses for crop c per unit of land for farm i , s_{ci} is the acreage share of crop c for farm i , y_{ci} denotes the yield of crop c and p_{ci} denotes its price for farm i . The input allocation problem consists in recovering input quantities x_{ci} for $c=1, \dots, C$.

Several approaches have been used or proposed for solving this allocation problem. We distinguish two main groups in the literature: the first group includes approaches that consider solely input allocation equation(s) as the one defined above. In these models, input allocations are treated as parameters to be estimated, along the lines of Just, Zilberman, Hochman and Bar-Shira (1990) terminology. These are, by far, the most widely used in practice. In the second group, input allocation equations belong to a system of equations including crop supply and acreage functions, or production functions (Chambers and Just, 1989). In what follows, we describe the first group type of approaches, along with their advantages and limits. These limits provide arguments for using the approaches of the second type.

3.2.1 Approaches based on single input allocation equations

Among the available methods for allocating inputs to activities or crops, the most widely used is the regression method that considers variable input allocation x_{ci} as parameters:

$$(1) \quad x_i = \sum_c^C s_{ci} x_{ci} + \eta_i \quad \text{with} \quad E[\eta_i / \mathbf{s}_i] = 0,$$

or as parametric functions:

$$(2) \quad x_i = \sum_c^C s_{ci} x_{ci}(\mathbf{z}_i; \mathbf{a}) + \eta_i \quad \text{with} \quad E[\eta_i / \mathbf{s}_i, \mathbf{z}_i] = 0,$$

where \mathbf{z}_i is the vector of exogenous variables such as farm's characteristics and activities, \mathbf{a} the vector of corresponding unknown parameters and \mathbf{s}_i is the vector of acreage shares.

Ordinary Least Square (OLS) for a single input model or seemingly unrelated regression (SUR) for a system of input allocation equations provide consistent estimators of x_{ci} and \mathbf{a} under the assumption that the conditional expectation of η_i is zero. See for example the behavioural model of Just et al. (1990) and the vast majority of the related literature.

Later, these models have been generalized by adding random terms to the crop input use models to account for the effects of unobserved determinants of input choices. Models (1) and (2) are then respectively written:

$$(3) \quad x_i = \sum_c^C s_{ci} [x_{ci} + u_{ci}^x] + \eta_i \quad \text{with} \quad E[\eta_i / \mathbf{s}_i] = E[u_{ci}^x / \mathbf{s}_i] = 0,$$

$$(4) \quad x_i = \sum_c^C s_{ci} [x_{ci}(\mathbf{z}_i; \mathbf{a}) + u_{ci}^x] + \eta_i \quad \text{with} \quad E[\eta_i / \mathbf{z}_i, \mathbf{s}_i] = E[u_{ci}^x / \mathbf{z}_i, \mathbf{s}_i] = 0,$$

where η_i terms include measurement errors, stock variations (...) and the u_{ci}^x terms are defined as the difference between the « true » values of the unobserved input uses and the values what can be « explained » by the \mathbf{z}_i variables. Models (3) and (4) are input allocation equations with random parameters. In these models, the error terms, $\sum_c^C s_{ci} u_{ci}^x + \eta_i$ are heteroskedastic, and feasible generalized OLS or SUR estimations will provide efficient estimators of the parameter vector \mathbf{a} under the assumption that the error terms u_{ci}^x and η_i have constant variances and covariances (Dixon, Batte and Sonka, 1984; Hornbaker, Dixon and Sonka, 1989; Dixon and Hornbaker 1992).¹⁷

The approaches just described are easy to implement and can provide satisfactory results (Just, Zilberman, Hochman and Bar-Shira, 1990). However, the consistency of the regression estimators of \mathbf{a} in the generalized input allocation equation system relies on the assumption that acreage shares \mathbf{s}_i are exogenous with respect to u_{ci}^x , i.e.:

¹⁷ Surry and Peeters (2001) consider a similar equation system but exploit the flexibility of the Maximum Entropy (GME) statistical framework to compute crop input use estimates per farm. The ME framework also permits to easily impose positivity constraints on the input allocation and to make use of information provided by extension services.

$$(5) \quad E[u_{ci}^x / \mathbf{z}_i, \mathbf{s}_i] = 0$$

These conditional mean conditions are unlikely to hold with farm data, for the simple reason that input use x_{ci} partly determines profitability of crop c , which itself is a determinant of crop c acreage. Since x_{ci} are determinants of the acreage choices, any part of x_{ci} is a determinant of the choice of s_{ci} . As a result, the conditions:

$$(6) \quad E[u_{ci}^x / \mathbf{s}_i] = 0$$

hold if and only if $u_{ci} = 0$, i.e. in the unrealistic case where \mathbf{z}_i are "perfect" control variables for the heterogeneity of x_{ci} . Of course the biases due the endogeneity of \mathbf{s}_i are reduced by the use of « imperfect » control variables. These biases are also likely to be limited if the elements of the x_{ci} vectors represents small amounts when compared to the crop returns.

These approaches based on single input allocation equations suffer from the same limits. Hence, the specification of a complete production model (describing land allocation, use of variable inputs and crop yields) is necessary in order to account for the link between the input uses and acreages choices.

3.2.2 Approaches based on multicrop econometric models

We discuss here models in which input allocation equations are estimated jointly with other equations, such as production technology or models describing acreage choices. Multicrop models dealing with production dynamics (*e.g.*, Ozarem and Miranowski, 1994), risk aversion (*e.g.*, Coyle, 1992, 1999 ; Chavas and Holt, 1990) and price uncertainty (*e.g.*, Coyle, 1992, 1999 ; Moro and Sckokai, 2006) or models based on plot per plot discrete choice (*e.g.*, Wu and Segerson, 1995) are not considered here. Also, we focus

on models in which land is considered as an allocatable fixed input (Shumway, Pope and Nash, 1984), *i.e.*, models designed for analyzing farmers' short run decisions.

In papers falling into this category, the problem of variable input allocation is considered as a by product or not considered in further details. Lence and Miller (1998), in a Maximum Entropy framework, estimate jointly crop production function models and crop input uses. Their use of the flexible maximum entropy estimators enables them to allocate the farm input uses by using a system of production function models (one for each crop) and constraining the crop input uses to sum to the farms' input uses. This approach ensures the consistency of the determined input allocation with a system of production functions. Note also that this approach does not rely on the modelling of farmers' economic choices. In this respect, Lence and Miller's (1998) approach lies in between the « atheoretical » approach of Dixon *et al.* (1984), Hornbaker *et al.* (1989) and the approach based on the specification of a complete model of farmers' choices. They use production functions but they don't use farmers' production choice models. But, as acknowledged by Lence and Miller, their approach as well as the other « atheoretical » approaches share the same drawback: they do not consider input uses and acreages (or production levels in Lence and Miller's approach) as (partly) simultaneous choices.

The first econometric models designed to model crop acreage decisions explicitly consider the variable input use allocation problem (Just, Zilberman and Hochman, 1983; Chambers and Just, 1989). Just *et al.* (1983) and Chambers and Just (1989) also consider variable input allocation issues although these are not the main focus of their studies. They employ an approach similar to the one used here in the sense that they determine the variable input allocation by considering a complete model of farmers' choices. The variable input allocation is merely only a by-product of their modelling exercise. Nevertheless their multicrop econometric model differs from ours in important respects. First, their use of Cobb-Douglas crop yield functions (Just *et al.* 1983) or Translog crop profit (Chambers and Just 1989) functions facilitates their determination of the variable input allocations.

Second, their multicrop econometric models is consistent in their deterministic part but they are not consistent in their random parts. Their econometric models are derived from their economic models basically by adding error terms to the deterministic equations derived from the economic model although Just *et al.* (1983) added random terms with structural interpretations. These points are further detailed below. But it is worth noting that the approach considered here builds on the two critics of the previous approaches presented above.

Acreage allocation models considered in the 1990's mostly use the model designed by Moore and Negri (1992) (see *e.g.* Moore, Gollehon and Carey, 1994 ; Moore and Dinar, 1995 ; Guyomard, Baudry and Carpentier, 1996; Oude Lansink and Peerlings, 1996; Bel Haj Hassine and Simioni, 2000; Bel, Lacroix, Salanié et Thomas, 2006). Moore and Negri's (1992) model is a variant of Chambers and Just's (1989) model for input non-joint multicrop technology where restricted Translog profit functions are replaced by restricted Normalized Quadratic Profit functions at the crop level. Function of profit, production and input demand functions in levels are much easier to use than their counterparts in logarithm because the total profit and the land constraint are defined as sums of the crop profits and the crop acreage levels (or shares). Variable input uses are usually considered at the farm level in most of studies employing multi-crop econometric models (Paris, 1989).

3.2.3 Outline of the control function approach

The starting point of this research is that the exogeneity conditions $E[u_{ci}^x / \mathbf{z}_i, \mathbf{s}_i] = 0$ required for the consistency of the regression based approaches are unlikely to hold in applied work. The argument for this claim is simple. The acreage choices \mathbf{s}_i depend on the relative (marginal) profitability of the crops. This profitability depends on input uses and, consequently, \mathbf{s}_i depends on how x_{ci} affects this profitability. Furthermore, this endogeneity problem cannot be solved by using standard instrumental variable (VI)

techniques, because the error term $s_{ci}u_{ci}^x + \eta_i$ contains the endogenous explanatory variables \mathbf{s}_i . The use of equation (1) as an estimating equation requires the control of the terms $E[u_{ci}^x / \mathbf{z}_i, \mathbf{s}_i]$.

The approach used to control these terms is based on control functions approach. The principle of the control function approach is now standard to account for endogenous sample selection (Heckman, 1974, 1979), correlated fixed effects in panel data models (Chamberlain, 1982) or endogenous explanatory variables in linear (Hausman, 1978) or non-linear models (Smith and Blundell, 1986; Petrin and Train, 2008; see also Imbens and Wooldridge, 2007 for a recent survey).

This section describes briefly the principle of the control function approach. Let assume that the considered model allows to define the $E[u_{ci}^x / \mathbf{z}_i, \mathbf{s}_i]$ terms known functions of \mathbf{z}_i , \mathbf{s}_i and of a vector of unknown parameters $\boldsymbol{\theta}$. Let assume also that there exists a consistent estimator of $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}$. The input allocation equation (1) can be transformed as:

$$(7) \quad x_i = \sum_c s_{ci} \left[x_{ci}(\mathbf{z}_i; \mathbf{a}) + c_c^x(\mathbf{z}_i, \mathbf{s}_i; \boldsymbol{\theta}) \right] + \omega_i^x,$$

$$(8) \quad \text{with } \omega_i^x = \sum_c s_{ci} \left[u_{ci}^x(\mathbf{z}_i; \mathbf{a}) - c_{ci}^x(\mathbf{z}_i, \mathbf{s}_i; \boldsymbol{\theta}) \right] + \eta_i,$$

where $c_{ci}(\mathbf{z}_i, \mathbf{s}_i; \boldsymbol{\theta})$ are the control functions and where the conditional expectation of $E[\omega_i^x / \mathbf{z}_i, \mathbf{s}_i]$ is null by construction. Since the $c_{ci}(\mathbf{z}_i, \mathbf{s}_i; \hat{\boldsymbol{\theta}})$ terms are consistent estimators of the corresponding $c_{ci}(\mathbf{z}_i, \mathbf{s}_i; \boldsymbol{\theta})$ terms, equation (7) can be used to construct consistent regression based estimators of \mathbf{a} . The control function approach basically splits the error terms u_{ci}^x in two terms: the control function $c_{ci}(\mathbf{z}_i, \mathbf{s}_i; \boldsymbol{\theta}) = E[u_{ci}^x / \mathbf{z}_i, \mathbf{s}_i]$ which « captures » and thus controls the links between u_{ci}^x and the endogenous variable vector

\mathbf{s}_i ; and a « new » error term $u_{ci}^x(\mathbf{z}_i; \mathbf{a}) - c_{ci}^x(\mathbf{z}_i, \mathbf{s}_i; \boldsymbol{\theta})$. By construction, \mathbf{s}_i is exogenous with respect to the « new » error term.

The crucial point is then to define the control functions $c_{ci}(\mathbf{z}_i, \mathbf{s}_i; \boldsymbol{\theta})$ for $c=1, \dots, C$. This requires assumptions about the error terms of the multi-crop econometric model. In the case where the acreage share function model is defined by:

$$(9) \quad s_{ci} = s_{ci}(\mathbf{z}_i; \mathbf{b}) + \omega_{ci}^s \text{ with } E[\omega_{ci}^s / \mathbf{z}_i] = 0,$$

The control functions are determined by the following conditional expectations:

$$(10) \quad c_{ci}^x(\mathbf{z}_i, \mathbf{s}_i; \boldsymbol{\theta}) = E[u_{ci}^x / \mathbf{z}_i, \mathbf{s}_i] = E[u_{ci}^x / \mathbf{z}_i, s_{ci}(\mathbf{z}_i; \mathbf{b}) + \omega_{ci}^s] = E[u_{ci}^x / \mathbf{z}_i, \omega_{ci}^s],$$

As a result, it is necessary to define the relationship between the error term vectors u_{ci}^x and ω_{ci}^s . It is thus necessary to define a « structural » multi-crop econometric model, *i.e.* a model in which the error terms are specified as unknown determinants of the modelled choices, and not just random terms added to « make statistical noise ».

3.3 Econometric model specification

Although the proposed approach can be applied with other multi-crop models (with more or less adaptations), a specific multi-crop econometric model is considered to « concretely » illustrate the basic features of the approach. This model combined standard quadratic yield functions with crop acreage (share) functions derived along the line of Heckelee and Wolff (2003). The model considered is the one described in chapter two. It is chosen because of its fairly simple interpretation and its flexibility. The error terms of the econometric model are defined as integral parts of this model (see, *e.g.*, McElroy, 1987).

The model is considered in its simplest version, *i.e.* with constant parameters. In empirical work most of the defined parameters may usefully defined as parametric functions of

observed exogenous variable to control (as much as possible) for the heterogeneity of the farms and farmers. Finally the single variable input is considered for simplicity.

3.3.1 Yield functions

The yield y_{ci} of each crop c ($c=1, \dots, C$) for farm i ($i=1, \dots, N$) is assumed to be a quadratic function of the single variable input (for simplicity). This function represents the short term « agronomic » yield function and is defined as:

$$(11) \quad y_{ci} = \alpha_{ci} - 0.5\gamma_c^{-1}(\beta_{ci} - x_{ci})^2 \quad \text{with} \quad \alpha_{ci} = \alpha_{0c} + v_{ci}^y \quad \text{and} \quad \beta_{ci} = \beta_{0k} + v_{ci}^x,$$

where x_{ci} is the quantity of variable input used per hectare by farm i devoted to crop c , and α_{ci} , γ_c and β_{ci} are parameters to be estimated with $\alpha_{ci} > 0$, $\gamma_c > 0$ and $\beta_{ci} > 0$. This alternative specification of the standard quadratic function is also used by Pope and Just (2003) albeit for other purposes. The yield function is strictly concave if $\gamma_c > 0$. Under this assumption the term α_{ci} can be interpreted as the maximum yield of crop c for farm i . The variable input quantity required for achieving this maximum yield is given by β_{ci} . The estimates of these yield functions can thus be checked with agricultural scientists or extension agents.

The v_{ci}^y and v_{ci}^x are random terms. These terms are split into two parts for simplifying their interpretation:

$$(12) \quad v_{ci}^y = e_{ci}^y + \varepsilon_{ci}^y \quad \text{and} \quad v_{ci}^x = e_{ci}^x + \varepsilon_{ci}^x.$$

The terms e_{ci}^y and e_{ci}^x are denoted as heterogeneity terms. They represent the effects on the yield of crop c of factors that are known to farmer i at the time he chooses his acreages (rotation effects, soil quality, but also quasi-fixed input availabilities...). These terms are closely related to the so-called « fixed effects » in the panel data econometrics literature (see, *e.g.*, Griliches and Mairesse, 1995), but they may not be « permanent » in the current

framework. They are considered as random because they are unknown to the econometrician. The terms ε_{ci}^y and ε_{ci}^x are denoted as stochastic events. They represent the effects on the yield of crop c of factors that are unknown to farmer i at the time he chooses his acreages (climatic conditions, pest infestations...). These factors are considered as random because they vary across farms and years, and are unknown to the econometrician. Their expectations are normalized to be null.

The production of crop c is sold at price p_{ci} and the input is bought at price w_i by the farmer i . These prices are assumed to be known at the beginning of the production process, *i.e.* when acreages are chosen. Farmers are supposed risk-neutral. Farmer i is assumed to choose his input use by maximizing the following gross margins π_{ci} for each crop c :

$$(13) \quad p_{ci}y_{ci} - w_i x_{ci}.$$

Variable input and “target” yields choices are assumed based on output and input prices and adjusted to specific production condition, *i.e.* after farmer has observed ε_{ci}^y and ε_{ci}^x . The maximisation of this profit function under technological constraints leads to the following per hectare variable input demand and supply functions:

$$(14) \quad x_{ci} = \beta_{0c} - \gamma_c (w_i / p_{ci}) + v_{ci}^x$$

$$(15) \quad y_{ci} = \alpha_{0c} - 0.5 \gamma_c (w_i / p_{ci})^2 + v_{ci}^y.$$

Consequently v_{ci}^x can be interpreted as the effects production conditions that can be « corrected » by variable input uses while v_{ci}^y represents the effects of fully undergone production conditions. The quadratic yield have a main practical advantage: they provide yield supply and variable input demand functions with additive error terms. This feature appears to be very useful for analysing the error term structure of the econometric model (see, *e.g.*, McElroy, 1987, and Pope and Just, 2003, in other contexts). Distinguishing the

heterogeneity effects and the stochastic events in the yield function allows to determine the gross margins of the crops as they are expected by the farmers at the time they choose their acreages:

$$(16) \quad \pi_{ci}^e = p_{ci}\alpha_{0c} - w_i\beta_{0c} + 0.5\gamma_c(w_i/p_{ci})^2 + p_{ci}e_{ci}^y - w_ie_{ci}^x.$$

The farmers' gross margin expectations can not depend on the ε_{ci}^y and ε_{ci}^x terms because these terms are unknown when farmers choose their acreages.

3.3.2 Acreage functions

Farmers' acreage choices are modelled within the framework developed by Heckelee and Wolff (2003). This framework is simple, flexible and links the econometric and mathematical programming literature on production choice modelling. Farmer i is assumed to allocate his total land quantity S_i by maximizing the following indirect restricted profit function:

$$(17) \quad \sum_{c=1}^C s_{ci}\pi_{ci}^e - C(\mathbf{s}_i),$$

where \mathbf{s}_i is the vector of acreage share for farmer i . According to this model, the implicit management cost of the chosen acreage $C(\mathbf{s}_i)$ is the motive for crop diversification for farmers. This cost function is increasing and quasi-convex in \mathbf{s}_i . It is used in the mathematical programming literature. It can be interpreted as a reduced form function smoothly approximating the unmeasured and implicit costs associated with a given land allocation and farm specific constraints. This cost function is assumed to have this form:

$$(18) \quad C(\mathbf{s}_i) = g + \sum_c s_{ci}g_{ci} + a^{-1} \sum_c s_{ci} \ln s_{ci} \quad \text{with} \quad g_{ci} = g_{0c} + e_{ci}^g,$$

where g, g_{ci} and a are parameters to be estimated. This form of the cost function is used and explained in the chapter one and two. The g_{ci} parameter are « fixed » cost (in the short run) per unit of land devoted to crop c of farmer i . This parameter is split into two parts: g_{0c} a parameter and e_{ci}^g a random term accounting for the cost heterogeneity term known to farmer i but unknown to the econometrician.

Crop 1 is the « reference » crop. The maximisation of this indirect restricted profit function under the land use constraint leads to the closed form of the acreage shares functions:

$$(19) \quad s_{ci} = \frac{\exp\left[a(\pi_{ci}^e - g_{0c} - e_{ci}^g)\right]}{\sum_{m=1}^C \exp\left[a(\pi_m^e - g_{0m} - e_{mi}^g)\right]} = \frac{\exp\left[a(\pi_{ci}^e - \pi_{1i}^e - g_{0c} - e_{ci}^g)\right]}{1 + \sum_{m=2}^C \exp\left[a(\pi_{mi}^e - \pi_{1i}^e - g_{0m} - e_{mi}^g)\right]} .$$

The main problem of these equations is the non-linearity of error terms. Thanks to the log-linear transformation, the acreage share functions can be defined as:

$$(20a) \quad \ln \Delta s_{ci} = a \left[\Delta \pi_{ci}^a - g_c \right] + v_{ci}^s, \text{ for } c = 2, \dots, C ,$$

with :

$$(20b) \quad \ln \Delta s_{ci} = \ln(s_{ci} / s_{1i}) ,$$

$$(20c) \quad \Delta \pi_{ci}^a = p_{ci} \alpha_{0c} - \gamma_c \left(\frac{w_i}{p_{ci}} \right) + \frac{1}{2} \beta_{0c} p_{ci} \left(\frac{w_i}{p_{ci}} \right)^2 - \left(p_{1i} \alpha_{01} - \gamma_1 \left(\frac{w_i}{p_{1i}} \right) + \frac{1}{2} \beta_{01} p_{1i} \left(\frac{w_i}{p_{1i}} \right)^2 \right)$$

$$(20d) \quad v_{ci}^s = a \left[p_{ci} (e_{ci}^y - e_{1i}^y) - w_i (e_{ci}^x - e_{1i}^x) - e_{ci}^g \right] .$$

These equations have two interesting feature. First, they have additive error terms. Second, these errors terms contain the heterogeneity parameters of the input demand and yield supply functions e_{ci}^x and e_{ci}^y . The logarithm of the acreage shares of the c crops with $c = 2, \dots, C$ is now denotes as $\ln \Delta s_{ci}^-$.

3.3.3 “Complete” multi-crop econometric model

The multi-crop econometric model is composed of three subsets of equations, yield equations, acreage equations and an input allocation equation. The total variable input X is define as the sum of the acreage share devoted to each crop c multiplied by the per hectare variable input quantity used for each crop k : $X_i = \sum_{c=1}^C s_{ci} x_{ci}$. This input variable allocation equation takes part into the econometric model, which is defined as:

$$(21) \quad \begin{cases} y_{ci} = \alpha_{0c} - 0.5 \gamma_c (w_i / p_{ci})^2 + \omega_{ci}^y \\ X_i = \sum_c s_{ci} [\beta_{0c} - \gamma_c (w_i / p_{ci})] + \omega_{ci}^x \\ \ln(\Delta s_{ci}^-) = a [\Delta \pi_{ci}^a - g_c] + \omega_{ci}^s \end{cases}$$

The error terms of the econometric equation systems are provided by:

$$(22) \quad \begin{cases} \omega_{ci}^y = v_{ci}^y = e_{ci}^y + \varepsilon_{ci}^y \\ \omega_{ci}^x = \sum_c s_{ci} [v_{ci}^x] + \eta_i = \sum_c s_{ci} [e_{ci}^x + \varepsilon_{ci}^x] + \eta_i \\ \omega_{ci}^s = v_{ci}^s = a [p_{ci} (e_{ci}^y - e_i^y) - w_i (e_{ci}^x - e_i^x) - e_{ci}^g] \end{cases}$$

An error term η_i is added in the input allocation equation and represents the effects of measurement errors due, *e.g.*, to stock variations.

We denote \mathbf{z}_i the output and input prices for farmer i . The preceding interpretations of the error terms allow to define the following mean assumptions: $E[\omega_{ci}^y / \mathbf{z}_i] = 0$, $E[\varepsilon_{ci} / \mathbf{z}_i] = 0$, $E[e_{ci}^g / \mathbf{z}_i] = 0$, $E[\eta_i / \mathbf{z}_i] = 0$ and $E[s_{ci}' \varepsilon_{ci}^x / \mathbf{z}_i] = 0$. This implies that each component of ω_{ci} has a null expectation conditionally on prices excepted the $s_{ci}' e_{ci}^x$ term in the input allocation equation. The main problem is then $E[s_{ci}' e_{ci}^x / \mathbf{z}_i] \neq 0$ or $E[e_{ci}^x / \mathbf{z}_i, \mathbf{s}_i] \neq 0$ because of the (partial) simultaneity of the input quantity and the

acreage choices. Farmers decide input uses and acreage at the beginning of the production process. Consequently their choices are simultaneous. But we assume that they can adjust their input use given stochastic events (ε_{ci}^x). Thus their choices are partially simultaneous.

These terms need thus to be determined. Before proceeding to the determination of the control functions two remarks are in order. First, the yield supply and the acreage choice functions identify almost the entire set of parameters. Only the term β_{01} can not be identified. Second, the heterogeneity terms $\mathbf{e}_i \equiv (\mathbf{e}_i^y, \mathbf{e}_i^x, \mathbf{e}_i^g)$ are the « interest error terms » for determining the control functions while ε_i and η_i can be viewed as « disturbances ».

3.4 Control function approach

The econometric model considered is fully consistent, *i.e.* consistent with respect to its deterministic parts and with respect to its error terms. It provides thus explicit forms of the relationship between the error term vectors of the yield supply, input demand allocation and acreage equations. The main problem is the link between the acreage and the input use choices in the variable input allocation equation. The control function idea is to determine explicitly this link and its estimator, and integrate this term in the fully multi-crop econometric model.

3.4.1 *Distributional assumptions*

The construction of control functions relies on two main approaches: the use of more or less restrictive distributional assumptions with respect to the error terms and/or the use of the linear projection techniques (see, *e.g.*, Chamberlain, 1982; Wooldridge, 2004).

Both types of approach rely on the additional conditional mean assumptions:

$$(23a) \quad E[\mathbf{e}_i / \mathbf{z}_i] = 0 \quad \text{and} \quad E[\varepsilon_i / \mathbf{z}_i, \mathbf{e}_i] = 0$$

and on the following homoskedasticity assumption:

$$(23b) \quad V[\mathbf{e}_i / \mathbf{z}_i] = \Psi.$$

The conditional mean assumption appears reasonable. The homoskedasticity assumption seems reasonable as far as prices are considered. It is further assumed that $\mathbf{e}_i^z \equiv (\mathbf{e}_i^x, \mathbf{e}_i^y)$ and \mathbf{e}_i^g are not correlated. This assumption is not necessary but it simplifies the approach and may appear empirically reasonable. As a result, the variance-covariance matrix of \mathbf{e}_i has the following structure:

$$(23c) \quad V[\mathbf{e}_i / \mathbf{z}_i] = V[\mathbf{e}_i] = \Psi = \begin{bmatrix} \Psi_{yy} & \Psi_{yx} & 0 \\ \Psi_{yx}' & \Psi_{xx} & 0 \\ 0 & 0 & \Psi_{gg} \end{bmatrix} = \begin{bmatrix} \Psi_{yz} & 0 \\ \Psi_{xz} & 0 \\ 0 & \Psi_{gg} \end{bmatrix} = \begin{bmatrix} \Psi_{zz} & 0 \\ 0 & \Psi_{gg} \end{bmatrix}.$$

The main implications of the assumptions (23) for the control function purpose concern the conditional variance-covariance structure of the error terms of the econometric model:

$$(24a) \quad a^{-2} V[\boldsymbol{\omega}_i^s \boldsymbol{\omega}_i^{s'} / \mathbf{z}_i] = \Delta' \mathbf{M}_i \Psi_{zz} \mathbf{M}_i' \Delta + \Psi_{gg}$$

$$(24b) \quad a^{-1} V[\boldsymbol{\omega}_i^y \boldsymbol{\omega}_i^{s'} / \mathbf{z}_i] = -\Psi_{yz} \mathbf{M}_i' \Delta$$

$$(24c) \quad a^{-1} V[\boldsymbol{\omega}_i^x \boldsymbol{\omega}_i^{s'} / \mathbf{z}_i] = -\Psi_{xz} \mathbf{M}_i' \Delta$$

where $\boldsymbol{\omega}_i^s$, $\boldsymbol{\omega}_i^y$ and $\boldsymbol{\omega}_i^x$ are the vector of error terms $\omega_{ci}^s, \omega_{ci}^y$ and ω_{ci}^x for $c=1, \dots, C$. Matrix \mathbf{M}_i and Δ are defined such as $\omega_{ci}^s = a \left[p_{ci} (e_{ci}^y - e_1^y) - w_i (e_{ci}^x - e_1^x) - e_{ci}^g \right]$ is denoted in matrix notations by : $\boldsymbol{\omega}_i^s = a \left[\Delta' \mathbf{M}_i \mathbf{e}_i^z - \mathbf{e}_i^g \right]$ (see appendix A for more details). These second order moment conditions can be used to define regression estimators of the useful parts of the variance-covariance matrix Ψ (see section on the implementation of the approach).

The linear projection techniques combined with limited assumptions on the distribution of the heterogeneity terms can be used in some special cases. However, distributional assumptions are the most frequent basis for determining control functions (see, *e.g.*, Imbens and Wooldridge, 2007). Due to its properties with respect to the conditioning operation, the normal distribution appears to be a « convenient » choice. We assume thus that \mathbf{e}_i is jointly normal conditional on \mathbf{z}_i , which denotes the input and output prices of farmer i . Since the error terms ω_i^y and ω_i^s are linear transformations of \mathbf{e}_i , they are thus normally distributed conditionally on \mathbf{z}_i and \mathbf{e}_i . The input allocation equation error term ω_i^x is normal distributed conditionally on \mathbf{z}_i , \mathbf{e}_i and \mathbf{s}_i .

3.4.2 Control function approach under normality assumptions

Two options can be used to determine control functions. In the first case, we need to compute the terms $E[\mathbf{s}_i \mathbf{e}_i^x / \mathbf{z}_i]$. In the second case, we need to compute $E[\mathbf{e}_i^x / \mathbf{z}_i, \mathbf{s}_i]$. The non-linearity in \mathbf{e}_i of \mathbf{s}_i in acreage shares equation (19) is the main problem for the first option. This option relies on less restrictive assumptions but more involved computations. Thanks to the equivalence of the conditioning by \mathbf{s}_i and by $\ln \Delta \mathbf{s}_i$, the non-linearity in \mathbf{e}_i of \mathbf{s}_i is not a problem in the second option. The important facts are that the log-linear transformation of acreage shares (20) are linear in the elements of \mathbf{e}_i . The problem is then to determine:

$$(25) \quad E[\mathbf{e}_i^x / \mathbf{z}_i, \mathbf{s}_i, s_{li}] = E[\mathbf{e}_i^x / \mathbf{z}_i, \ln \Delta \mathbf{s}_i, s_{li}].$$

The conditioning properties of normally distributed vectors and the zero conditional mean of \mathbf{e}_i^x , ω_i^y , \mathbf{e}_i^y and ω_i^s allow then to show that :

$$(26a) \quad E[\mathbf{e}_i^x / \mathbf{z}_i, \mathbf{s}_i, s_{li}] = a^{-1} \Psi_{xz}' C_i (\Psi) \omega_i^s$$

with :

$$(26b) \quad C_i(\Psi) = \mathbf{M}_i' \Delta \left[\Delta' \mathbf{M}_i \Psi_{zz} \mathbf{M}_i \Delta + \Psi_{gg} \right]^{-1}.$$

This function $a^{-1} \Psi_{xz}' C_i \omega_i^s$ can be used as control functions in the input demand allocation equations.

3.4.3 A three-stage inference procedure

This section considers the implementation of the control function approach. It presents a simple three-stage inference procedure. The description of the procedure mainly focuses on identification and consistencies issues. It only briefly considers the efficiency issues.

In the first stage the equation system composed of the yield supply and acreage choice equations is estimating. The objective is to construct a consistent estimator of all identifiable parameters θ , *i.e.* all the parameters except β_{01} . This system is a Seemingly Unrelated Regression (SUR) equation system:

$$(27) \quad \begin{cases} y_{ci} = \alpha_{0c} - 0.5 \gamma_c (w_i / p_{ci})^2 + \omega_{ci}^y, & \text{for } c = 1, \dots, C \\ \ln(\Delta s_{ci}^-) = a(\Delta \pi_{ci}^a - g_{0c}) + \omega_{ci}^s, & \text{for } c = 2, \dots, C \end{cases}$$

In the second stage, these estimators $\hat{\theta}$ are assumed to be available for constructing a consistent estimator of an useful part of the variance-covariance matrix Ψ . This stage is similar to the second stage of the construction of a standard Generalized Least Square (GLS) estimator. It relies on the second order moment conditions (defined equations (24)) and uses the (linear in its parameters) SUR system:

$$(28) \quad \begin{cases} a^{-2} \omega_i^s(\hat{\theta}) \omega_i^s(\hat{\theta})' = \Delta' \mathbf{M}_i \Psi_{zz} \mathbf{M}_i \Delta + \Psi_{gg} + \xi_i^{ss} \\ a^{-1} \omega_i^s(\hat{\theta}) \omega_i^y(\hat{\theta})' = \Psi_{yz} \mathbf{M}_i \Delta + \xi_i^{ys} \end{cases} \quad \text{with} \quad \begin{cases} E[\xi_i^{ss} / \Delta \mathbf{M}_i'] = 0 \\ E[\xi_i^{ys} / \Delta \mathbf{M}_i'] = 0 \end{cases},$$

where ξ_i^{ss} and ξ_i^{ys} are error terms. The estimates of variance-covariance matrix $\hat{\Psi}$ are used to construct the control functions $a^{-1} \Psi_{xz}' C_i (\hat{\Psi}) \omega_i^s (\hat{\theta})$. The elements of matrix Ψ_{xz} are thus re-estimated in the third stage.

The third stage of the procedure considers the estimation of all the interest parameters a_0, b_0, g_0 and a and the auxiliary parameters of the matrix Ψ_{xz} . The corresponding estimating equation system uses the control functions:

$$(29) \quad \begin{cases} y_{ci} = \alpha_{0c} - 0.5 \gamma_c (w_i / p_{ci})^2 + \omega_{ci}^y, \text{ for } c = 1, \dots, C \\ X_i = \sum_c^C s_{ci} [\beta_{0c} - \gamma_c (w_i / p_{ci})] + [a^{-1} \mathbf{s}_i' \Psi_{xz}' C_i (\hat{\Psi}) \omega_i^s (\hat{\theta})] + \mu_i^x, \text{ for } c = 1, \dots, C, \\ \ln(\Delta s_{ci}^-) = a(\Delta \pi_{ci}^a - g_{0c}) + \omega_{ci}^s, \text{ for } c = 2, \dots, C \end{cases}$$

with $E[\mu_i^x / \mathbf{z}_i, \mathbf{s}_i] = 0$, $E[\omega_{ci}^y / \mathbf{z}_i, \mathbf{s}_i] = 0$ and $E[\omega_{ci}^s / \mathbf{z}_i] = 0$. The Generalized Method of Moments (GMM) estimator is used to construct consistent estimator. A few remarks are in order for the implementation of this final stage. First, the estimator $\hat{\theta}$ in the first stage provides a useful set of starting values for the empirical implementation of the GMM estimator of the parameters in the third stage. Second, this approach can be interpreted as a generalized version of the « augmented regression » technique controlling for the endogeneity of explanatory variables in models linear in their explanatory variables. The augmented regression test can be used to test the endogeneity of \mathbf{s}_i in the input demand allocation equations. The null hypothesis is then $\Psi_{xz} = 0$. This is a test of the interest of the approach proposed in this study.

3.5 Empirical application

The approach described in this article aims to determine the link between error terms accounting for farms' heterogeneity and the acreage choices. The interest of this empirical application is twofold. First, two models are estimated: one with control function and one

without control function. The first objective is to evaluate the effects of acreage choice endogeneity on parameters and variable inputs allocation. Second, two others models are estimated: one with explanatory variables which control farm heterogeneity, and one without these explanatory variables. The second objective is to show that the biases due the endogeneity of acreage shares are reduced by the use of « imperfect » control variables.

3.5.1 Data

The three-stage procedure is applied to the French grain crop producer over 1989-2006 using rotating panel data sample of the French Farm Accountancy Data Network (FADN). It contains approximately 6000 observations. The information available is acreage, yield and price for each crop, subventions for cereals, oilseeds and protein crops, and variable input expenditures at the farm level. Six different crops group are considered: wheat, other cereals (mainly barley and corn), oilseeds (mainly rapeseed) and protein crops (mainly peas), sugar beets, potatoes and miscellaneous crops, and fodder crops¹⁸. Acreages of sugar beets, potatoes and miscellaneous crops, and fodder crops were considered as exogenous since most of them are contract crops. The different variable inputs (fertilizers, pesticides, seeds) are aggregated into a single variable input for simplicity. The corresponding price index is obtained from French agricultural statistics. All economic quantities are defined in € of 2000.

Some variable were introduced in the yield and input use equations to account for technical changes and farms' heterogeneity. In particular, a « production potential index » is included to control for farm heterogeneity. This index is defined by :

$$(30) \quad q_{it} = (y_{li,t-1} - y_{li,t-1}^{Med}) / (y_{li,t-1}^{Max} - y_{li,t-1}^{Min}),$$

where $y_{li,t-1}^{Med}$, $y_{li,t-1}^{Max}$ and $y_{li,t-1}^{Min}$ denote, respectively, the median, 99% quantile and 1% quantile of the yield of wheat in the sample year $t-1$. It is based on wheat yields due to the

¹⁸ All farmers of the sample cultivate wheat, other cereals, and oilseeds and/or protein crop.

specialization of the sampled farms, and it is defined on a year per year basis to control for year specific conditions. While this index mostly accounts for persistent production conditions, farmers' choices and yields also depend on crop rotation effects. The lagged acreage shares of root crops are introduced to account for the beneficial effects of the induced crop rotations. Since considered crops are aggregated, the acreage share of cereals except corn (in the total acreage of cereals except wheat) and the acreage share of protein crops (in the total acreage of oilseeds and protein crops) are also introduced.

3.5.2 Main results

Table 1 presents the estimates of the yield, input demand and acreage shares equations with the control function approach and without this approach. As discussed above, several explanatory variables are included in models to control farm heterogeneity: the production index, which represents the soil quality of each farm; the acreage share of protein crops in the oilseeds and protein crops aggregate and the acreage share of cereals in the cereals and corn aggregate; and the acreage share of sugar beets and potatoes of the preceding year.

The R^2 obtained with these two approaches are similar. They are not completely satisfactory, since they range from 11% to 42%, although this is a common result when dealing with farm data. However, almost 90% of the parameters are statistically significant at the 5% level. The parameter estimates in yield equations are also quite similar between these two approaches: the average potential yield changes from 8.83 to 8.73 € (in year 2000) per are for wheat, from 8.10 to 8.31 for the other cereals and is to 6.77 for oilseeds and protein crops. Note that production index and root crop acreages have positive effects on potential yield.

The main change generated by the use of control function concerns the input allocation equation: the average input requirement to obtain potential yield changes from 6.03 to 5.20 € (in year 2000) per are for wheat and from 4.33 to 5.78 for oilseeds and protein crops. This means that a correlation exists between acreage choices decisions and first,

input use for wheat and second, input use for oilseeds and protein crops. When these correlations are not controlled, then input use requirement for wheat is underestimated and the input use requirement for oilseeds and protein crops is overestimated.

Note that explanatory variables used to control farm heterogeneity have interesting effects on input requirement. Production index effects are positive on input use, indicating an intensification process. Results confirm also simple agronomic rules: cereals require less variable inputs than corn, and protein crops less variable inputs than oilseeds; sugar beets and potatoes are generally cultivated as preceding crops to wheat, that allow to decrease its input requirement. On the other hand, sugar beets and potatoes acreages have a positive effect on input use for oilseeds and protein crops¹⁹.

The other important change caused by the use of control function concerns the parameter estimates representing fixed costs in acreage equations. These parameters must be interpreted with precaution because they represent the difference between the fixed costs of the considered crop and wheat. This parameter for oilseeds and protein crops changes from 4.10 to 1.92 € par are.

A Wald test is used to test the joint significance of the elements of the matrix Ψ_{xz} . This test is significant, then we reject the null hypothesis $\Psi_{xz} = 0$ and conclude that the parameters are jointly statistically different from zero.

Table 2 presents the same estimation methods as above but explanatory variables used in the preceding estimation to control farm heterogeneity are removed. The omission of these explanatory variables should lead to an increase of error terms accounting for farms' heterogeneity, and consequently accentuate the correlation between these error terms and acreage choices. In this case, the importance of control function approach use should be more evident. Results show that with the standard regression method, the omission of the considered explanatory variables implies important changes in parameter estimates,

¹⁹ Sugar beets and rapeseed, which is the main oilseed in France, are not generally cultivated on the same plot because they are both preceding crops to wheat and they face with the same pest infestation. However, because of the increasing needing of biofuel and the complementary of these crops, these type of rotation recently develop in the north of France and can be explained the increasing input requirement for oilseeds and protein crops.

especially in the average input requirement. This parameter decreases by 17% for the other cereals and by 29% for oilseeds and protein crops, whereas it decreases only by 5% and 4% respectively, with the control function approach. This means that the control functions have captured correlations between error terms accounting for farms' heterogeneity (due in part to omitted variables) and acreage decisions in yield and input demand equations. These results are confirmed by the estimates of the control functions effects presented in the table 3. These parameter estimates correspond to the elements of the variance-covariance matrix of production model's residuals. Results show that more parameters are statistically significant at the 5% level when the farm heterogeneity is not well-controlled, and their values are higher.

All these results confirm that (1) correlations between farms' heterogeneity error terms and acreage decisions exist in input demand equation, and can biased parameter estimates; (2) the control function approach can reduce bias associated with these correlations. This problem is even more important that a lot of factors known by farmer at the time he chooses his acreages (rotation effects, soil quality, quasi-fixed input availabilities...) are not available in standard database and/or are complex to evaluate.

Conclusion

The contribution of this research is threefold. First it shows that the standard regression based approaches for allocating variable input uses to crops are potentially biased to the (partial) simultaneity of the (expected) crop variable input and acreage choices. Second, it proposes a structural econometric multi-crop model, *i.e.* a model which is consistent in its deterministic and random parts, for determining the origin of these biases and providing potential solutions. Third, it proposes an approach based on the use of control functions to eliminate these biases. The empirical application shows that the econometric multi-crop model used is well-behaved and provide interesting results, and it confirms the acreage

endogeneity in input allocation equations and thus, shows the usefulness of the proposed approach.

The proposed approach is described within the context of crop production but could be applied in other contexts where inputs need to be allocated to activities. It could also be applied by using other structural econometrics models with an explicit specification of (deterministic and random) links between production, input uses and activity level choices. Note however that the error term additivity plays a crucial role in the proposed approach.

The proposed approach has potentially three main drawbacks. First, as it is « fully » structural it is thus subject to specification biases. A potential useful extension would replace the structural activity choice model by a more flexible model of the expected gross margin. The second drawback is linked to the first: the econometric model used cannot account for corner solutions of activity choices. This is a potentially important weakness of this framework, particularly in the crop production context. But, the specification of a fully structural model for activity choices with corner solutions is an involved exercise. This highlights the usefulness of « acceptable approximations » to replace a fully structural framework. Third, the identification of the control functions relies on models of the square and cross products of the crop and input prices. As a result, the empirical identification of these functions requires price data at the farm level of good quality.

References

- Antle, J.M. 1983. "Sequential Decision Making in Production Models." *American Journal of Agricultural Economics* 65:282-290.
- Antle, J.M. and S.A. Hatchett. (1986). "Dynamic Input Decisions in Econometric Production Models." *American Journal of Agricultural Economics* 68:939-949.
- Bel, F., Lacroix, A., Salanié, F. and A. Thomas. 2006. "Evaluating the Impact of CAP Reforms on Land Use and the Environment: a Two-Step Estimation with Multiple Selection Rules and Panel Data", Working paper.
- Bel Haj Hassine and M. Simioni. 2000. "Estimation of Two-Stage Models of Multicrop Production: With and Application to Irrigated Water Allocation in Tunisian Agriculture" *Région et Développement*,12, 2000, p. 121-141
- Blackorby, C., Primont, D. and R.R. Russell. 1978. *Duality, Separability, and Functional Structure: Theory and Economic Implications*. Elsevier, North-Holland:New-York.
- Blundell, R. and S. Bond. 2000. "GMM Estimation with Persistent Panel Data: An Application to Production Functions," *Econometric Reviews* 19, 321-340.
- Carpentier, A. and R.D. Weaver. 1997. "Damage Control Productivity: Why specification matters." *American Journal of Agricultural Economics* 52:11-22.
- Carpentier, A. and R.D. Weaver. 1996. "Intertemporal and Interfirm Heterogeneity: Implications for Pesticide Productivity." *Canadian Journal of Agricultural Economics* 44:219-236.
- Coyle, B. T. 1992. "Risk Aversion and Price Risk in Duality Models of Production: A Linear Mean-Variance Approach." *American Journal of Agricultural Economics* 74:849-859.
- Coyle, B. T. 1999. "Risk Aversion and Price Risk in Duality Models of Production: A Mean-Variance Approach." *American Journal of Agricultural Economics* 81:553-567.
- Chamberlain, G. 1987. "Asymptotic Efficiency in Estimation with Conditional Moment Restrictions", *Journal of Econometrics* 34, 305-334.
- Chamberlain, G. 1982. "Multivariate Regression Models for Panel Data." *Journal of Econometrics* 18:5-46.

- Chambers, R. and R. Just. 1989. "Estimating Multioutput Technologies." *American Journal of Agricultural Economics* 71:980-95.
- Dixon, B.L., Batte, M.T. and S.T. Sonka. 1984. "Random Coefficients Estimation of Average Total Product Costs For Multiproduct Firms." *Journal of Business Economics and Statistics* 2: 360-366.
- Dixon, B. and R.H. Hornbaker. 1992. "Estimating the Technology Coefficients in Linear Programming Models." *American Journal of Agricultural Economics* 74:1029-1038.
- Griliches, Z. et J. Mairesse. 1995 "Production Functions: the Search for Identification" National Bureau of Economic Research, Cambridge, MA, NBER Working Paper No. 5067.
- Guyomard, H., Baudry, M. and A. Carpentier. 1996. "Estimating Crop Supply Response in the Presence of Farm Programmes: Application to the CAP." *European Review of Agricultural Economics* 23:401-420.
- Hausman, J. (1978) "Specification Tests in Econometrics," *Econometrica*, 46, 1251-1271.
- Heckelei, T. and H. Wolff. 2003. "Estimation of Constrained Optimisation Models for Agricultural Supply Analysis Based on Generalised Maximum Entropy." *European Review of Agricultural Economics* 30:27-50.
- Heckman, J. J. 2008. "Econometric Causality." *Cemmap working paper* CWP1/08.
- Heckman, J. J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47: 153-161.
- Heckman, J. J. 1974. "Shadow Prices, Market Wages, and Labor Supply." *Econometrica* 42: 679-693.
- Heckman, J.J., Tobias, J.L. and E. Vytlacil. 2003. "Simple Estimators for Treatment Parameters in a Latent-Variable Framework." *Review of Economics and Statistics* 85: 748-755.
- Heckman, J. J. and E. Vytlacil. 1998. "Instrumental Variables Methods for the Correlated Random Coefficient Model." *Journal of Human Resources* 33:974-987.

- Hornbaker, R. H., Dixon, B. L. and S. T. Sonka. 1989. "Estimating Production Activity Costs for Multioutput Firms with a Random Coefficient Regression Model." *American Journal of Agricultural Economics* 71:167-177.
- Imbens, G. and J. M. Wooldridge. 2007. "Control Functions and Related Methods" Lecture notes NBER Summer Institute '07, What's new in econometrics?, NBER.
- Just, R.E., Zilberman, D. and E. Hochman. 1983. "Estimation of Multicrop Production Functions." *American Journal of Agricultural Economics* 65:770-780.
- Just, R.E., Zilberman, D., Hochman, E. and Z. Bar-Shira. 1990. "Input Allocation in Multicrop Systems." *American Journal of Agricultural Economics* 72:200-209.
- Lence, S.H. and D.J. Miller. 1998a. "Estimation of Multi-Output Production Functions with Incomplete Data: A Generalised Maximum Entropy Approach." *European Review of Agricultural Economics* 25:188-209.
- Lence, S.H. and D.J. Miller. "Recovering Output-Specific Inputs from Aggregate Input Data: A Generalized Cross-Entropy Approach." *American Journal of Agricultural Economics* 80:852-867.
- McElroy, M.B. 1987. "Additive General Error Models for Production, Cost, and Derived Demand or Share systems." *Journal of Political Economy* 95:737-757.
- Moore, M.R. and A. Dinar. 1995. "Water and Land as Quantity-Rationed Inputs in California Agriculture: Empirical Tests and Water Policy Implications." *Land Economics* 74(4):445-461.
- Moore, M.R., Gollehon, N.R. and M.B. Carey. 1994. "Multicrop Production Decisions in Western Irrigation Agriculture: The Role of Water Price." *American Journal of Agricultural Economics* 11:143-158.
- Moore, M. R. and D.H. Negri. 1992. "A Multicrop Production Model of Irrigated Agriculture, Applied to Water Allocation Policy of the Bureau of Reclamation." *Journal of Agricultural and Resource Economics* 17:29-43.
- Moro, D. and P. Sckokai. 1999. "Modelling the CAP Arable Crop Regime in Italy: Degree of Decoupling and Impact of Agenda 2000." *Cahiers d'économie et sociologie rurales* 53:50-73.

- Newey, W. and D. McFadden. 1994. "Large Sample Estimation and Hypothesis Testing", Chapter 36 Handbook of Econometrics, Vol 4, McFadden and Engle, (eds.), Elsevier, North Holland.
- Oude Lansink, A. and J. Peerlings. 1996. "Farm-specific Impacts of Quantitative Restrictions of N-Fertiliser Use in Dutch Arable Farming." *Journal of Agricultural Economics* 52(3):38-52.
- Ozarem, P.F. and J.A. Miranowski. 1994. "A Dynamic Model of Acreage Allocation with General and Crop-Specific Soil Capital." *American Journal of Agricultural Economics* 76:385-395.
- Paris, Q. 1989. "A Sure Bet on Symmetry." *American Journal of Agricultural Economics* 71:344-351.
- Peeters, L. and Y. Surry. 1993. "Estimating Feed Utilisation Matrices Using a Cost Function Approach." *Agricultural Economics* 9:109-126.
- Petrin, A. and K. E. Train. 2008. "A Control Function Approach to Endogeneity in Consumer Choice Model." *Forthcoming, Journal of Marketing Research*.
- Pope, R. D. and R. E. Just. 2003. "Distinguishing Errors in Measurement from Errors in Optimization." *American Journal of Agricultural Economics* 85:348-358.
- Reiss, P. C. and F. A. Wolak. 2007. "Structural Econometric Modeling: Rationales and Examples from Industrial Organization." Handbook of Econometrics, in: J.J. Heckman and E.E. Leamer (eds), Handbook of Econometrics, Vol 6, Elsevier, North Holland.
- Rivers, D. and Q. H. Vuong. 1988. "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." *Journal of Econometrics* 39, 347-366.
- Shumway, C.R., Pope, R.D. and E.K. Nash. 1984. "Allocatable Fixed Inputs and Jointness in Agricultural Production: Implications for Economic Modeling". *American Journal of Agricultural Economics* 66:72-78.
- Smith, R. J. and R. W. Blundell. 1986. "An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply." *Econometrica* 54, 679-686.

- Wooldridge, J. M. 2008. "Instrumental Variables Estimation of the Average Treatment Effect in Correlated Random Coefficient Models," in *Advances in Econometrics*, Volume 21, Millimet D., J. Smith, and E. Vytlacil (eds.), 93-117. Amsterdam: Elsevier.
- Wooldridge, J.M. 2004. "Estimating Average Partial Effects under Conditional Moment Independence Assumptions." *The Institute for Fiscal Studies, Dpt of Economics, UCL. Cemmap*, WP CWP03/04.
- Wooldridge, J.M. 2002. *The Econometrics of Cross-Sections and Panel Data*. The MIT Press: Cambridge MA.
- Wooldridge, J.M. 1996. "Estimating Systems of Equations with Different Instruments for Different Equations." *Journal of Econometrics* 74, 387-405.

Table 1. Estimates of the Yield, Input Demand and Acreage Shares Equations
(with heterogeneity control).

Explanatory Variable	with function control			without function control		
	Wheat	Other cereals	Oilseeds protein crops	Wheat	Other cereals	Oilseeds protein crops
Yield supply						
Price effects (γ)	2.07***	0.81**	1.81***	1.87***	1.25**	1.84***
Constant	2.22***	0.87**	1.76***	2.02***	1.30***	1.77***
Production index	0.25	-0.51*	3.16***	0.57*	-0.38	3.13***
Average potential yield (α)	8.83***	8.10***	6.77***	8.73***	8.31***	6.77***
Constant	7.91***	8.12***	5.89***	7.83***	8.33***	5.88***
Trend	0.12***	0.05***	0.13***	0.12***	0.04***	0.13***
Production index	2.62***	1.80***	2.75***	2.72***	1.83***	2.74***
Root crop acreage	2.37***	2.01***	0.95***	2.50***	1.93***	1.12***
Aggregate composition	-	-1.10**	0.92***	-	-1.29***	0.87***
Irrigation	-0.04	2.05***	0.22**	-0.04	1.99***	0.87***
<i>R-square</i>	<i>0.42</i>	<i>0.35</i>	<i>0.32</i>	<i>0.42</i>	<i>0.35</i>	<i>0.32</i>
Input demand						
Average input requirement (β)	6.03***	5.67***	4.33***	5.20***	5.68***	5.78***
Constant	6.51***	6.34***	3.60***	5.69***	6.40***	5.01***
Production index	1.03***	0.68*	2.75***	1.41***	0.78*	2.79***
Root crop acreage	-5.68***	-0.99	8.63***	-5.60***	-1.11	8.55***
Aggregate composition	-	-2.90***	-4.23***	-	-2.98***	-4.03***
Irrigation	-0.46	-0.43**	1.16***	-0.35	-0.48**	1.14***
Sugar beets		8.80***			9.14***	
Potatoes		12.09***			11.89***	
Fodder crops		1.97			2.43	
<i>R-square</i>		<i>0.40</i>			<i>0.40</i>	
Acreage shares						
Fixed costs (c)	-	2.75***	4.10***	-	2.43***	1.92***
Cost weight (a)			0.139***			0.140***
<i>R-square</i>	-	<i>0.11</i>	<i>0.19</i>	-	<i>0.11</i>	<i>0.19</i>

Note: (*), (**), and (***) denote parameter estimates statistically different from 0 at, respectively, 10%, 5% and $\leq 1\%$ confidence levels.

**Table 2. Estimates of the Yield, Input Demand and Acreage Shares Equations
(without heterogeneity control).**

Explanatory Variable	with function control			without function control		
	Wheat	Other cereals	Oilseeds protein crops	Wheat	Other cereals	Oilseeds protein crops
Yield supply						
Price effects (γ)	2.21***	0.10	1.44**	2.01***	0.56**	1.19***
Average potential yield (α)	8.87***	7.76***	6.59***	8.79***	7.97***	6.49***
Constant	7.93***	7.20***	6.13***	7.88***	7.28***	6.11***
Trend	0.12***	0.05**	0.13***	0.13***	0.06***	0.11***
Irrigation	0.35 [*]	2.27***	0.34**	0.32	2.17***	0.32***
<i>R-square</i>	<i>0.07</i>	<i>0.12</i>	<i>0.04</i>	<i>0.03</i>	<i>0.12</i>	<i>0.01</i>
Input demand						
Average input requirement (β)	5.72***	5.44***	3.87***	5.56***	4.86***	4.47***
Constant	5.72***	5.47***	3.84***	5.55***	4.88***	4.44***
Irrigation	0.96 [*]	-0.41**	0.53 [*]	1.14***	-0.36**	0.55 [*]
Sugar beets		14.77***			15.14***	
Potatoes		8.51***			8.63***	
Fodder crops		-1.02			-0.30	
<i>R-square</i>		<i>0.35</i>			<i>0.38</i>	
Acreage shares						
Fixed costs (c)	-	3.16***	5.00***	-	4.35***	4.40***
Cost weight (a)		0.109***		0.105***		
<i>R-square</i>	-	<i>0.05</i>	<i>0.01</i>	-	<i>0.05</i>	<i>0.008</i>

Note: (^{*}), (^{**}) and (^{***}) denote parameter estimates statistically different from 0 at, respectively, 10%, 5% and $\leq 1\%$ confidence levels.

Table 3. Estimates of variance-covariance matrix.

	with heterogeneity control			without heterogeneity control		
	Wheat	Other cereals	Oilseeds protein crops	Wheat	Other cereals	Oilseeds protein crops
Yield (v_{cm}^{yy})						
Wheat	1.61**	2.70***	1.72***	0.61	3.20***	3.37***
Other cereals	2.70***	6.38***	4.84***	3.20***	7.07***	6.50***
Oilseeds protein crops	1.72***	4.84***	2.38**	3.37***	6.50***	5.21**
Input use \times Yield (v_{cm}^{yx})						
Wheat	-0.75	0.19	0.69	-5.94***	-2.90**	-2.43**
Other cereals	1.04	4.27***	4.68***	1.52	4.18***	4.74***
Oilseeds protein crops	-0.40	2.75***	2.01*	0.87	3.67***	3.27***
Input use ($v_{cm}^{xx} - v_{11}^{xx}$)						
Wheat	0	0.10	-0.29	0	4.69**	2.04
Other cereals	0.10	5.99***	3.93***	4.69**	11.73***	7.98***
Oilseeds protein crops	-0.29	3.93***	0.33	2.04	7.98***	6.51***

Note 1: (*), (**) and (***) denote parameter estimates statistically different from 0 at, respectively, 10%, 5% and $\leq 1\%$ confidence levels.

Note 2: v_{cm}^{xx} , v_{cm}^{yy} and v_{cm}^{yx} are elements of the variance-covariance matrix Ψ (see appendix A).

Note 3: A control function is included in yield equations in order to test endogeneity of acreages choices.

Appendix A : Matrix notations

$$\mathbf{M}_i = \begin{bmatrix} p_{1i} & & 0 & -w_i & & 0 \\ & \ddots & & & \ddots & \\ 0 & & p_{Ki} & 0 & & -w_i \end{bmatrix} \text{ and } \Delta' \mathbf{M}_i = \begin{bmatrix} -p_1 & p_2 & & 0 & w & -w & & 0 \\ \vdots & & \ddots & & \vdots & & \ddots & \\ -p_1 & 0 & & p_K & w & 0 & & -w \end{bmatrix}.$$

$$\Psi = \begin{bmatrix} \Psi_{zz} & 0 \\ 0 & \Psi_{gg} \end{bmatrix} = \begin{bmatrix} \Psi_{yz} & 0 \\ \Psi_{xz} & 0 \\ 0 & \Psi_{gg} \end{bmatrix} = \begin{bmatrix} \Psi_{yy} & \Psi_{yx} & 0 \\ \Psi_{xy} & \Psi_{xx} & 0 \\ 0 & 0 & \Psi_{gg} \end{bmatrix}$$

$$\Psi_{zz} = \begin{bmatrix} v_{11}^{yy} & \cdots & v_{1C}^{yy} & v_{11}^{yx} & \cdots & v_{1C}^{yx} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ v_{1C}^{yy} & \cdots & v_{CC}^{yy} & v_{C1}^{yx} & \cdots & v_{CC}^{yx} \\ v_{11}^{yx} & \cdots & v_{C1}^{yx} & v_{11}^{xx} & \cdots & v_{1C}^{xx} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ v_{1C}^{yx} & \cdots & v_{CC}^{yx} & v_{1C}^{xx} & \cdots & v_{CC}^{xx} \end{bmatrix}.$$

Chapitre IV

Un Modèle Econométrique Flexible pour l'Allocation de la Terre et des Intrants Variables ²⁰

Résumé

Les modèles micro-économétriques multi-produits, basés sur la fonction de coût implicite de gestion des assolements, développés dans les chapitres précédents permettent de décrire les choix de production des agriculteurs sur le court terme. Ils possèdent des caractéristiques intéressantes pour envisager des collaborations pluridisciplinaires, notamment celles d'avoir une structure simple et des paramètres facilement interprétables. Ils reposent néanmoins sur quelques hypothèses qui peuvent sembler restrictives, et en particulier l'hypothèse de rendement marginal de la terre constant. L'objectif de cet article est de construire un modèle plus général et plus flexible, et qui permet de relâcher deux hypothèses : l'hypothèse de rendement marginal de la terre constant et l'hypothèse sur la distribution des termes d'erreurs. Pour cela, nous proposons de spécifier la fonction de coût implicite sous une forme quadratique et d'intégrer des effets d'échelle dans les équations de rendement. L'approche par les fonctions de contrôle est utilisée pour contrôler l'endogénéité des surfaces mais à partir des techniques de la projection linéaire. Une application empirique est réalisée sur des données d'exploitations françaises de 1989 à 2006.

²⁰ Ce chapitre est basé sur un papier écrit conjointement avec Alain Carpentier.

4.1 Introduction

Le modèle micro-économétrique multi-produits développé dans les chapitres 2 et 3 permet de décrire les choix de production de court terme des agriculteurs. Ce modèle est composé d'équations de rendement, d'une équation de répartition des intrants et d'équations de parts de surface. Une de ses particularités est que la fonction de profit total d'un agriculteur est définie comme étant la somme des marges brutes à l'hectare de chaque culture multipliées par la surface allouée à ces cultures moins une fonction de coût. Cette fonction de coût représente les coûts de gestion implicites d'un assolement. Elle mesure deux effets : 1) l'effet liés aux limites imposées sur les choix d'assolement par les schémas de rotation utilisés par les agriculteurs, 2) et les effets liés aux limites dans les choix d'assolement imposées par la disponibilité des facteurs fixes autres que la terre. Ces effets sont considérés dans nos modèles comme des motifs de diversification des cultures. Cette fonction de coût est également utilisée dans les modèles de programmation mathématique positive pour tenir compte des contraintes auxquelles les agriculteurs font face mais qui ne sont pas observées (Howitt, 1995 ; Heckeleï et Wolff, 2003).

La forme de la fonction de coût implicite utilisée dans les chapitres 2 et 3 est basée sur une fonction d'entropie. Son principal avantage est de permettre la dérivation d'équations de parts de surface de forme Logit multinomial, qui ont une forme simple avec peu de paramètres facilement interprétables. Néanmoins, la dérivation des équations de parts de surface de forme Logit a nécessité de définir les fonctions de rendements et donc les marges brutes de chaque culture comme étant indépendantes des surfaces allouées à chaque culture. Ceci revient à poser l'hypothèse de rendement marginal de la terre constant. Les modèles utilisés généralement pour décrire les choix de production des agriculteurs (Moore et Negri, 1992 ; Guyomard et al., 1996, Moro et Sckokai, 1999) considèrent, au contraire, que les rendements marginaux de la terre sont décroissants. Cette hypothèse traduit l'idée selon laquelle le profit associé à une culture diminue en fonction de la sole qui lui est consacrée. Les arguments avancés en faveur de cette l'hypothèse par les économistes agricoles sont, d'une part, l'hétérogénéité de conditions de

production sur une exploitation, et d'autre part, les effets des rotations culturales. Choisir d'accroître la surface d'une culture contraint l'agriculteur à implanter cette culture sur des parcelles avec des effets de précédents moins bénéfiques, voire négatifs. Ces effets sont considérés dans ces modèles comme des motifs de diversification des cultures.

Un autre avantage du modèle développé dans les chapitres 2 et 3 est qu'il permet de répartir les intrants variables entre les cultures. Cette répartition a nécessité de corriger le problème d'endogénéité des surfaces dans l'équation d'allocation des intrants par les fonctions de contrôle. Le principe de cette méthode est de déterminer explicitement la forme de la corrélation entre les surfaces et le terme d'erreur dans l'équation d'allocation des intrants. Pour cela, l'hypothèse de normalité de ce terme d'erreur a été posée. Cette hypothèse de distribution est souvent nécessaire pour déterminer des fonctions de contrôle dans des modèles multi-output (Imbens et Wooldridge, 2007), mais elle n'est pas automatique. Dans certains cas, les techniques de projection linéaire suffisent à l'application de l'approche par les fonctions de contrôle.

L'objectif de ce travail est de construire un modèle plus flexible et plus général que celui développé dans les chapitres précédents. Pour cela, nous donnons une forme quadratique à la fonction de coût implicite de gestion des assolements. L'intérêt de cette forme fonctionnelle est triple : elle permet de relâcher l'hypothèse de rendement marginal constant de la terre ; l'approche par les fonctions de contrôle peut être appliquée sans hypothèse sur la distribution des termes d'erreur ; elle est plus flexible et les choix de surfaces sont déterminés à partir de plus de paramètres. Bien que les équations de parts de surfaces dérivées de ce modèle soient plus complexes que celles de forme Logit multinomial, ce modèle est plus général et plus flexible.

L'article est organisé comme suit. La première section présente le modèle de choix de production construit à partir d'une fonction de coût implicite de forme quadratique et en intégrant des effets d'échelle dans les fonctions de rendement. La deuxième section

présente l'approche par les fonctions de contrôle utilisée pour corriger le problème d'endogénéité des surfaces dans l'équation d'allocation des intrants, à partir des techniques de la projection linéaire. Dans la troisième section, une application empirique est réalisée sur des données d'exploitations françaises spécialisées en céréales sur la période 1989-2006 (ce sont les mêmes données que les applications précédentes). Plusieurs modèles sont estimés afin d'étudier l'importance des deux motivations utilisées dans ce modèle pour décrire les choix diversifiés des assolements : les effets d'échelle et la fonction de coût.

4.2 Le modèle économétrique multi-produits

La plupart des modèles économétriques construits pour décrire les décisions d'applications d'intrants et/ou les choix d'assolements utilisent l'hypothèse de rendement marginal de la terre décroissant comme incitation pour les agriculteurs à diversifier leurs cultures (Moore et Negri, 1992 ; Guyomard et al., 1996, Moro et Sckokai, 1999). Nous proposons d'élargir le modèle défini dans les chapitres 2 et 3 en intégrant des effets d'échelle dans les fonctions de rendements. Ce modèle tient compte également explicitement de l'hétérogénéité inobservée (par les économistes) des exploitations via des termes d'erreur (McElroy, 1987).

4.2.1 Les fonctions de rendement

Les notations utilisées dans cet article sont : y_{ci} le rendement de la culture c obtenu par l'agriculteur i ; x_{ci} la quantité de l'intrant variable utilisée par hectare par l'agriculteur i pour la culture c ; X_i la quantité totale d'intrant variable utilisée par l'agriculteur i ; s_{ci} la part de surface allouée à la culture c par l'agriculteur i ; p_{ci} et w_i les prix de la culture c et le prix de l'intrant pour l'agriculteur i . Le rendement de la culture c ($c=1, \dots, C$) pour l'exploitation i ($i=1, \dots, N$) est supposé être une fonction quadratique à un seul input (sans perte de généralité):

$$(1a) \quad y_{ci} = \alpha_{ci} - 0.5\gamma_c^{-1} (\beta_{ci} - x_{ci})^2,$$

avec :

$$(1b) \quad \alpha_{ci} = \alpha_{0c} + 0.5\alpha_{1c}s_{ci} + v_{ci}^y \quad \text{et} \quad \beta_{ci} = \beta_{0c} + 0.5\beta_{1c}s_{ci} + v_{ci}^x,$$

avec α_{ci} , γ_c et β_{ci} les paramètres à estimer avec $\alpha_{ci} > 0$, $\gamma_c > 0$ et $\beta_{ci} > 0$. Cette fonction de production est choisie pour sa flexibilité et sa facilité d'interprétation. Cette forme quadratique « translaturée » est également utilisée par Pope et Just (2003). Le paramètre α_{ci} représente le rendement maximum de la culture c pour l'agriculteur i . Le paramètre β_{ci} représente la quantité d'intrant nécessaire pour obtenir le rendement maximum de la culture c pour l'agriculteur i . Le paramètre γ_c contrôle la concavité de la fonction de production ($\gamma_c > 0$).

Le rendement maximum α_{ci} et les intrants nécessaires pour atteindre ce rendement maximum β_{ci} sont spécifiés comme des fonctions de la surface allouée à la culture pour tenir compte des effets d'échelle potentiels α_{0c} et β_{0c} , et de termes aléatoires v_{ci}^y et v_{ci}^x .

Ces effets d'échelle α_{0c} et β_{0c} visent à approximer, les effets des successions culturales, voire de l'hétérogénéité de la terre à l'échelle de l'exploitation, sur les profits des cultures. Ils sont supposés être négatifs. Ils traduisent donc l'idée des rendements marginaux de la terre décroissants. La décroissance du profit des cultures en s_c provient de deux types d'effets. Le premier type d'effets regroupe ceux liés à la structure du sol et à des problèmes sanitaires sans solutions chimiques pour les agriculteurs. Ces effets concernent directement le potentiel de rendement de la plante et sont représentés par le paramètre α_{0c} . Le deuxième type d'effets regroupe ceux liés aux utilisations d'intrants. Les schémas de rotation recommandés visent à exploiter les effets de la culture précédente pour réduire les utilisations d'intrants. S'écarter de ces schémas tend à accroître les quantités d'intrants nécessaires au développement et à la protection de la culture. Ces effets sont donc représentés par β_{0c} .

Les termes v_{ci}^y et v_{ci}^x représentent des termes aléatoires. Ils sont séparés en deux termes d'erreur distincts pour simplifier leur interprétation :

$$(2) \quad v_{ci}^y = e_{ci}^y + \varepsilon_{ci}^y,$$

$$(3) \quad v_{ci}^x = e_{ci}^x + \varepsilon_{ci}^x.$$

Les termes e_{ci}^y et e_{ci}^x représentent les effets de l'hétérogénéité des exploitants et des exploitations sur la production, comme la qualité des sols ou les caractéristiques de l'exploitant. Les termes ε_{ci}^y et ε_{ci}^x représentent les effets liés aux événements stochastiques sur la production, comme des conditions climatiques ou des maladies. Ces termes d'erreur e_{ci} et ε_{ci} ne sont pas observés par l'économètre. Par contre les effets de l'hétérogénéité sur la production e_{ci} sont connus de l'agriculteur quand il prend toutes ses décisions de production. Les événements stochastiques ε_{ci} ne sont pas observés par l'agriculteur au début du processus de production mais le sont pendant le processus. L'espérance des termes ε_{ci} est supposée nulle, et les termes e_{ci} et ε_{ci} sont supposés non corrélés.

Les effets de l'hétérogénéité sont généralement contrôlés dans les fonctions de production par des variables observées (comme des surfaces décalées, des effets fixes ou des caractéristiques du sol) mais sont le plus souvent inclus directement dans les termes d'erreur, qui sont alors ajoutés aux modèles de production de façon *ad hoc* (Just et Pope, 2001). L'hétérogénéité des exploitations est pourtant une des sources majeures de variabilité des niveaux de production agricoles entre les exploitations et entre les périodes (Carpentier et Weaver, 1996). La prise en compte explicite des effets de l'hétérogénéité non-observée dans les fonctions de production est donc importante dans les études agricoles empiriques.

L'agriculteur i , supposé neutre au risque, vend sa culture c au prix p_{ci} et achète son intrant au prix w_i . Les prix sont supposés être connus au début du processus de production. La maximisation de la marge brute de chaque culture sous contrainte technologique conduit aux équations d'offre de productions, de demande d'intrants et de marge brute suivantes :

$$(4) \quad x_{ci} = \beta_{0c} + 0.5\beta_{1c}s_{ci} - \gamma_c (w_i / p_{ci}) + v_{ci}^x$$

$$(5) \quad y_{ci} = \alpha_{0c} + 0.5\alpha_{1c}s_{ci} - 0.5\gamma_c (w_i / p_{ci})^2 + v_{ci}^y$$

$$(6) \quad \pi_{ci}^e = p_{ci}\alpha_{0c} - w_i\beta_{0c} + 0.5(p_{ci}\alpha_{1c} - w_i\beta_{1c})s_{ci} + 0.5\gamma_c (w_i / p_{ci})^2 + p_{ci}e_{ci}^y - w_i e_{ci}^x$$

L'agriculteur est supposé pouvoir ajuster son utilisation d'intrant variable pendant le processus de production, c'est-à-dire après avoir observé ε_{ci}^y et ε_{ci}^x . Par contre, la marge brute espérée pour chaque culture ne dépend pas des termes ε_{ci}^y et ε_{ci}^x parce que l'agriculteur ne connaît pas au début du processus de production les événements stochastiques qui vont se produire. Le système composé des équations de rendement, des demandes d'inputs variables et des marges brutes espérées est défini en forme matricielle pour simplifier l'écriture du modèle complet décrit par la suite (voir annexe A) :

$$(7) \quad \mathbf{y}_i = \mathbf{a}_0^y + 0.5\mathbf{B}_0^y \mathbf{s}_i + \mathbf{v}_i^y$$

$$(8) \quad \mathbf{x}_i = \mathbf{a}_0^x + 0.5\mathbf{B}_0^x \mathbf{s}_i + \mathbf{v}_i^x$$

$$(9) \quad \pi_i^e = \mathbf{M}_i \left[\mathbf{a}_0 + 0.5\mathbf{B}_0 \mathbf{s}_i + \mathbf{e}_i \right]$$

avec $\mathbf{a}_0 \equiv (\mathbf{a}_0^y, \mathbf{a}_0^x)$, $\mathbf{B}_0 \equiv (\mathbf{B}_0^y, \mathbf{B}_0^x)$ et $\mathbf{e}_i \equiv (\mathbf{e}_i^y, \mathbf{e}_i^x)$. Les paramètres indicés par 0 sont les paramètres à estimer. Les paramètres \mathbf{a}_0^y et \mathbf{a}_0^x peuvent être définis comme des fonctions de la matrice \mathbf{z}_i , qui correspond aux prix des outputs et des inputs. Les paramètres \mathbf{B}_0^y et

\mathbf{B}_0^x sont les matrices des effets d'échelle. La matrice \mathbf{M}_i est définie telle que $\mathbf{M}_i \mathbf{e}_i = \mathbf{P}_i' \mathbf{e}_i^y - w_i \mathbf{e}_i^x$ avec \mathbf{P}_i la matrice des prix des outputs.

4.2.2 Les fonctions de parts de surface

Les choix d'assolement des agriculteurs sont décrits à partir du modèle développé par Heckelevi et Wolff (2003). L'agriculteur i est supposé allouer sa quantité totale de terre entre les cultures en maximisant la fonction de profit totale suivante :

$$(10) \quad \sum_{c=1}^C s_{ci} \pi_{ci}^e(s_i) - C(s_i).$$

Selon ce modèle, l'agriculteur a deux raisons de diversification des assolements : les effets d'échelle $0.5(p_{ci} \alpha_{1c} - w_i \beta_{1c}) s_{ci}$ dans les marges brutes et le coût implicite de gestion des assolements $C(s_i)$. Cette fonction de coût représente des contraintes de disponibilité du travail et du capital et des contraintes agronomiques, auxquelles les agriculteurs font face mais qui ne sont pas observées. De telles fonctions de coût sont utilisées dans les modèles de choix d'assolement de la programmation mathématique positive (Howitt, 1995) et ont été utilisées dans le cadre de modélisations économétriques par Arnberg et Hansen (2007). Cette fonction de coût est supposée avoir une forme quadratique :

$$(11a) \quad C(s_i) = g + \sum_{c=1}^C s_{ci} g_{ci} + 0.5 \sum_{c=1}^C \sum_{m=1}^C g_{cm} s_{ci} s_{mi}$$

avec :

$$(11b) \quad g_{ci} = g_{0c} + e_{ci}^g,$$

avec g, g_{ci} et g_{cm} les paramètres à estimer. Cette fonction de coût est convexe dans les surfaces (si la matrice composée des paramètres g_{cm} est symétrique et définie positive). Le paramètre g_{ci} représente le coût fixe (hors intrants variables) par unité de surface de la culture c pour l'agriculteur i . Il est séparé en deux parties : un paramètre g_{0c} et un terme

aléatoire e_{ci}^g , qui tient compte des coûts liés à l'hétérogénéité, connu de l'agriculteur i mais non de l'économètre. Le paramètre g_{cm} indique si les cultures sont concurrentes ou non en termes d'utilisation des intrants fixes. Si la culture c et la culture m sont concurrentes, alors le paramètre g_{cm} est positif. L'agriculteur est donc incité à diversifier son assolement, ou plus précisément à choisir des cultures qui n'ont pas les mêmes besoins de travail et de matériel au même moment. Si ces cultures ne sont pas concurrentes, alors le paramètre g_{cm} est nul. Ce sont alors seulement les rendements d'échelle qui incitent les agriculteurs à diversifier leurs cultures.

La fonction de profit est réécrite en intégrant directement la contrainte de disponibilité de la terre $1 = \sum_{c=1}^C s_{ci}$ et en écriture matricielle (par souci de simplicité) :

$$(12) \quad \mathbf{f}_i + \mathbf{s}_i^- ' \left[\Delta ' \mathbf{M}_i [\mathbf{a}_0 + \mathbf{e}_i] - [\mathbf{g}_0 + \mathbf{e}_i^g] - \mathbf{F}_i \right] + 0.5 \mathbf{s}_i^- ' \mathbf{Q}_0 \mathbf{s}_i^- ,$$

avec \mathbf{s}^- le vecteur des surfaces de dimension $C-1$, c'est-à-dire sans la culture de « référence », qui est supposée être la culture 1. La matrice Δ est une matrice de différentiation telle que $\Delta'q$ (par exemple) soit égale à un vecteur-colonne de dimension $C-1$, avec comme éléments $q_c - q_1$ pour $c=1, \dots, C$. Le vecteur \mathbf{g}_0 correspond aux paramètres g_{0c} pour $c=2, \dots, C$ et le vecteur \mathbf{e}_i^g contient les termes d'erreurs e_{ci}^g pour $c=2, \dots, C$. Le terme \mathbf{Q}_0 est défini par $\mathbf{A}(\mathbf{b}_0) - \Delta' \mathbf{G}_0 \Delta$. La matrice \mathbf{G}_0 contient les paramètres g_{cm} et la matrice $\mathbf{A}(\mathbf{b}_0)$ est une fonction des paramètres β_{1c} . Le terme \mathbf{F}_i contient les effets d'échelle des marges brutes sur les choix de surfaces. Tous ces termes sont définis dans l'annexe A. La fonction de profit est strictement concave en \mathbf{s} si la forme quadratique $\mathbf{s}_i^- ' \mathbf{Q}_0 \mathbf{s}_i^-$ est semi-définie négative.

La maximisation en \mathbf{s} de cette fonction de profit conduit aux équations de parts de surfaces suivantes :

$$(13a) \quad \mathbf{s}_i^- = -\mathbf{Q}_0^{-1} [\Delta' \mathbf{M}_i \mathbf{a}_0 - \mathbf{g}_0 - \mathbf{F}_{1i}] + \mathbf{v}_i^s$$

avec :

$$(13b) \quad \mathbf{v}_i^s = -\mathbf{Q}_0^{-1} [\Delta' \mathbf{M}_i \mathbf{e}_i - \mathbf{e}_i^g].$$

Ces équations de parts de surface sont plus complexes que celles de forme Logit multinomial dérivées dans les chapitres précédents, mais elles sont plus flexibles. La forme de ces équations de parts de surfaces a deux autres intérêts : elle a des termes d'erreurs additifs, et ces termes d'erreur contiennent les paramètres d'hétérogénéité des fonctions d'offre d'outputs et de demande d'input \mathbf{e}_i^x et \mathbf{e}_i^y .

4.2.3 Le modèle économétrique multi-produits « complet »

Le modèle économétrique multi-output est composé de trois ensembles d'équations les équations de rendements, les équations de surfaces et l'équation de répartition de l'intrant entre les cultures. Le système complet est défini par :

$$(14) \quad \begin{cases} \mathbf{y}_i = \mathbf{a}_0^y + 0.5 \mathbf{B}_0^y \mathbf{s}_i + \boldsymbol{\omega}_i^y \\ X_i = \mathbf{s}_i' [\mathbf{a}_0^x + 0.5 \mathbf{B}_0^x \mathbf{s}_i] + \boldsymbol{\omega}_i^x \\ \mathbf{s}_i^- = -\mathbf{Q}_0^{-1} [\Delta' \mathbf{M}_i \mathbf{a}_0 - \mathbf{g}_0 - \mathbf{F}_{1i}] + \boldsymbol{\omega}_i^s \end{cases}$$

avec les termes d'erreur :

$$(15) \quad \begin{cases} \boldsymbol{\omega}_i^y = \mathbf{v}_i^y = \mathbf{e}_i^y + \boldsymbol{\varepsilon}_i^y \\ \boldsymbol{\omega}_i^x = \mathbf{s}_i' \mathbf{e}_i^x + \mathbf{s}_i' \boldsymbol{\varepsilon}_i^x + \eta_i \\ \boldsymbol{\omega}_i^s = \mathbf{v}_i^s = -\mathbf{Q}_0^{-1} [\Delta' \mathbf{M}_i \mathbf{e}_i - \mathbf{e}_i^g] \end{cases}$$

avec $\Delta' \mathbf{M}_i \mathbf{e}_i - \mathbf{e}_i^g = \boldsymbol{\omega}_i^\pi$. Un terme d'erreur η_i est ajouté à l'équation de répartition des intrants variables et représente les effets des erreurs de mesure, dus au stock, par exemple.

L'interprétation des paramètres permet de poser les hypothèses suivantes : $E[\boldsymbol{\omega}_i^y / \mathbf{z}_i] = 0$, $E[\boldsymbol{\varepsilon}_i / \mathbf{z}_i] = 0$, $E[\mathbf{e}_i^s / \mathbf{z}_i] = 0$, $E[\eta_i / \mathbf{z}_i] = 0$ et $E[\mathbf{s}_i' \boldsymbol{\varepsilon}_i^x / \mathbf{z}_i] = 0$.

Deux problèmes sont à résoudre dans ce modèle. Le premier est l'endogénéité des \mathbf{s}_i dans les équations de rendements. Ce problème peut être contrôlé par la méthode des Variables Instrumentales (VI). Le second est l'endogénéité des \mathbf{s}_i dans l'équation de répartition des intrants. La méthode des VI ne peut pas être utilisée parce que le terme d'erreur $\mathbf{s}_i' \boldsymbol{\varepsilon}_i^x + \mathbf{s}_i' \boldsymbol{\varepsilon}_i^s + \eta_i$ contient la variable explicative endogène \mathbf{s}_i . La condition $E[\mathbf{s}_i' \boldsymbol{\varepsilon}_i^x / \mathbf{z}_i] = 0$ est vérifiée puisque les agriculteurs ne connaissent pas les événements $\boldsymbol{\varepsilon}_i^x$ au moment des choix d'assolements. Par contre, la condition d'exogénéité $E[\mathbf{s}_i \boldsymbol{\varepsilon}_i^x / \mathbf{z}_i] = 0$ n'est vérifiée que si les variables \mathbf{z}_i contrôlent « parfaitement » l'hétérogénéité de x_{ci} . Or dans les bases de données standards, l'information contenue dans les variables \mathbf{z}_i n'est généralement pas suffisante pour satisfaire cette condition. La seule méthode adaptée est donc l'approche par les fonctions de contrôle. Le principe de cette approche est de déterminer explicitement ce terme $E[\mathbf{s}_i \boldsymbol{\varepsilon}_i^x / \mathbf{z}_i] = 0$ ou $E[\boldsymbol{\varepsilon}_i^x / \mathbf{z}_i, \mathbf{s}_i]$ et de l'intégrer dans l'équation d'allocation des intrants pour contrôler l'endogénéité des \mathbf{s}_i .

4.3 L'approche par les fonctions de contrôle

L'approche par les fonctions de contrôle est utilisée dans le cas d'endogénéité de variables explicatives (Imbens et Wooldridge, 2007 ; Petrin et Train, 2008) ou dans le cas de sélection des échantillons (Heckman, 1974, 1979). La construction des fonctions de contrôle s'appuie sur deux méthodes principales : l'utilisation d'hypothèse sur la distribution des termes d'erreurs et/ou l'utilisation des techniques de la projection linéaire (Chamberlain, 1982; Wooldridge, 2004). Les hypothèses de distribution sont généralement nécessaires pour les modèles économétriques multi-outputs (Imbens and Wooldridge, 2007). L'hypothèse de normalité des termes d'erreurs du modèle a été posée pour déterminer les fonctions de contrôle dans le chapitre 3.

4.3.1 Hypothèses de distribution

Les deux types d'approches nécessitent de poser des hypothèses sur l'espérance conditionnelle et une hypothèse d'homoscédasticité : $E[\mathbf{e}_i / \mathbf{z}_i] = 0$, $V[\mathbf{e}_i / \mathbf{z}_i] = \Psi$ et $E[\mathbf{e}_i / \mathbf{z}_i, \mathbf{e}_i] = 0$. Il est aussi supposé que $\mathbf{e}_i^x, \mathbf{e}_i^y$ et \mathbf{e}_i^g ne sont pas corrélés. Cette hypothèse n'est pas obligatoire mais simplifie l'approche et l'application empirique. La matrice de variance-covariance des \mathbf{e}_i a donc la structure suivante :

$$(16) \quad \Psi = \begin{bmatrix} \Psi_{zz} & 0 \\ 0 & \Psi_{gg} \end{bmatrix} = \begin{bmatrix} \Psi_{yz} & 0 \\ \Psi_{xz} & 0 \\ 0 & \Psi_{gg} \end{bmatrix} = \begin{bmatrix} \Psi_{yy} & \Psi_{yx} & 0 \\ \Psi_{xy} & \Psi_{xx} & 0 \\ 0 & 0 & \Psi_{gg} \end{bmatrix}.$$

Ces hypothèses permettent de construire la structure des variances et des covariances conditionnelles des termes d'erreur du modèle économétrique :

$$(17a) \quad V[\omega_i^\pi \omega_i^{\pi'} / \mathbf{z}_i] = \Delta' \mathbf{M}_i \Psi_{zz} \mathbf{M}_i' \Delta + \Psi_{gg}$$

$$(17b) \quad V[\omega_i^y \omega_i^{\pi'} / \mathbf{z}_i] = -\Psi_{yz} \mathbf{M}_i' \Delta$$

$$(17c) \quad V[\omega_i^x \omega_i^{\pi'} / \mathbf{z}_i] = -\Psi_{xz} \mathbf{M}_i' \Delta$$

avec $\omega_i^\pi = -\mathbf{Q}_0 \omega_i^s$. Ces moments du second ordre seront utilisés pour construire des estimateurs d'une partie de la matrice de variance-covariance Ψ (détaillés dans la procédure d'estimation).

La forme quadratique de la fonction de coût a un avantage dans la détermination des fonctions de contrôle : elle permet de dériver des équations de parts de surfaces dont les termes d'erreur sont une combinaison linéaire des termes d'hétérogénéité des équations de rendement et de répartition des intrants. De plus, dans le cas où $\mathbf{B}_{0c}^x = 0$ (c'est-à-dire l'effet d'échelle dans l'équation de répartition des intrants est nul), l'équation de répartition des intrants est linéaire dans les \mathbf{s}_i . L'additivité des termes d'erreurs dans les

équations de surface, combinée à la linéarité des surfaces dans l'équation d'intrant, permettent d'utiliser directement les techniques de la projection linéaire, pour déterminer la forme de la corrélation entre les surfaces et les termes d'erreurs dans l'équation de répartition des intrants. Aucune hypothèse supplémentaire sur la distribution sur ces termes d'erreurs n'est donc nécessaire.

4.3.2 Fonctions de contrôle basées sur la projection linéaire

La projection linéaire est un outil mathématique utile dans les modèles linéaires. Son principal avantage est qu'elle peut être utilisée pour n'importe quelle variable aléatoire. La projection linéaire du terme e_i^x sur la variable s_i conditionnellement à z_i permet de décomposer le terme e_i^x en deux parties :

$$(18) \quad e_i^x = L[e_i^x / s_i ; z_i] + \zeta_i$$

avec le terme $[e_i^x / s_i ; z_i]$ la projection linéaire de e_i^x sur s_i . Le terme d'erreur ζ_i est orthogonal à cette projection par construction et on a : $E[\zeta_i] = 0$ et $E[s_i \zeta_i] = 0$. Par construction, cette projection linéaire s'écrit :

$$(19) \quad L[e_i^x / s_i ; z_i] = E[e_i^x / z_i] + \frac{\text{cov}[e_i^x, s_i / z_i]}{V[s_i / z_i]} [s_i - E[s_i / z_i]]$$

Il est montré que cette projection est la fonction affine de s_i qui prédit au mieux e_i^x conditionnellement à z_i , au sens de l'erreur quadratique moyenne.

La structure du modèle complet ainsi que celle des termes d'erreur permettent de déterminer explicitement les termes de cette projection linéaire :

$$(20a) \quad L[e_i^x / s_i ; z_i] = \Psi_{xz} C_i(\Psi) \omega_i^x,$$

avec :

$$(20b) \quad C_i = -\mathbf{M}_i' \Delta (\Delta' \mathbf{M}_i \Psi_{zz} \mathbf{M}_i' \Delta + \Psi_{gg})^{-1}.$$

Cette fonction de contrôle $L[\mathbf{e}_i^x / \mathbf{s}_i ; \mathbf{z}_i]$ peut alors être utilisée dans l'équation de répartition des intrants.

Une fonction de contrôle peut également être intégrée dans chaque équation de rendement, et ce pour trois raisons : pour contrôler l'endogénéité de \mathbf{s}_i (cette endogénéité peut être contrôlée par la méthode des Variables Instrumentales, contrairement au problème d'endogénéité des surfaces dans l'équation d'intrant) ; pour tenir compte des facteurs non-observés par l'économètre qui influencent à la fois les décisions de surfaces et les rendements, comme la qualité des sols, par exemple ; et enfin, pour estimer les variances des termes d'erreur des équations de rendements Ψ_{yy} . L'endogénéité des surfaces dans l'équation de répartition des intrants et dans les équations de rendement pourra être testée en posant l'hypothèse nulle $\Psi_{xx} = \Psi_{yy} = 0$. De la même façon que celles définies pour l'équation d'intrant, les fonctions de contrôle dans les équations de rendement sont définies par :

$$(21a) \quad L[\mathbf{e}_i^y / \mathbf{s}_i ; \mathbf{z}_i] = \Psi_{yz} C_i(\Psi) \omega_i^\pi,$$

avec :

$$(21b) \quad C_i = -\mathbf{M}_i' \Delta (\Delta' \mathbf{M}_i \Psi_{zz} \mathbf{M}_i' \Delta + \Psi_{gg})^{-1}.$$

Cette expression $L[\mathbf{e}_i^y / \mathbf{s}_i ; \mathbf{z}_i]$ peut être utilisée comme fonction de contrôle dans les équations de rendements.

4.3.3 Procédure d'estimation en trois étapes

Une procédure en trois étapes est proposée pour estimer le modèle complet. La première étape consiste à estimer les équations de rendements et de parts de surface avec la méthode des moments généralisés :

$$(22) \quad \begin{cases} \mathbf{y}_i = \mathbf{a}_0^y + 0.5 \mathbf{B}_0^y \mathbf{s}_i + \boldsymbol{\omega}_i^y \\ \mathbf{s}_i^- = -\mathbf{Q}_0^{-1} [\Delta' \mathbf{M}_i \mathbf{a}_0 - \mathbf{g}_0 - \mathbf{F}_{1i}] + \boldsymbol{\omega}_i^s \end{cases}$$

Ce système d'équations est un système d'équations simultanées à cause de l'endogénéité des surfaces dans les équations de rendements. Il faut donc construire des variables instrumentales valides pour tenir compte de l'endogénéité des \mathbf{s}_i et obtenir des estimateurs cohérents de tous les paramètres $\boldsymbol{\theta}$ (excepté du paramètre β_{01}).

La seconde étape consiste à estimer les équations des moments du second ordre à partir des résidus obtenus de la première étape $\boldsymbol{\omega}_i^y(\hat{\boldsymbol{\theta}})$ et $\boldsymbol{\omega}_i^s(\hat{\boldsymbol{\theta}})$ (et donc $\boldsymbol{\omega}_i^\pi(\hat{\boldsymbol{\theta}})$ car $\boldsymbol{\omega}_i^\pi = -\mathbf{Q}_0 \boldsymbol{\omega}_i^s$). Cette étape permet d'obtenir des estimateurs cohérents d'une partie des éléments de la matrice variance-covariance $\boldsymbol{\Psi}$. Elle est similaire à la seconde étape nécessaire pour construire les estimateurs des moindres carrés généralisés (MCG). Ce système d'équations est un système d'équations apparemment non reliés et est donc estimé par la méthode SUR :

$$(23) \quad \begin{cases} \boldsymbol{\omega}_i^\pi(\hat{\boldsymbol{\theta}}) \boldsymbol{\omega}_i^\pi(\hat{\boldsymbol{\theta}})' = \Delta' \mathbf{M}_i \boldsymbol{\Psi}_{zz} \mathbf{M}_i' \Delta + \boldsymbol{\Psi}_{gg} + \boldsymbol{\xi}_i^{ss} \\ \boldsymbol{\omega}_i^\pi(\hat{\boldsymbol{\theta}}) \boldsymbol{\omega}_i^y(\hat{\boldsymbol{\theta}})' = \boldsymbol{\Psi}_{yz} \mathbf{M}_i' \Delta + \boldsymbol{\xi}_i^{ys} \end{cases}$$

avec $E[\boldsymbol{\xi}_i^{ss} / \Delta \mathbf{M}_i'] = 0$ et $E[\boldsymbol{\xi}_i^{ys} / \Delta \mathbf{M}_i'] = 0$. Les estimateurs obtenus $\hat{\boldsymbol{\Psi}}$ permettent de construire une partie des fonctions de contrôle $\boldsymbol{\Psi}_{yz} \mathbf{C}_i(\hat{\boldsymbol{\Psi}}) \boldsymbol{\omega}_i^\pi(\hat{\boldsymbol{\theta}})$ et $\boldsymbol{\Psi}_{xz} \mathbf{C}_i(\hat{\boldsymbol{\Psi}}) \boldsymbol{\omega}_i^\pi(\hat{\boldsymbol{\theta}})$. Les éléments $\boldsymbol{\Psi}_{yz}$ et $\boldsymbol{\Psi}_{xz}$ de la matrice variance-covariance $\boldsymbol{\Psi}$ sont ré-estimés dans la troisième étape.

Dans la troisième étape, le modèle complet, avec les fonctions de contrôle, est estimé avec la méthode des moments généralisés. Tous les paramètres du modèle sont estimés, et les éléments $\boldsymbol{\Psi}_{yz}$ et $\boldsymbol{\Psi}_{xz}$ de la matrice de variance-covariance sont ré-estimés. Le modèle complet est défini par :

$$(24) \quad \begin{cases} \mathbf{y}_i = \mathbf{a}_0^y + 0.5 \mathbf{b}_0^y \mathbf{s}_i + \Psi_{yz} \mathbf{C}_i^y (\hat{\Psi}) \boldsymbol{\omega}_i^\pi (\hat{\boldsymbol{\theta}}) + \boldsymbol{\mu}_i^y \\ X_i = \mathbf{s}_i' [\mathbf{a}_0^x] + \mathbf{s}_i' \Psi_{xz} \mathbf{C}_i^x (\hat{\Psi}) \boldsymbol{\omega}_i^\pi (\hat{\boldsymbol{\theta}}) + \boldsymbol{\mu}_i^x \\ \mathbf{s}_i^- = -\mathbf{Q}_0^{-1} [\Delta' \mathbf{M}_i \mathbf{a}_0 - \mathbf{g}_0 - \mathbf{F}_{1i}] + \boldsymbol{\omega}_i^s \end{cases}$$

avec $E[\boldsymbol{\mu}_i^y / \mathbf{s}_i, \mathbf{z}_i] = 0$, $E[\boldsymbol{\mu}_i^x / \mathbf{s}_i, \mathbf{z}_i] = 0$ et $E[\boldsymbol{\omega}_i^y / \mathbf{z}_i] = 0$. Il est à noter que, quelquefois, l'estimation des équations non-linéaires dans les paramètres, comme les équations de parts de surfaces définies dans ce modèle, est difficile, voire impossible. Les conditions du premier ordre peuvent alors remplacer les équations de surfaces :

$$(25) \quad Q_0 s_i^- + [\Delta' M a_0 - g_0 - F_{ki}] + w_i^\pi = 0$$

Un inconvénient d'utiliser cette spécification est qu'il n'y aucune valeur prédite associée à ces équations et donc que les R^2 ne peuvent être calculés.

4.4 Application empirique

Le modèle économétrique multi-produit est estimé à partir de la procédure en trois étapes décrite dans la section précédente. Plusieurs spécifications du modèle sont également estimées afin d'étudier l'importance des deux motivations utilisées dans ce modèle pour décrire les choix diversifiés des assolements : les effets d'échelle et la fonction de coût.

4.4.1 Données

Les données utilisées sont des données issues du Réseau d'Information Comptable Agricole (RICA). Nous considérons des exploitations françaises spécialisées en céréales et grandes cultures sur la période 1989-2006. L'échantillon contient environ 6000 observations. Cette base de données fournit les rendements, les surfaces et les prix de chaque culture, la quantité totale d'intrants variables utilisés dans chaque exploitation, ainsi que les subventions perçues par les agriculteurs pour les céréales et les oléoprotéagineux. On distingue six cultures : le blé, les autres céréales (principalement du maïs et de l'orge), les

oléoprotéagineux (principalement du colza), des betteraves, des pommes de terre et le fourrage. Les betteraves, les pommes de terre et le fourrage sont supposés exogènes²¹. On considère un seul input variable regroupant les engrais, les produits phytosanitaires et les semences. Le prix de l'input agrégé est obtenu par les Statistiques Agricoles françaises. Toutes les quantités économiques sont exprimées en € de l'année 2000 par are.

Des variables sont introduites dans les équations de rendement et d'intrant pour tenir compte du progrès technique et de l'hétérogénéité des exploitations. Ce sont les mêmes que celles utilisées dans l'application du chapitre 3. En particulier, « l'indice de qualité » est inclus pour contrôler l'hétérogénéité des exploitations. Il est défini par :

$$(26) \quad q_{it} = (y_{li,t-1} - y_{li,t-1}^{Med}) / (y_{li,t-1}^{Max} - y_{li,t-1}^{Min}),$$

avec $y_{li,t-1}^{Med}$ la médiane, $y_{li,t-1}^{Max}$ le quantile à 99% et $y_{li,t-1}^{Min}$ le quantile à 1% du rendement du blé tendre pour l'agriculteur i à l'année $t-1$. Cet indice est basé sur le rendement du blé puisque les agriculteurs de cet échantillon sont spécialisés en céréales, et principalement dans la culture du blé. Il est défini année par année afin de contrôler les conditions spécifiques de chaque année. Les surfaces cultivées de betteraves et de pommes de terre l'année précédente sont également intégrées pour représenter les effets de ces cultures en tant que précédents culturels. Enfin, puisque les cultures appartenant à chaque agrégat n'ont pas les mêmes caractéristiques en termes de rendement et de besoin d'intrants, une variable représentant la part de céréales dans la quantité totale de céréales (*i.e.* avec le maïs) et une variable représentant la part de protéagineux dans la quantité totale d'oléoprotéagineux sont incluses dans le modèle.

4.4.2 Les principaux résultats

Le tableau 1 présente les estimateurs des paramètres des équations de rendement, de demande d'intrants et de parts de surfaces du modèle économétrique décrit dans cet

²¹ La culture des betteraves sucrières est soumise à un système de quota mis en place par l'Union Européenne, et les surfaces en pommes de terre sont majoritairement commercialisées sous contrats.

article. Le tableau 2 présente les élasticités prix correspondantes. Les résultats obtenus montrent que ce modèle décrit de façon satisfaisante le comportement des agriculteurs.

Le R^2 de l'équation de rendement du blé tendre est de .42, ce qui signifie que le modèle explique environ 42% de la variation du rendement du blé à partir des prix et des variables explicatives retenues. Les R^2 des équations de rendements sont de .35 pour les autres céréales et .32 pour les oléoprotéagineux et celui de l'équation d'intrant est de .36. Ces coefficients de détermination sont assez satisfaisants. Il est toujours difficile d'obtenir de bons R^2 à partir d'estimations sur les prix et avec des données à l'échelle microéconomique. Cependant, près de 95% des paramètres estimés sont statistiquement différents de zéro dans un intervalle de confiance de 5%.

Un test de Wald est réalisé pour tester l'existence de l'endogénéité des surfaces dans les équations de rendement et dans l'équation d'allocation des intrants. L'hypothèse nulle $\Psi_{yz} = \Psi_{xz} = 0$ est rejetée, confirmant l'endogénéité des surfaces et donc l'utilisation des fonctions de contrôle dans le modèle.

Toutes les conditions de concavité des fonctions de rendement et de profit sont respectées. Les fonctions de rendement sont concaves puisque les paramètres γ_k sont positifs. La fonction de profit total est concave puisque la forme quadratique associée $s_i^- 'Q_0 s_i^-$ est semi-définie négative : les paramètres q_1 et q_2 de la matrice Q_0 sont négatifs et le déterminant de Q_0 est strictement positif.

Les paramètres représentant les effets d'échelle α_{1k} (dans les fonctions de rendement) étant négatifs, l'hypothèse de rendement marginal de la terre décroissant est également vérifiée. Ces paramètres sont de -5.66 pour les oléoprotéagineux, -1.45 pour le blé et -0.93 pour les céréales. Ces facteurs d'échelle montrent que le rendement d'une culture décroît avec la quantité de surface allouée à cette culture. Ceci est particulièrement vrai pour le colza qui a un délai de retour conseillé sur une parcelle de trois ans.

Les autres paramètres estimés semblent cohérents avec les principes agronomiques. Le rendement maximum moyen du blé est de 9.72 € par are, celui des céréales est de 8.74 €

par are et celui des oléoprotéagineux de 7.75 € par are. Les quantités d'intrants moyennes nécessaires pour obtenir le rendement maximum du blé sont de 6.63 € par are, de 5.97 € par are pour les céréales et de 6.85 € par are pour les oléoprotéagineux.

Les variables utilisées pour contrôler l'hétérogénéité des exploitations ont les effets attendus sur le rendement et les utilisations d'intrants. L'«indice de qualité» a un effet positif sur le rendement et sur la demande d'intrant. Ce résultat décrit un processus d'intensification. Les agriculteurs qui obtiennent les meilleurs rendements grâce à leurs conditions de production spécifiques (qualité du sol...), sont également ceux qui utilisent les plus grandes quantités d'intrants chimiques. Les surfaces cultivées de betteraves et de pommes de terre de l'année précédente ont un effet positif sur le rendement de toutes les cultures.

La part de céréales (dans la surface totale allouée aux céréales et au maïs) a un effet négatif sur le rendement (-1.18) et sur les demandes d'intrants (-2.83). Ces résultats sont ceux attendus, puisque le maïs a un rendement plus élevé que celui des céréales et nécessite plus d'intrants. De même, la part des protéagineux (dans la surface totale allouée aux oléoprotéagineux) a un effet positif sur le rendement (1.20) et a un effet important et négatif sur la quantité d'intrants (-4.08). Ces résultats sont conformes également avec l'agronomie puisque les pois protéagineux ont un rendement légèrement supérieur aux oléagineux, et nécessite très peu d'intrants chimiques, et en particulier aucun apport d'azote.

Les paramètres g_{0c} représentent les coûts fixes liés à la gestion de la culture c par rapport à ceux de la culture de référence, qui est le blé. Ces résultats montrent que les oléoprotéagineux sont les cultures avec les coûts fixes les plus élevés. Les agriculteurs étant spécialisés en céréales, et principalement en blé, les coûts fixes associés à la culture des céréales et surtout du blé sont moins importants. Les éléments de la matrice Q_0 dépendent des paramètres g_{km} . Ces paramètres n'ont pas été différenciés pour simplifier l'estimation, et donc ces résultats ne sont pas directement interprétables.

Les élasticités-prix sont calculées à la moyenne de l'échantillon (tableau 2). Les rendements des cultures sont inélastiques aux prix : une augmentation de 10% du prix du blé entraîne une augmentation de 3.7 % de son rendement. Guyomard et al. (1996) ont estimé également des élasticités-prix sur des données françaises (mais sur des données agrégées), et leurs résultats montrent que les élasticités prix propres des rendements varient entre 0.2 et 0.4 pour les céréales et les oléoprotéagineux. La variation du prix d'une culture influence plus la quantité d'intrant utilisée pour cette culture que son rendement : une augmentation de 10% du prix du blé entraîne une augmentation de 3.7 % de son rendement et de 8.4% des quantités d'intrants appliquées. Toutes les élasticités-prix propres de demande de surface sont positives, et sont supérieures à 1 pour le blé et les céréales : une augmentation de 10% du prix du blé entraîne une augmentation de 11% de la surface allouée au blé et une diminution de 10% de la surface allouée aux autres céréales. Guyomard et al. (1996) ont également estimé que la surface de maïs est élastique à son propre prix (1.27). Cela montre également que le blé et les autres céréales sont considérés comme des substituts.

4.4.3 Rendements marginaux de la terre décroissants versus coûts implicites des assolements

Heckelei et Wolff (2003) ont proposé deux cas particuliers de la fonction de profit espéré décrite dans ce modèle par l'équation (10). Le premier cas correspond à notre modèle sans fonction de coût ($g_c = g_{cm} = 0$ pour $c, m=1, \dots, C$). Dans ce cas, la diversification de l'assolement n'est motivée que par les rendements marginaux décroissants de la terre. Le second cas correspond à notre modèle sans les effets d'échelle ($\alpha_{1c} = \beta_{1c} = 0$ pour $c=1, \dots, C$). Dans ce cas, la diversification de l'assolement n'est due qu'aux effets décrits par la fonction de coût.

Dans cette application, nous estimons ces deux cas particuliers. Le modèle 1 correspond au modèle avec seulement les effets d'échelle comme motif de diversification des cultures, le modèle 2 correspond au modèle avec seulement la fonction de coût et le modèle 3 est le

modèle général avec ces deux effets, décrit dans la section précédente. Les tableaux 3, 4 et 5 présentent les estimateurs des paramètres des équations de rendement, de l'équation d'intrants et des équations de parts de surface selon ces trois modèles. Le tableau 6 présente leurs élasticités-prix. L'endogénéité des surfaces est contrôlée dans tous ces modèles à partir de l'approche par les fonctions de contrôle.

Les résultats nous montrent qu'il existe des différences significatives entre les modèles dans les paramètres estimés, et particulièrement pour les oléoprotéagineux. Le rendement maximum moyen des oléoprotéagineux est estimé à 7 € par are dans le modèle 1 contre 8.25 € par are dans le modèle 2 (tableau 3). Les quantités d'intrants moyennes nécessaires pour obtenir ce rendement maximum sont de 6.87 € par are dans le modèle 1 contre 4.76 € dans le modèle 2 (tableau 4). La marge brute de chaque agrégat est calculée. A partir du modèle 2, la marge brute moyenne des oléoprotéagineux est de 5.79 € par are et est plus élevée que celle du blé (4.91 € par are) et des céréales (4.69 € par are), ce qui n'est pas cohérent d'un point de vue agronomique. A partir du modèle 1, la marge brute du blé est de 6.64 € par are devant les céréales (5.19 € par are) et devant les oléoprotéagineux (4.31 € par are). Le modèle 2 explique l'attrait des agriculteurs pour les oléoprotéagineux seulement à travers la marge brute, ce qui explique qu'elle soit surestimée. La fonction de coût implicite utilisée dans les modèles 1 et 3 permet de tenir compte de plus d'informations sur les motivations des agriculteurs dans leur choix d'assolement.

Nous constatons également que la prise en compte de la fonction de coût implicite dans le modèle modifie les effets d'échelle et les effets des précédents culturaux. Les effets d'échelle (des équations de rendements) sont compris entre -2.21 (pour les céréales) à -9.08 (pour les oléoprotéagineux) dans le modèle 2, alors qu'ils sont compris entre -0.93 (pour les céréales) et -5.66 (pour les oléoprotéagineux) dans le modèle 3. La surface de betteraves et de pommes de terre cultivée l'année précédente a un effet significatif et positif sur le rendement de toutes les cultures dans les modèles 1 et 3, et a un effet significatif (positif ou négatif) sur les quantités d'intrants nécessaires pour toutes les

cultures. Par contre, dans le modèle 2, elle a un effet négatif sur le rendement des oléoprotéagineux et n'a pas d'effet significatif sur les quantités d'intrants nécessaires.

En termes d'élasticités (tableau 6), on observe que les modèles 1 et 3 qui intègrent une fonction de coût, ont des élasticités-prix très proches. La prise en compte des rendements d'échelle n'a donc pas une forte influence sur les élasticités lorsqu'une fonction de coût implicite est définie. Par contre, les élasticités sont différentes entre ces modèles et le modèle 2, et particulièrement pour les surfaces. Une augmentation du prix du blé de 10 % entraîne une augmentation de 20 % de la surface allouée au blé avec le modèle 2, contre une augmentation de 11% avec les modèles 1 et 3.

Les économistes qui estiment des modèles de choix de production intègrent généralement un paramètre de « calage », ce qui permet de rajouter de la flexibilité au modèle empirique, et de mieux approximer le comportement des agriculteurs. Ce paramètre, bien qu'il n'ait pas d'interprétation directe, joue le même rôle que le paramètre de coûts fixes dans notre fonction de coût. Cette pratique (ainsi que les résultats obtenus) nous conforte dans notre choix de construire une fonction de coût implicite de façon complète.

Conclusion

Le modèle de choix de production construit dans ce travail s'appuie sur les modèles développés dans les chapitres précédents. L'originalité de ce travail est donner une forme quadratique à la fonction de coût implicite de gestion des assolements. Cette forme quadratique permet au modèle de choix de production d'être plus général, en relâchant quelques hypothèses posées dans les chapitres précédents : l'hypothèse de rendement marginal de la terre constant, et l'hypothèse de normalité des termes d'erreur du modèle. Elle permet également au modèle d'être plus flexible et de déterminer les choix d'assolements à partir de plus de paramètres.

Dans l'application empirique, nous comparons deux cas particuliers du modèle initial : l'un avec seulement les effets d'échelle et l'autre avec seulement la fonction de coût implicite.

La fonction de coût permet de tenir compte de façon implicite, des contraintes de gestion du travail et du capital, et de contraintes agronomiques. Les effets d'échelle permettent de tenir compte principalement des méfaits de la monoculture par rapport aux rotations culturales. Nous montrons que l'utilisation de la fonction de coût implicite a une influence marquante sur les paramètres estimés et sur les élasticités-prix, plus que les effets d'échelle.

Plusieurs extensions du modèle sont possibles. En adaptant l'approche de Röhm et Dabbert (2003), il est possible de spécialiser la forme de la fonction de coût quadratique pour tenir compte des similarités des rôles des différentes cultures pour ce qui concerne les successions culturales et l'organisation du travail. Une logique similaire a conduit à la construction des groupes de cultures utilisés dans les modèles de choix d'assolement MNL emboîtés (chapitre 2). La deuxième extension possible du modèle concerne le traitement des solutions en coin. Lorsqu'un agriculteur ne cultive pas certaines cultures, les surfaces en terre allouées à ces cultures prennent des valeurs nulles. Le modèle de choix d'assolement fondé sur la fonction de coût implicite quadratique, décrit dans cet article, semble adapté pour traiter ce type de problème, grâce notamment à l'utilisation de techniques d'intégration des fonctions de vraisemblance par des méthodes de simulation (Kao, Lee et Pitt, 2001).

References

- Carpentier, A., et R. D. Weaver. 1996. "Intertemporal and Interfirm Heterogeneity: Implications for Pesticide Productivity." *Canadian Journal of Agricultural Economics* 44:219-236.
- Chamberlain, G. 1982. "Multivariate Regression models for panel data." *Journal of Econometrics* 18:5-46.
- Chavas, J.-P. et T. Holt. 1990. "Acreage Decisions under Risk: The Case of Corn and Soybeans." *American Journal of Agricultural Economics* 72:529-538.
- Coyle, B. T. 1992. "Risk Aversion and Price Risk in Duality Models of Production: A Linear Mean-Variance Approach." *American Journal of Agricultural Economics* 74:849-859.
- Coyle, B. T. 1999. "Risk Aversion and Price Risk in Duality Models of Production: A Mean-Variance Approach." *American Journal of Agricultural Economics* 81:553-567.
- Guyomard, H., Baudry, M. et A. Carpentier. 1996. "Estimating Crop Supply Response in the Presence of Farm Programmes: Application to the CAP." *European Review of Agricultural Economics* 23:401-420.
- Heckelei, T. et H. Wolff. 2003. "Estimation of Constrained Optimisation Models for Agricultural Supply Analysis Based on Generalised Maximum Entropy." *European Review of Agricultural Economics* 30:27-50.
- Heckman, J. J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47, 153-161.
- Heckman, J. J. 1974. "Shadow Prices, Market Wages, and Labor Supply." *Econometrica* 42, 679-693.
- Howitt, E. 1995. "Positive Mathematical Programming." *American Journal of Agricultural Economics* 77:329-342.
- Imbens, G. et J. M. Wooldridge. 2007. "Control Functions and Related Methods" Lecture notes NBER Summer Institute '07, What's new in econometrics?, NBER.

- Just, R.E., et R.D. Pope. 2001. "The Agricultural Producer: Theory and Statistical Measurement." In B.L. Gardner and G.C. Rausser, ed. *Handbook of Agricultural Economics*, Vol. 1A. Amsterdam, Elsevier.
- Kao, C., L.-F. Lee and M.M. Pitt, 2001. "Simulated Maximum Likelihood Estimation of the Linear Expenditure System with Binding Non-negativity Constraints." *Annals of Economics and Finance* 2:215-55.
- McElroy, M. B. 1987. "Additive General Error Models for Production, Cost, and Derived Demand or Share systems." *Journal of Political Economy* 95:737-757.
- Moore, M. R. et D. H. Negri. 1992. "A Multicrop Production Model of Irrigated Agriculture, Applied to Water Allocation Policy of the Bureau of Reclamation." *Journal of Agricultural and Resource Economics* 17:29-43.
- Moro, D. et P. Sckokai. 1999. "Modelling the CAP Arable Crop Regime in Italy: Degree of Decoupling and Impact of Agenda 2000." *Cahiers d'économie et sociologie rurales* 53:50-73.
- Ozarem, P.F. et J.A. Miranowski. 1994. "A Dynamic Model of Acreage Allocation with General and Crop-Specific Soil Capital." *American Journal of Agricultural Economics* 76:385-395.
- Petrin, A. et K. E. Train. 2008. "A Control Function Approach to Endogeneity in Consumer Choice Model." *Forthcoming, Journal of Marketing Research*.
- Pope, R. D. et R. E. Just. 2003. "Distinguishing Errors in Measurement from Errors in Optimization." *American Journal of Agricultural Economics* 85:348-358.
- Röhm, O., et S. Dabbert. 2003. Integrating Agri-Environmental Programs into Regional Production Models: An Extension of Positive Mathematical Programming, *American Journal of Agricultural Economics* 85:254-265.
- Wooldridge, J.M. 2004. "Estimating Average Partial Effects under Conditional Moment Independence Assumptions." *The Institute for Fiscal Studies, Dpt of Economics, UCL. Cemmap*, WP CWP03/04.
- Wooldridge, J.M. 2002. *The Econometrics of Cross-Sections and Panel Data*. The MIT Press: Cambridge MA.

**Tableau 1. Estimateurs des paramètres des équations de rendements,
de demande d'intrants et de parts de surface.**

Variables explicatives	Blé tendre		Autres céréales		Oléo- protéagineux	
Offre de production						
Effets des prix (γ)	3.48	(0.32)	1.99	(0.18)	2.60	(1.22)
Constante	3.69	(0.20)	2.11	(0.29)	2.47	(0.15)
Indice de « qualité »	-0.07	(0.22)	-0.15	(0.21)	4.26	(0.26)
Rendement moyen potentiel	9.72	(1.06)	8.74	(1.00)	7.75	(1.12)
Constante (α_0)	8.50	(0.07)	8.24	(0.10)	6.58	(0.07)
Tendance	0.14	(0.01)	0.12	(0.01)	0.18	(0.01)
Indice de « qualité »	2.33	(0.09)	1.99	(0.11)	3.19	(0.12)
Effet précédent	1.30	(0.14)	1.87	(0.19)	0.75	(0.22)
Composition des agrégats	-		-1.18	(0.23)	1.20	(0.13)
Irrigation	0.41	(0.10)	1.64	(0.09)	-0.08	(0.06)
Effets d'échelle (α_1)	-1.45	(0.23)	-0.93	(0.28)	-5.66	(0.36)
<i>R-carré</i>	<i>0.42</i>		<i>0.35</i>		<i>0.32</i>	
Demande d'intrants						
Besoin moyen d'intrants (β)	6.63	(0.56)	5.97	(0.62)	6.85	(1.88)
Constante	6.54	(0.14)	6.25	(0.21)	5.85	(0.17)
Indice de « qualité »	0.94	(0.21)	0.44	(0.22)	3.79	(0.26)
Effet précédent	-2.31	(0.75)	1.29	(0.79)	2.66	(0.94)
Composition des agrégats	-		-2.83	(0.30)	-4.08	(0.30)
Irrigation	-0.49	(0.18)	0.84	(0.11)	0.57	(0.15)
Betteraves			9.19	(1.43)		
Pommes de terre			12.64	(2.07)		
Fourrage			2.86	(0.80)		
<i>R-carré</i>			<i>0.36</i>			
Parts de surface						
Coûts fixes (g_0)	-		2.97	(0.24)	3.33	(0.52)
Matrice Q_0	-		-7.99	(0.91)	-6.49	(0.57)
	-		-6.49	(0.57)	-31.90	(3.39)

Note: Les écarts-types sont entre parenthèses.

Tableau 2. Elasticités-prix moyennes.

	Blé tendre	Autres céréales	Oléo- protéagineux	Intrants
Offre de productions				
Blé tendre	0.366 (0.18)	-	-	-0.366 (0.18)
Autres céréales	-	0.241 (0.12)	-	-0.241 (0.12)
Oléoprotéagineux	-	-	0.352 (0.23)	-0.352 (0.23)
Demande d'intrants				
Blé tendre	0.894 (0.32)	-	-	-0.894 (0.32)
Autres céréales	-	0.460 (0.15)	-	-0.460 (0.15)
Oléoprotéagineux	-	-	0.536 (0.35)	-0.536 (0.35)
Parts de surface				
Blé tendre	1.062	-0.951	-0.047	0.077
Autres céréales	-1.004 (0.15)	1.195 (0.19)	-0.207 (0.03)	-0.057 (0.09)
Oléoprotéagineux	-0.058 (0.01)	-0.244 (0.04)	0.254 (0.03)	-0.020 (0.05)

Note: Les écarts-types sont entre parenthèses.

Tableau 3. Estimateurs des paramètres des équations de rendements selon les 3 modèles.

Variables explicatives	Modèle 1 Avec la fonction de coût			Modèle 2 Avec les effets d'échelle			Modèle 3 Avec les deux effets		
	Blé tendre	Autres céréales	Oléo- protéa.	Blé tendre	Autres céréales	Oléo- protéa.	Blé tendre	Autres céréales	Oléo- protéa.
Offre de production									
Effets des prix (γ)	3.50***	2.05***	2.52***	3.21***	2.59***	2.88***	3.48***	1.99***	2.60***
Constante	3.66***	2.15***	2.45***	3.39***	2.70***	2.82***	3.69***	2.11***	2.47***
Indice de « qualité »	-0.15	-0.12	4.17***	0.35	-0.17	3.66***	-0.07	-0.15	4.26***
Rendement moyen potentiel	9.41***	8.64***	7.00***	9.73***	9.14***	8.25***	9.72***	8.74***	7.75***
Constante (α_0)	8.17***	8.14***	5.69***	8.55***	8.43***	7.16***	8.50***	8.24***	6.58***
Tendance	0.14***	0.12***	0.19***	0.14***	0.12***	0.18***	0.14***	0.12***	0.18***
Indice de « qualité »	2.27***	2.09***	3.16***	2.51***	1.97***	2.88***	2.33***	1.99***	3.19***
Effet précédent	1.26***	2.02***	2.20***	1.65***	1.81***	-0.46**	1.30***	1.87***	0.75***
Composition des agrégats	-	-1.44***	0.69***	-	-1.10***	1.31***	-	-1.18***	1.20***
Irrigation	0.50***	1.50***	0.07	0.39***	1.85***	-0.19***	0.41***	1.64***	-0.08
Effets d'échelle (α_1)	-	-	-	-2.57***	-2.21***	-9.08***	-1.45***	-0.93***	-5.66***
<i>R-carré</i>	<i>0.42</i>	<i>0.33</i>	<i>0.32</i>	<i>0.43</i>	<i>0.35</i>	<i>0.31</i>	<i>0.42</i>	<i>0.35</i>	<i>0.32</i>

Note : (*), (**), (***) signifient que les estimateurs des paramètres sont statistiquement différents de zéro à un seuil de confiance de, respectivement, 10%, 5% et <1%.

Tableau 4. Estimateurs des paramètres de l'équation de demande d'intrants selon les 3 modèles.

Variables explicatives	Modèle 1 Avec la fonction de coût			Modèle 2 Avec les effets d'échelle			Modèle 3 Avec les deux effets		
	Blé tendre	Autres céréales	Oléo- protéa.	Blé tendre	Autres céréales	Oléo- protéa.	Blé tendre	Autres céréales	Oléo- protéa.
Demande d'intrants									
Besoin moyen d'intrants (β)	6.56 ***	6.03 ***	6.87 ***	7.59 ***	6.64 ***	4.76 ***	6.63 ***	5.97 ***	6.85 ***
Constante	6.37 ***	6.27 ***	6.09 ***	7.36 ***	6.86 ***	3.99 ***	6.54 ***	6.25 ***	5.85 ***
Indice de « qualité »	0.84 ***	0.45 **	3.73 ***	1.45 ***	0.39 *	2.96 ***	0.94 ***	0.44 **	3.79 ***
Effet précédent	-3.12 ***	0.64	3.85 ***	-0.55	-1.06	-0.66	-2.31*	1.29 *	2.66 ***
Composition des agrégats	-	-2.80 ***	-4.53 ***	-	-2.49 ***	-1.98 ***	-	-2.83 ***	-4.08 ***
Irrigation	-0.40 ***	0.81 ***	-0.14	-0.43 ***	1.41 ***	0.01	-0.46 ***	0.84 ***	0.57 **
Betteraves		9.45 ***			8.33 ***			9.19 ***	
Pommes de terre		11.93 ***			11.21 ***			12.64 ***	
Fourrage		3.20 ***			3.82 ***			2.86 ***	
<i>R-carré</i>		<i>0.35</i>			<i>0.36</i>			<i>0.36</i>	

Note : (*), (**), (***) signifient que les estimateurs des paramètres sont statistiquement différents de zéro à un seuil de confiance de ; respectivement, 10%, 5% et <1%.

Tableau 5. Estimateurs des paramètres des équations de parts de surface selon les 3 modèles.

Variables explicatives	Modèle 1 Avec la fonction de coût			Modèle 2 Avec les effets d'échelle			Modèle 3 Avec les deux effets		
	Blé tendre	Autres céréales	Oléo- protéa.	Blé tendre	Autres céréales	Oléo- protéa.	Blé tendre	Autres céréales	Oléo- protéa.
Parts de surface									
Coûts fixes (g_0)	-	2.82 ***	4.17 ***	-	-	-	-	2.97 ***	3.33 ***
Matrice Q_0	-	-7.09 ***	-6.16 ***	-	-2.08 ***	-1.96 ***	-	-7.99 ***	-6.49 ***
	-	-6.16 ***	-32.16 ***	-	-1.96 ***	-8.93 ***	-	-6.49 ***	-31.90 ***

Note : (*), (**), (***) signifient que les estimateurs des paramètres sont statistiquement différents de zéro à un seuil de confiance de, respectivement, 10%, 5% et <1%.

Tableau 6. Elasticités-prix moyennes calculées à partir des 3 modèles.

	Modèle 1			Modèle 2			Modèle 3		
	Avec la fonction de coût			Avec les effets d'échelle			Avec les deux effets		
	Price			Price			Price		
	Blé tendre	Autres céréales	Oléo-protéa.	Blé tendre	Autres céréales	Oléo-protéa.	Blé tendre	Autres céréales	Oléo-protéa.
Offre de production									
Blé tendre	0.356	-	-	0.336	-	-	0.366	-	-
Autres céréales	-	0.257	-	-	0.314	-	-	0.241	-
Oléoprotéagineux	-	-	0.347	-	-	0.393	-	-	0.352
Demande d'intrants									
Blé tendre	0.873	-	-	0.608	-	-	0.894	-	-
Autres céréales	-	0.495	-	-	0.584	-	-	0.460	-
Oléoprotéagineux	-	-	0.524	-	-	0.397	-	-	0.536
Parts de surface									
Blé tendre	1.079	-0.988	-0.044	2.028	1.596	-0.29	1.062	-0.951	-0.047
Autres céréales	-1.027	1.247	-0.217	-1.681	1.963	-0.324	-1.004	1.195	-0.207
Oléoprotéagineux	-0.052	-0.259	0.261	-0.347	-0.367	0.614	-0.058	-0.244	0.254

Annexe A – Les notations matricielles.

$$\mathbf{a}_0^y = \begin{bmatrix} \alpha_{01} - 0.5\gamma_1(w/p_1)^2 \\ \vdots \\ \alpha_{0K} - 0.5\gamma_K(w/p_K)^2 \end{bmatrix}, \quad \mathbf{a}_0^x = \begin{bmatrix} \beta_{01} - \gamma_1(w/p_1) \\ \vdots \\ \beta_{0K} - \gamma_K(w/p_K) \end{bmatrix}.$$

$$\mathbf{B}_0^y = \begin{bmatrix} \alpha_{11} & 0 \\ & \ddots \\ 0 & \alpha_{1K} \end{bmatrix}, \quad \mathbf{B}_0^x = \begin{bmatrix} \beta_{11} & 0 \\ & \ddots \\ 0 & \beta_{1K} \end{bmatrix}.$$

$$\mathbf{M} = \begin{bmatrix} p_1 & 0 & -w & 0 \\ & \ddots & & \ddots \\ 0 & p_K & 0 & -w \end{bmatrix} \quad \text{et} \quad \Delta' \mathbf{M} = \begin{bmatrix} -p_1 & p_2 & 0 & w & -w & 0 \\ \vdots & & \ddots & \vdots & & \ddots \\ -p_1 & 0 & p_K & w & 0 & -w \end{bmatrix}.$$

$$\mathbf{g}_0 = \begin{bmatrix} g_{02} \\ \vdots \\ g_{0K} \end{bmatrix}, \quad \mathbf{G}_0 = \begin{bmatrix} g_{11} & \cdots & g_{1K} \\ \vdots & \ddots & \vdots \\ g_{1K} & \cdots & g_{KK} \end{bmatrix}.$$

$$\mathbf{A}(\mathbf{b}_0) = \begin{bmatrix} (p_1\alpha_{11} - w\beta_{11}) + (p_2\alpha_{12} - w\beta_{12}) & \cdots & (p_1\alpha_{11} - w\beta_{11}) \\ \vdots & \ddots & \vdots \\ (p_1\alpha_{11} - w\beta_{11}) & \cdots & (p_K\alpha_{1K} - w\beta_{1K}) + (p_K\alpha_{1K} - w\beta_{1K}) \end{bmatrix}$$

$$\Delta' \mathbf{G}_0 \Delta = \begin{bmatrix} g_{22} + g_{11} - 2g_{12} & \cdots & g_{2K} + g_{11} - g_{12} - g_{1K} \\ \vdots & \ddots & \vdots \\ g_{2K} + g_{11} - g_{12} - g_{1K} & \cdots & g_{KK} + g_{11} - 2g_{1K} \end{bmatrix} \quad \text{et} \quad \mathbf{Q}_0 = \begin{bmatrix} q_1 & q_3 \\ q_3 & q_2 \end{bmatrix}$$

$$f_1 = s_e \pi_1^e + s_e^2 [0.5(p_1\alpha_{11} - w\beta_{11}) - 0.5g_{11}] - g$$

$F_{1c} = s_e [(g_{1c} - g_{11}) + (p_1\alpha_{11} - w\beta_{11})]$ avec s_e la part de surface allouée aux cultures endogènes.

$$\Psi_{yy} = \begin{bmatrix} \omega_{11}^y & \cdots & \omega_{1K}^y \\ \vdots & \ddots & \vdots \\ \omega_{1K}^y & \cdots & \omega_{KK}^y \end{bmatrix}, \quad \Psi_{yx} = \begin{bmatrix} \omega_{11}^{yx} & \cdots & \omega_{1K}^{yx} \\ \vdots & \ddots & \vdots \\ \omega_{K1}^{yx} & \cdots & \omega_{KK}^{yx} \end{bmatrix} \quad \text{et} \quad \Psi_{gg} = \begin{bmatrix} \omega_{11}^g & \cdots & \omega_{1K}^g \\ \vdots & \ddots & \vdots \\ \omega_{1K}^g & \cdots & \omega_{KK}^g \end{bmatrix}.$$

Conclusion

Les modèles micro-économétriques multi-produit développés dans cette thèse contribuent à la littérature existante sur les modèles de choix de production des exploitants agricoles. Ils permettent de décrire le comportement de court terme des agriculteurs dans un cadre statique et plus précisément, leurs choix d'assolements et leurs décisions d'utilisations d'intrants variables par culture. Les apports de ces modèles sont résumés en quatre points :

1) Ils permettent une description originale du comportement des agriculteurs, prenant en compte la séquence dans leurs décisions. L'agriculteur est supposé pouvoir adapter les quantités d'intrants variables qu'il applique à chaque culture en fonction des événements qui surviennent pendant la campagne agricole.

2) Ils permettent une meilleure communication des résultats, grâce aux spécifications particulières des équations de rendements et de parts de surfaces. Les paramètres du modèle ont une interprétation directe « agronomique » ce qui facilite les collaborations pluridisciplinaires.

3) Ils permettent d'obtenir des résultats cohérents avec la théorie économique (les contraintes de concavité des fonctions de production et de profit sont vérifiées), valides sur le plan économétrique (les différentes méthodes utilisées donnent des estimateurs non biaisés), et en accord avec les effets attendus (et notamment, d'un point de vue agronomique).

- 4) Ils permettent d'envisager des améliorations indispensables pour expliquer les choix des agriculteurs, et notamment la prise en compte des solutions en coin et l'intégration de la dynamique. Ces travaux d'améliorations sont en cours ou en phase de réflexion. Ils sont détaillés ci-dessous.

La première voie d'amélioration des modèles présentés dans la thèse concerne le traitement des solutions en coin. Lorsqu'un agriculteur ne cultive pas certains types de cultures, les surfaces en terre allouées à ces cultures prennent des valeurs nulles. Le modèle de choix d'assolement fondé sur des fonctions de coût quadratiques peut être utilisé pour l'analyse des solutions en coin. Les développements récents des techniques d'inférence statistique assistée par des méthodes de simulation offre des perspectives intéressantes dans ce contexte.

La deuxième voie d'amélioration du modèle développé dans cette thèse est la prise en compte de la dynamique de production. Cette dernière a été rarement introduite dans la modélisation des choix de production (sans doute parce que les solutions analytiques sont difficiles à obtenir dans les modèles d'optimisation dynamique) ou alors de manière assez *ad hoc* (Eckstein, 1984). Quelques travaux néanmoins ont utilisé les outils de l'optimisation dynamique pour traiter de problèmes spécifiques, essentiellement de la gestion des reliquats d'engrais (Ozarem et Miranowski, 1994 ; Thomas, 2003 ; Arnberg et Hansen, 2007). Or les effets des rotations culturales concernent également la gestion du risque phytosanitaire et la gestion de la structure des sols. Considérer tous ces effets selon les principes d'optimisation dynamique s'avère très complexe. Une possibilité envisagée pour prendre en compte des effets généraux de rotations est de caractériser les choix d'assolement optimaux par des choix discrets. La littérature sur les modèles de choix discrets dynamiques se développe et offre d'importantes perspectives de modélisation (Rust, 1987 ; Train, 2003, Livingston et al., 2008). Cette extension est envisagée à partir du modèle de choix discret présenté dans le chapitre 2.

La troisième voie d'amélioration de notre modèle est d'y intégrer des variables climatiques et pédologiques pour mieux capter les effets dus à l'hétérogénéité des exploitations. Quelques premières estimations (à partir des mêmes données que dans les applications empiriques) ont montré que des variables climatiques, telles que la température maximale, minimale, le rayonnement et la pluviométrie, peuvent expliquer à elles-seules 50% de la variabilité du rendement du blé tendre observée en France.

Enfin, une étape de validation du modèle par les résultats est nécessaire. Cette thèse est principalement méthodologique et les applications réalisées ont quelquefois plus un rôle d'illustrations que de test de validation du modèle sur un plan économique. Les résultats en matière agronomique semblent conformes aux attentes. Il serait cependant utile de les confronter à d'autres bases de données. Par exemple, nous disposons de bases de données (enquêtes sur les pratiques culturales, base de données sur les céréaliers de la Meuse) dans lesquelles les allocations des intrants entre les cultures sont connues. Un moyen de valider notre approche d'allocation des intrants (chapitre 3) serait de confronter les allocations estimées aux allocations observées.

Il nous semble important de noter également que les résultats obtenus dans ce travail de thèse sont à interpréter avec précaution. D'une part, nous avons regroupé les intrants variables dans un seul agrégat, et ce, pour simplifier les estimations. Maintenant que les modèles « fonctionnent », il serait intéressant de séparer les principaux intrants variables, que sont les engrais, les produits phytosanitaires et les semences. D'autre part, nous avons agrégé également les cultures afin d'éviter le problème des solutions en coin. Regrouper les oléagineux et les protéagineux n'a pas grand sens d'un point de vue agronomique, de même pour les céréales et le maïs. Ces cultures n'ont pas les mêmes caractéristiques, (en termes de rendement ou de besoin d'intrants chimiques), ni les mêmes rôles (en terme de précédent cultural). Une fois nos modèles adaptés pour traiter les solutions en coin, nous pourrions alors regrouper les cultures de façon plus cohérente d'un point de vue agronomique.

Au-delà des améliorations possibles du modèle développé dans cette thèse, le projet initial de collaboration avec des agronomes, qui a dicté nos choix de modélisation, concerne l'analyse de la durabilité des variétés rustiques de blé tendre associées à des itinéraires techniques à bas niveaux d'intrants. Des travaux sur l'analyse et la prévision de l'interaction entre variétés et itinéraires techniques ont permis de rendre compte des possibilités de désintensification en adaptant l'ensemble d'un itinéraire technique au choix variétal (Saulas et Meynard, 1998 ; Loyce et al., 2001 a et b ; Rolland et al., 2003, 2006 ; Félix et al., 2002, 2003). Des itinéraires techniques (ou encore ITK) à bas niveaux d'intrants permettent de réduire les excédents de fertilisation azotée et l'usage de produits phytosanitaires. Afin d'éviter le risque de verse et de maladies lié à la diminution de l'utilisation des produits phytosanitaires, ces itinéraires sont associés à des variétés rustiques, plus résistantes aux maladies. Ces pratiques, économes en intrants, tendent à diminuer les rendements obtenus mais, dans le même temps, permettent de diminuer les charges en intrants chimiques, si bien que les marges brutes induites sont souvent comparables à celles obtenues en production conventionnelle intensive (Rolland et al., 2003). Ces combinaisons « itinéraires techniques à bas niveaux d'intrants/variétés rustiques » pourraient donc répondre à la fois à l'objectif de rentabilité économique et à l'objectif environnemental.

Notre projet, en collaboration avec des agronomes, dont le modèle développé dans cette thèse sera un outil central, s'articulera autour de deux questions principales : l'analyse des effets de l'utilisation des combinaisons « variétés rustiques / ITK à bas niveaux d'intrants » en terme économique (offre, revenu des agriculteurs,..) ; et l'analyse des déterminants économiques et politiques de l'utilisation de ces pratiques innovantes, le but étant de comprendre comment les pouvoirs publics, à travers divers instruments de politique, peuvent inciter les agriculteurs à changer leurs pratiques agricoles.

La première étape de ce projet, qui consistait à spécifier et estimer des fonctions de production intégrées dans un modèle de comportement de l'agriculteur à partir de

données économiques, a fait l'objet de cette thèse.

La deuxième étape consistera à estimer ces mêmes fonctions de production sur la base de données agronomiques : des données expérimentales issues du réseau ITK blé rustique et des données issues d'un modèle agronomique, le modèle Betha (Loyce et al., 2001 a et b). Le réseau de test d'itinéraires techniques adaptés aux variétés rustiques de blé tendre (réseau ITK blé rustique), regroupe l'INRA, Arvalis-institut du végétal et les sélectionneurs privés du Groupe d'Intérêts Economiques GIE « Club des 5 ». Il a été lancé en 1999, à l'initiative de l'INRA, pour tester les performances agronomiques et économiques de différents couples « variété-itinéraires techniques ». Le réseau ITK blé rustique compare quatre itinéraires techniques : un itinéraire « potentiel » où l'on cherche à atteindre le rendement maximum permis par le sol et le climat, un itinéraire « raisonné » qui vise un rendement élevé avec un objectif calé sur le rendement régional, un itinéraire « intégré » où l'on vise une diminution des intrants dans une logique calculée des risques et où la baisse des rendements est acceptée et un itinéraire « extensif » impliquant une réduction très importante des intrants. L'utilisation du modèle Betha (Loyce et al., 2001 a et b) vise à élargir le champ des itinéraires techniques considérés, en simulant, à partir du nombre limité d'expérimentations disponibles, le comportement de différentes variétés de blé sous d'autres ITK. L'estimation des fonctions de production sur ces données agronomiques permettra alors d'intégrer dans notre modèle l'innovation que représentent les nouvelles variétés de blé associées à des itinéraires techniques à bas niveaux d'intrants.

Enfin, dans une troisième étape, la mise en œuvre de divers instruments de politique sera simulée afin d'analyser leurs impacts sur les pratiques des agriculteurs d'une part, de mesurer les conséquences de la prise en compte de connaissances agronomiques au regard de l'analyse de l'impact des politiques publiques, d'autre part. L'analyse sera centrée sur trois types d'instruments de politique différents. Les pouvoirs publics ont plusieurs possibilités pour inciter les agriculteurs à adopter une nouvelle technologie, soit ils font en sorte que l'ancienne technologie devienne moins rentable pour les agriculteurs, en

augmentant les coûts liés à cette ancienne technologie (taxation des intrants, par exemple), ou en limitant l'intérêt d'utiliser l'ancienne technologie (suppression des raccourcisseurs, par exemple), soit ils privilégient la nouvelle technologie (subventions à l'achat des semences de variétés rustiques, par exemple).

Ce projet de collaboration avec les agronomes a été élaboré dans un contexte de prix bas du blé, donc dans un contexte compatible avec des baisses de rendement. La crise de fin 2007 début 2008 a changé la donne et l'idée de la baisse des rendements est moins bien acceptée, surtout par les agriculteurs. Cependant, ce contexte économique pourra être testé et permettre de simuler l'impact de différents instruments de politique sur l'adoption des variétés rustiques/ITK bas intrants, dans différents contextes de prix. De plus, la réduction de moitié de l'utilisation de pesticides décidée à l'issue du Grenelle de l'environnement va inciter les agriculteurs à revoir leurs itinéraires techniques. Dans ce cadre, la prise en compte des rotations devient encore plus pertinente puisque la rotation permet de réduire la dépendance de la production vis-à-vis de l'utilisation des pesticides. Ce projet reste donc en plein dans l'actualité.

Références.

- Arnberg, S. et L.G. Hansen. 2007. "The Dynamics of Farm Land Allocation – Short and Long Run Reactions in a Long Micro Panel." American Agricultural Economics Association Annual Meeting, Portland, July 29 - August 1, 2007.
- Bouchard, C., Bernicot, M.H., Félix, I., Guérin, O., Loyce, C., Omon, B. et B. Rolland. 2008. "Associer des Itinéraires Techniques de Niveau d’Intrants Variés à des Variétés Rustiques de Blé Tendre : Evaluation Economique, Environnementale et Energétique." *Courrier de l’environnement de l’INRA* 55:53-77.
- Debaeke, P. et J.M. Nolot. 2002. "Evaluation de Systèmes de Cultures à Bas Niveaux d’Intrants : Aspects Agronomiques et Environnementaux." *Dossier de l’environnement de l’INRA* 24:59-61.
- Eckstein, Z. 1984. "A Rational Expectations Model of Agricultural Supply." *The Journal of Political Economy* 92(1):1-19.
- Félix, I., Loyce, C., Bouchard, C., Meynard, J.M., Bernicot, M.H., Rolland, B. et H. Haslé. 2002. "Associer des Variétés Rustiques à des Niveaux d’Intrants Réduits : Intérêts Economiques et Perspectives Agronomiques." *Perspectives Agricoles* 79:30-35.
- Félix, I., Loyce, C., Bouchard, C., Meynard, J.M., Rolland, B., Bernicot, M.H. et H. Haslé. 2003. "Une des Voies pour s’Adapter aux Baisses de Prix du Blé : Des Variétés Rustiques Conduites à Faible Coût." *Perspectives Agricoles* 290:22-29.
- Livingston, M., Roberts, M.J. et J. Rust. 2008. "Optimal Corn and Soybean Rotations." Paper presented at AAEA annual meeting, Orlando, Florida, 27-29 July.
- Loyce, C., Bouchard, C., Meynard, J.M., Doussinault, G., Rolland, B., Bernicot, M.H. et H. Haslé. 2001. "Les Variétés Tolérantes aux Maladies : Une Innovation Majeure à Valoriser par des Itinéraires Techniques Economes." *Perspectives Agricoles* 268:50-57.
- Loyce, C. et J.M. Meynard. 1997. "Low Input Wheat Management Techniques are More Efficient in Ethanol Production." *Industrial Crops and Products* 6:271-283.
- Loyce, C., Rellier, J.P. et J.M. Meynard. 2002. "Management Planning for Winter with Multiple Objectives (1) : The BETHA System." *Agricultural Systems* 72:9-31.

- Loyce, C., Rellier, J.P. et J.M. Meynard. 2002. "Management Planning for Winter with Multiple Objectives (2) : Ethanol-Wheat Production." *Agricultural Systems* 72:33-57.
- Meynard J.M. et I. Savini. 2003. "La Désintensification : Point de Vue d'un Agronome." *Dossier de l'environnement de l'INRA* 24:23-33.
- Moore, M. R. et D. H. Negri. 1992. "A Multicrop Production Model of Irrigated Agriculture, Applied to Water Allocation Policy of the Bureau of Reclamation." *Journal of Agricultural and Resource Economics* 17:29-43.
- Ozarem, P.F. et J.A. Miranowski. 1994. "A Dynamic Model of Acreage Allocation with General and Crop-Specific Soil Capital." *American Journal of Agricultural Economics* 76(3):385-395.
- Rolland, B., Bouchard, C., Loyce, C., Meynard, J.M., Guyomard, H., Lonnet, P. et G. Doussinault. 2003. "Des Itinéraires Techniques à Bas Niveaux d'Intrants pour des Variétés Rustiques de Blé Tendre : Une Alternative pour Concilier Economie et Environnement." *Courrier de l'environnement de l'INRA* 49:53-77.
- Rolland, B., Oury, F.X., Bouchard, C. et C. Loyce. 2006. "Vers une Evolution de la Création Variétale pour Répondre aux Besoins de l'Agriculture Durable? L'Exemple du Blé Tendre." *Dossier de l'environnement de l'INRA* 30:79-90.
- Rust, J. 1987. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica* 55:999-1033
- Saulas, P. et J.M. Meynard. 1998. "Production Intégrée et Extensification sont-elles Compatibles ? Cas des Céréales à Paille." *Dossier de l'environnement de l'INRA* 16:9-15.
- Tegene, A., Huffman, W.E. et J. Miranowski. 1988. "An Econometric Model of Dynamic and Stochastic Production with Explicit Optimization: Farmers' Decisions on Corn and Soybean" *American Journal of Agricultural Economics* 70:103-111.
- Thomas, A. 2003. "A Dynamic Model of On-Farm Integrated Nitrogen Management." *European Review of Agricultural Economics* 30(4):439-460.
- Train, K. 2003. *Discrete Choice Methods with Simulation*. Cambridge University Press.