



Le marché de la donnée au service de la transparence du marché du logement ?

Alexandre Coulondre, Claire Juillard

► To cite this version:

Alexandre Coulondre, Claire Juillard. Le marché de la donnée au service de la transparence du marché du logement ?. 2022. halshs-03931580

HAL Id: halshs-03931580

<https://shs.hal.science/halshs-03931580>

Preprint submitted on 9 Jan 2023

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.

Working Paper

> N°01/2022

Le marché de la donnée au service de la transparence du marché du logement ?

Alexandre Coulondre
Claire Juillard

SciencesPo

CITIES AND DIGITAL TECHNOLOGY CHAIR

The “Cities and Digital Technology” Chair of Sciences Po’s Urban School has been launched in March 2017 to better grasp the impact of digital technologies on urban governance. Funded by four sponsoring firms (La Poste, RTE, Caisse des Dépôts), the Chair aims to create new research fields exploring the interaction between digital technology and cities in an empirical and comparative perspective.

Le marché de la donnée au service de la transparence du marché du logement ?

Alexandre Coulondre, chercheur associé au LATTS (Ecole des Ponts Paristech) et chercheur indépendant

alexandre.coulondre@gmail.com

Claire Juillard, chercheuse indépendante, fondatrice et présidente d'OGGI Conseil

claire.juillardm@gmail.com

Résumé

Les données de prix des logements accessibles gratuitement aux particuliers sont de plus en plus nombreuses. L'accroissement de l'offre tient à l'intervention conjointe de deux catégories d'acteurs : d'une part l'Etat, dont la mise en œuvre de la politique d'Open Data dans le champ de l'immobilier a abouti à l'ouverture récente de données massives sur le logement (notamment sur les transactions avec le fichier Demande de Valeurs Foncières - DVF) ; d'autre part les diffuseurs de données eux-mêmes, des entreprises de la *Proptech* dont le nombre et l'offre d'informations augmentent avec l'accélération de la digitalisation de l'immobilier.

Ce working paper décrit comment l'action conjuguée de ces acteurs a suscité l'émergence d'un marché de la donnée de prix à destination des particuliers. Dans ce marché, l'information aux particuliers alimente un flux de données sur les particuliers eux-mêmes que les opérateurs exploitent à des fins commerciales. Sa valeur renvoie au modèle qui domine dans l'écosystème : l'économie de plateforme.

L'Etat et les entreprises du marché de la donnée de prix affichent un objectif commun : la transparence du marché du logement. Ils promeuvent l'idée, empruntée à la théorie économique, selon laquelle l'information de tous concourt à la bonne évaluation et à la bonne fixation du prix des logements. Ce working paper illustre comment cette notion de transparence se confronte dans les faits aux pratiques concrètes des acteurs et notamment à celles des particuliers vendeurs de logements. L'ensemble apporte des nuances significatives à la théorie. Premièrement, la consommation de données est certes une pratique répandue mais elle s'avère socialement et géographiquement située. Deuxièmement, quand elle est utilisée, la donnée a un effet conditionné sur la transparence des marchés. Elle procure un sentiment de transparence aux vendeurs qui en consomment avec modération mais floute la vision de ceux qui en consomment beaucoup. Troisièmement, elle favorise les paris gagnants mais n'a pas forcément un effet modérateur sur les prix et peuvent même favoriser un alignement sur les valeurs hautes du marché.

Mots clés : marchés du logement, données, open data, PropTech.

Introduction

Ce working paper s'inscrit dans un programme de recherche sur les données de prix des logements en France. Il fait suite à un premier working paper consacré à l'écosystème de la production de ces données et à la place et au rôle que l'Etat y tient (Juillard 2020). Il prolonge et déplace le regard de deux manières : en s'intéressant spécifiquement aux données destinées au grand public, pour partie différentes des données adressées aux acteurs de l'immobilier qui font l'objet du premier papier ; en se penchant sur les usages de ces données, plutôt que sur les modalités de leur production. Il s'arrête par ailleurs aux données relatives au prix des biens sur le marché de l'ancien, contrairement au précédent working paper qui avait vocation à couvrir l'ensemble des marchés privés du logement.

Le déplacement opéré s'appuie sur deux constats. Premièrement, les données de prix des logements accessibles gratuitement aux particuliers sont de plus en plus nombreuses. Elles composent une mosaïque qui prend place dans un vaste paysage informationnel marqué par la diversité des types d'informations disponibles et par la multiplicité de leurs canaux d'accès. La donnée n'est pas toute l'information disponible mais elle en représente la principale variable d'extension.

Deuxièmement, l'accroissement de l'offre de données tient à l'intervention conjointe de deux catégories d'acteurs : l'Etat, dont la mise en œuvre de la politique d'Open Data dans le champ de l'immobilier a abouti à l'ouverture récente des données exhaustives sur les ventes de logements (Demande de Valeurs Foncières - DVF) ; les diffuseurs de données eux-mêmes, dont le nombre et l'offre d'informations augmentent avec l'accélération de la digitalisation de l'immobilier.

Les uns comme les autres affichent un objectif commun : la transparence du marché. Ils promeuvent l'idée, empruntée à la théorie économique, selon laquelle l'information de tous concourt à la bonne évaluation et à la bonne fixation du prix des logements. L'enjeu est d'intérêt général. Il vise la correction d'un marché haussier marqué par des tensions croissantes. Sa formulation illustre en particulier la figure d'un Etat qui tend à gouverner par le marché (Dubuisson-Quellier 2016 ; Ansaloni, Trompette, Zilio 2017) et favorise pour ce faire l'innovation et la création de services à destination des citoyens et sources d'activité économique.

Trois ans après l'ouverture de DVF, ce working paper illustre combien les fournisseurs de données de prix ont tiré parti de la politique d'Open Data. Les usages qu'ils font des données ouvertes révèlent une caractéristique du marché du logement à l'heure de la digitalisation de l'immobilier : l'information sur les prix est moins une condition de fonctionnement du marché tel que l'établit la théorie qu'un marché à part entière. Une raison tient à ce que l'information aux particuliers tirée ou consolidée par DVF alimente un « flux de données » sur les particuliers (Migozzi 2020) que les opérateurs exploitent à des fins commerciales. Sa valeur renvoie au modèle qui domine dans l'écosystème : l'économie de plateforme. Elle croît avec

l'extension du périmètre d'intervention des entreprises de la PropTech (pour *Property Technology*) dont les fournisseurs de données font partie pour une large part.

Ce working paper s'attache également à soumettre la théorie économique à l'épreuve des pratiques d'une autre catégorie d'acteurs : les particuliers vendeurs de logements. Il est tiré d'une étude qui les a choisis pour objet en raison de leur rôle central dans la fixation des prix des logements. Cette étude a principalement consisté en la réalisation d'une enquête par questionnaires destinée à saisir les conditions concrètes dans lesquelles les vendeurs approchent la valeur possible de leur bien et la place éventuelle de la donnée dans leur démarche¹.

L'ensemble apporte des nuances significatives à la théorie économique qu'il révèle plutôt comme un argument que comme un corpus scientifique explicatif des phénomènes. Sur cette base, le présent working paper se décompose en trois parties : une première, de cadrage, sur le marché de la donnée, ses acteurs et leur modèle de développement ; une seconde, centrale, de confrontation de l'objectif de transparence aux pratiques de documentation des vendeurs de logements ; une troisième, plus générale, qui dépasse les bornes du marché de la donnée et s'attache à l'articulation de l'ensemble du paysage informationnel des vendeurs avec le marché du logement.

1 La construction (politique) d'un marché de la donnée de prix

Les données de prix des logements sont de plus en plus nombreuses. Leur extension reconfigure le paysage informationnel dans lequel les vendeurs de logements puisent des ressources pour approcher la valeur possible de leur logement. Elle se nourrit de deux tendances : l'accélération de la politique d'Open Data de l'Etat qui a récemment ouvert DVF, sources d'informations inédites extraites des actes authentiques de vente de logements enregistrés par les offices notariaux pour le compte de l'administration fiscale ; le développement de la PropTech, écosystème d'entreprises (des « starts-ups ») qui, dans le contexte de la digitalisation de l'immobilier, essaient la chaîne de la transaction de nouveaux outils et services innovants à destination des particuliers.

Le déploiement de nouvelles données de prix révèle la construction d'un marché en soi. Sur ce marché, la donnée n'a pas de prix pour les particuliers, elle est gratuite. Mais dans de nombreux cas, elle crée de la valeur économique car elle est diffusée en échanges d'informations personnelles fournies par les ménages. Le modèle économique des

¹ L'enquête par questionnaires a été réalisée avec la souscription financière de Crédit Logement, du Groupe BPCE, du LIFTI et du PUCA et menée entre décembre 2020 et janvier 2021 auprès de personnes ayant vendu un logement en France en 2019 ou en 2020, qu'il s'agisse ou non d'une résidence principale. La rédaction du questionnaire a été assurée sur la base d'un premier volet qualitatif ayant permis de réaliser des entretiens semi-directifs avec une dizaine de vendeurs et autant d'agents immobiliers. La passation du questionnaire a été assurée par l'institut Opinion Way. Il s'agissait d'un questionnaire auto-administré en ligne. Pour le rapport tiré de son exploitation principale, voir : Coulondre, Juillard 2021.

fournisseurs de données tient à ce qu'ils tirent de l'information offerte *aux* particuliers des informations *sur* les particuliers qu'ils revendent sous forme de « contacts qualifiés » à des agences immobilières. L'opération répond à un enjeu central pour les acteurs : la captation de prospects et la signature de nouveaux mandats. Elle aligne l'ensemble des acteurs de la chaîne autour d'un intérêt commun et conforte leur partenariat plus qu'elle ne les divise. L'économie de plateforme en constitue le modèle sous-jacent.

De l'information sur les prix à la donnée de prix

Les informations à disposition des particuliers sur les prix des logements ne manquent pas. On peut les décomposer en quatre types (Figure 1) :

- **le prix des biens en vente**, disponible dans les annonces affichées en agence ou sur des portails internet dédiés ;
- **le prix des biens vendus**, qu'il soit disponible auprès de professionnels, au moyen de cartes ou de listes de ventes interactives sur internet, ou encore sous la forme d'un affichage de biens vendus en agence ou en ligne ;
- **l'estimation**, qu'elle soit réalisée « à la main » par un professionnel de l'immobilier à partir de son expérience et de sa connaissance du terrain ou par un outil automatique utilisé par le professionnel ou par le vendeur lui-même sur Internet ;
- **le prix moyen** ou médian au mètre carré à la commune ou au quartier généralement disponible sur Internet ou dans la presse spécialisée.

Les canaux possibles pour accéder aux informations sur les prix des logements sont également multiples et peuvent eux-mêmes se décomposer en quatre catégories :

- **le canal physique des professionnels** de l'immobilier (agents, mandataires, notaires) ;
- **le canal physique de l'entourage** ;
- **le canal digital**, qui regroupe sites Internet et applications mobiles ;
- **le canal papier**, fait de journaux et de magazines.

Le paysage informationnel formé à partir de la combinaison de ces types d'informations et de leurs canaux d'accès est particulièrement divers. Il a également tendance à s'étendre. Son extension prend principalement racine dans la digitalisation de l'immobilier. Concrètement, elle tient à l'éclosion du canal digital d'accès à l'information. Elle suit le rythme soutenu du développement de la PropTech depuis une dizaine d'années et plus récemment, profite de l'accélération de la mise en œuvre de la politique d'Open Data de l'Etat dans le domaine du logement.

Le canal digital irrigue une catégorie d'information particulière : la donnée. Celle-là prend la forme de prix des biens vendus, d'estimations et de prix moyens (ou médians) au mètre carré. Parmi les quatre types d'informations disponibles en d'autres termes, seuls les prix des biens en vente qui s'affichent dans les annonces immobilières en ligne n'en sont pas. Une raison en

est que les prix indiqués dans les offres de logement, contrairement aux autres types d'informations accessibles sur la toile, obéissent à aucune logique d'ingénierie statistique et informatique (Figure 1).

Sur le canal digital, les sites internet et applications qui mettent à disposition gratuitement des données de prix en France sont de l'ordre d'une soixantaine à ce jour (Figure 2). Pour une part significative d'entre eux, ils mêlent tout ou partie des trois catégories de données de prix des logements et peuvent donc fournir à la fois des prix des biens vendus, des estimations et des prix moyens (ou médians) au mètre carré. Cette propension à diversifier le type de données offertes contribue à son tour à étendre l'ensemble paysage informationnel des particuliers.

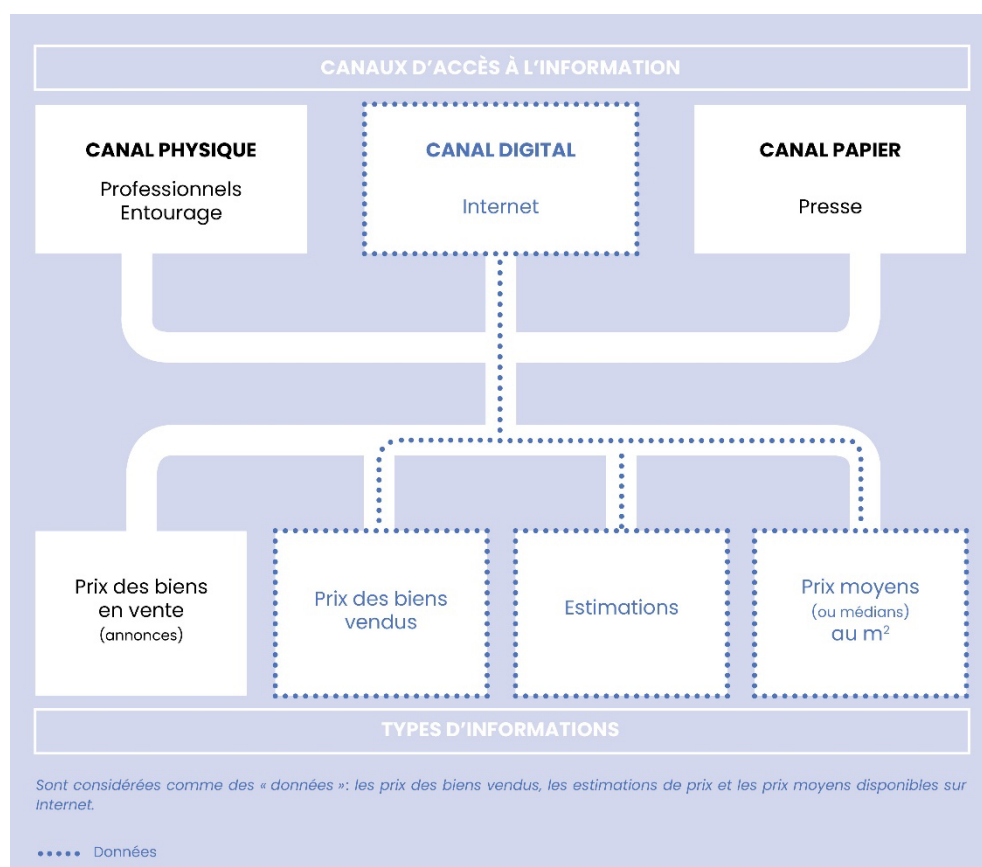


Figure 1. La donnée dans le paysage informationnel sur les prix des logements

La consolidation d'un marché de la donnée par la politique d'Open Data

Les fournisseurs de données sur le canal digital représentent plusieurs catégories d'acteurs : des portails d'annonces immobilières (leboncoin, SeLogger, Bien' Ici, etc.) ; des spécialistes de l'information immobilière (MeilleursAgents, La Cote Immo, Bien Estimer, etc.) ; des services en ligne (les agences digitales et ebuyeurs par exemple) ; des médias (Le Figaro immobilier, BFM Immo, Capital, etc.) ; les portails du notariat (source historique de références de prix et de statistiques immobilières) ; des banques (BNP Paribas, Banque Populaire, Caisse

d'Epargne) ; et l'Etat, à travers Etalab, département interministériel en charge de la mises à disposition des données DVF.

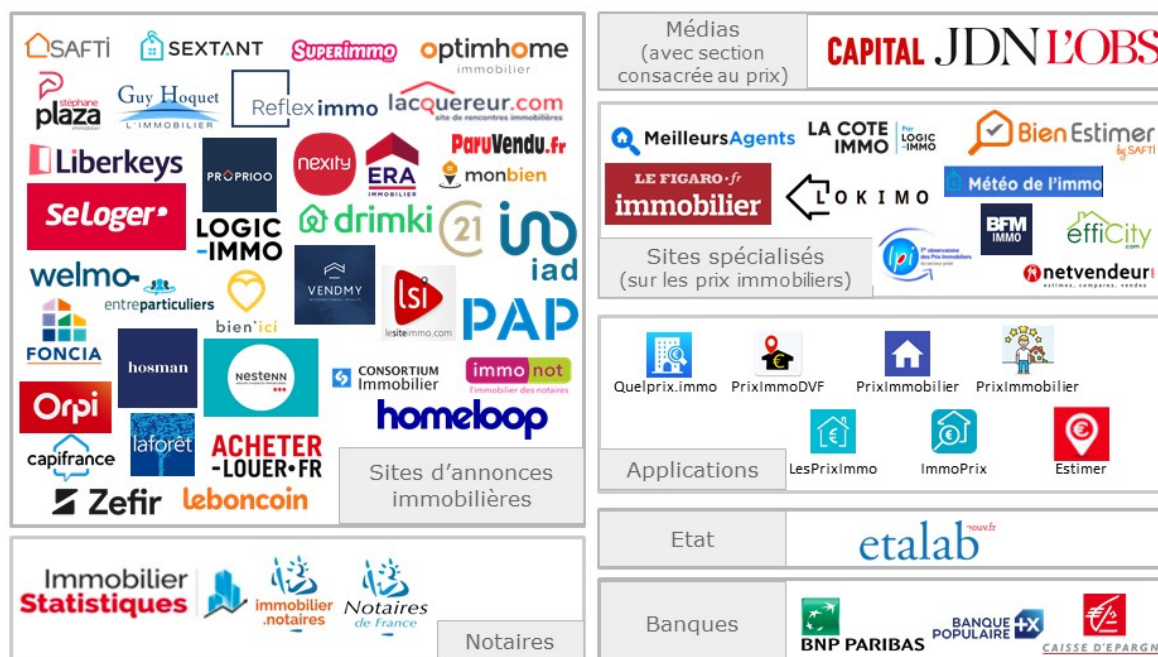


Figure 2. Sites Internet et applications fournissant des données de prix à destination des particuliers des logements en France (marché de l'ancien, juin 2021)

Parmi ces fournisseurs de données aux particuliers, l'Etat à travers Etalab joue un rôle qui lui est inhabituel. Il fait figure de nouvel entrant mais marque d'emblée un tournant dans l'information aux particuliers avec son site AppDVF qui fournit la première interface de consultation gratuite des données DVF. Le site a rencontré un succès d'audience inattendu au moment de son lancement (Juillard 2020). Pour autant, les vendeurs de logements ne l'utilisent que marginalement relativement aux autres services d'accès à DVF disponibles sur la toile, à commencer par les cartes interactives développées dans la foulée de l'AppDVF par MeilleursAgents et SeLoger, deux sites centraux dans le paysage.

Etalab n'est pas seulement fournisseur de données aux particuliers. Il joue aussi et en premier lieu le rôle de fournisseur aux professionnels. L'ouverture de DVF l'a effectivement placé au cœur de l'écosystème de la production de données de prix (Juillard 2020). Elle a agi de plusieurs manières. D'abord, on vient de le voir, elle a favorisé la création d'une nouvelle offre de données. Elle n'a d'ailleurs pas seulement profité aux acteurs déjà solidement ancrés dans le paysage. Elle a aussi suscité l'entrée de nouveaux, comme les applications quelpriximmo, PriximmoDVF, ImmoPrix etc. qui fournissent elles aussi une interface d'accès gratuit à DVF. Ensuite, de manière moins visible (et plus « opaque »²), DVF a alimenté d'une source officielle et exhaustive les modèles de production des fournisseurs de données déjà en place auparavant, à commencer par MeilleursAgents et SeLoger. Ces modèles statistiques apprenants (ils empruntent notamment au *machine learning*) reposent sur un *credo* : la fiabilité

² Boulay, Blanke, Casanova, Granié 2020.

de leurs résultats est fonction de la qualité et de la quantité d'informations qui les sous-tendent. L'ouverture à tous de DVF les a donc confortés.

Enfin, l'ouverture de DVF a accompagné le développement de nouveaux services qui, comme les agences en ligne (Hosman, Liberkeys, Propioo, etc.) et les ebuyeurs (Homeloop et Zefir par exemple), se centrent sur la digitalisation de l'immobilier. Parmi ces derniers d'ailleurs, les ebuyeurs accélèrent la vente immobilière en procédant eux-mêmes à l'achat dit « instantané » des biens des vendeurs et à leur revente à des candidats à l'achat. Ils utilisent DVF d'au moins deux manières en complément des autres sources dont ils disposent : non seulement pour produire des données à destination des particuliers (des estimations automatiques essentiellement), mais aussi pour conforter leur propre modèle d'estimation, une clé dans la réalisation du pari qu'ils font sur le prix de revente des biens dont ils font l'acquisition.

L'ouverture de DVF a donc porté ses fruits au-delà de l'information apportée aux particuliers. Cette ouverture a concrétisé une ambition forte bien que moins connue de la politique d'Open Data de l'Etat : l'incitation à l'innovation et à la création de services marchands à la fois utiles aux citoyens et sources d'activité économique pour les entreprises. Ainsi, bien que la politique d'ouverture des données immobilières amorcée dans les années 2010 soit relativement récente, elle a été déterminante dans la consolidation et l'extension de l'écosystème de la donnée immobilière. En ouvrant des données massives et exhaustives, cette politique a fourni une solution technique à la fois aux entreprises ayant déjà des données mais partielles (les portails d'annonces par exemple) et aux entreprises qui, ayant peu ou pas de données auparavant, ont pu se constituer en nouveaux entrants. Plus encore, la politique d'Open Data a apporté une légitimité aux données offertes aux particuliers en leur ajoutant une caution officielle : celle de leur source principale.

La donnée de prix : une économie de plateformes

Une question demeure : quelle logique économique y a-t-il pour les entreprises de la PropTech à fournir gratuitement autant de données aux particuliers ? Pour la plupart, les fournisseurs de données de prix font partie de la même famille : celle de l'économie numérique. Ils s'appuient sur un modèle de plateforme. Leur créneau est la mise en relation des particuliers vendeurs (et/ou acquéreurs) et des professionnels de la transaction immobilière. Leur succès dépend de leur capacité à attirer et à « garder à bord » chacune de ces deux catégories d'acteurs (Rochet et Tirole 2004 ; Brailly et Coulondre 2016). La donnée de prix en est une des principales clés. Elle complète un autre puissant levier de captation de la demande : l'annonce immobilière.

Le succès du modèle de plateforme en immobilier paraît aujourd'hui acquis dans un certain nombre de cas. Pour preuve, les sites et applications d'immobilier enregistrent une vingtaine de millions de visiteurs uniques par mois. Parmi les plus consultés, leboncoin affichent plus de douze millions de visiteurs uniques fin 2019. Bien que loin derrière SeLoger en compte tout de

même 6,5 millions, Bien'ici et Logic-Immo près de 4 millions, PAP et AVendreALouer environ 3 millions, MeilleursAgents et le Figaro plus de 2 millions³.

Le trafic que les sites et applications d'immobilier absorbent est d'abord un trafic de consultation d'annonces. Le record d'audience du site leboncoin tient précisément à la quantité inégalée d'annonces de logements à vendre qu'il publie⁴. Il est en partie assuré par un positionnement inédit dans le paysage, à la jonction du marché intermédié et du marché désintermédié (de particuliers à particuliers), plutôt que sur l'un ou sur l'autre segment tels que les autres plateformes se positionnent essentiellement. Le modèle économique du site repose sur le flux de consultations d'annonces qu'il génère. La visibilité que les agences immobilières et les particuliers vendeurs en retirent a un prix. Elle constitue la principale composante de l'abonnement qu'ils paient pour publier leur offre de biens à vendre sur la plateforme.

L'ensemble des fournisseurs de données immobilières fondent aujourd'hui leur activité sur ce modèle de création de valeur par la consultation d'annonces. Tous reposent également sur un autre levier : la consultation de données. Premier site dédié à l'information immobilière et leader de l'estimation en ligne, MeilleursAgents a montré la voie. Cette plateforme offre gratuitement aux internautes tous les types de données disponibles sur les prix des logements, c'est-à-dire à la fois des informations sur le prix des biens vendus, des prix moyens et des estimations automatiques.

Parmi ces données, l'estimation tient une place à part. Elle est fournie à l'internaute à condition qu'il accepte de renseigner un formulaire le décrivant et appréciant la maturité et la teneur de son projet immobilier. MeilleursAgents ce faisant tire de l'information qu'il offre aux particuliers des informations sur les particuliers eux-mêmes. Elle alimente ainsi des listes de prospects qualifiés qu'elle monnaie auprès de ses agences immobilières partenaires. L'établissement de ces listes répond à un enjeu central dans l'activité d'une agence : la signature de nouveaux mandats. Il aligne la plateforme et ses entreprises clientes autour de l'intérêt commun de capter des internautes.

C'est un fait, les plateformes immobilières diffusent à ce jour la quasi totalité de l'offre de logements à vendre. Mais elles se positionnent à l'interface des agences et de leurs clients et ont tendance à conforter la position des premiers sur le marché de la transaction. C'est d'ailleurs le cas des plateformes qui ciblent les vendeurs en premier lieu, à l'instar de MeilleursAgents, comme celles qui ciblent les acquéreurs de manière privilégiée, à l'instar de SeLoger qui adressent donc plutôt des acheteurs potentiels à ses agences partenaires. Au final, les plateformes ont créé un marché de la donnée à part entière. La concurrence qui se joue entre elles tient notamment à leur capacité à délivrer des données qui attirent des internautes à elles.

3 Internet Global, Médiamétrie-Netratings, octobre 2019. Les derniers chiffres d'audience publiés par Médiamétrie date du printemps 2021 mais ils ne sont pas retenus ici car jugés trop dépendants de l'effet de la crise sanitaire sur le marché du logement.

⁴ Ce chiffre n'est pas disponible pour l'année 2019 mais il avoisine le million d'annonces à ce jour.

Un révélateur des formes contemporaines d'action publique

En irriguant le marché immobilier de données sur les prix des logements, l'Etat affiche un premier objectif : la transparence des marchés. Au moment de l'ouverture de DVF, Gérard Darmanin, alors Ministre de l'action et des comptes publics, explicite lui-même l'effet attendu d'un surcroît d'informations économiques : « DVF permettra à tous les professionnels mais aussi à tous les citoyens (...) de bien évaluer les biens. Ils pourront ainsi je l'espère permettre à la concurrence pure et parfaite de s'exprimer » (allocution du 24 avril 2019 à la Direction Générale des Finances Publiques, Paris). Les propos du Ministre sont emblématiques de deux principes structurants de l'action publique contemporaine : d'une part, la figure d'un Etat qui entend moins gouverner « le » marché que gouverner « par » le marché (Dubuisson-Quellier 2016) ; d'autre part, l'usage performatif de la théorie économique (Muniesa et Callon, 2009). Les propos du Ministre sont clairs : en œuvrant à l'information de l'ensemble des parties prenantes de la transaction immobilière, l'Etat ambitionne de garantir l'établissement du juste prix et la correction d'un marché du logement haussier aux tensions croissantes. Les fournisseurs de données, non sans une certaine ambivalence lorsqu'ils tiennent la croissance des prix comme un signe de bonne santé du marché, marquent aussi leur adhésion à cet enjeu d'intérêt général en revendiquant leur propre contribution à la transparence. Parmi eux, certains ont même précédé l'Etat, à l'instar de MeilleursAgents qui, en pionnier de l'information immobilière aux particuliers, a d'emblée placé la transparence au cœur de sa stratégie marketing.

Alors que le politique a la possibilité d'emprunter des modes d'intervention directs et contraignants, ce qu'il fait parfois lorsqu'il agit sur le plafonnement des loyers par exemple, il intervient ici à travers des dispositifs doux qui ne limitent pas l'autonomie des acteurs. L'Etat entend agir indirectement sur la formation des prix *via* la consolidation d'un secteur économique centré sur le développement de services visant la transparence des marchés. A ce titre, l'Open Data peut être classée dans la catégorie de plus en plus large des politiques consistant à « produire de l'hétéronomie en jouant sur l'autonomie des acteurs » (Bergeron et al., 2014). Il s'agit de profiler des dispositifs marchands de telle sorte qu'ils puissent faire converger les intérêts économiques et les projets politiques (Ansaloni, Trompette, Zilio 2017). Cette forme d'action publique repose sur la croyance fondamentale en l'efficacité des marchés. En l'espèce, la notion de transparence est empruntée à la théorie économique néo-classique. L'allocution du Ministre y fait directement référence en citant la notion de « concurrence pure et parfaite ». Elle convoque la théorie non pas comme une grille de lecture pour décrypter les phénomènes existants mais comme un modèle idéal visé par les acteurs qui organisent préalablement le marché afin qu'il s'y conforme. En d'autres termes, la théorie ici ne vise pas tant à décrire le monde qu'à le provoquer dans le sens d'une prophétie auto-réalisatrice. On touche au caractère performatif de la science économique (Muniesa et Callon, 2009) et de son recours par les acteurs politiques.

La question centrale qui se pose est celle des conséquences en termes de régulation. Cette forme d'action publique centrée sur le marché et se mettant en œuvre dans des configurations d'acteurs multiples à la fois publics et privés, historiques et nouvellement entrés, peut produire

des effets imprévus. Qu'en est-il en dans le cas du logement ? La politique d'Open Data et la consolidation du marché de la donnée de prix ont-elles apporté plus de transparence ? Sont-elles en mesure de limiter les tendances haussières du marché dans les faits ? En d'autres termes, concourent-elles à l'équilibre du marché ?

2 Marché de la donnée et transparence du marché immobilier : la théorie à l'épreuve des faits

Le raisonnement qui relie la mise à disposition de données et la transparence des marchés s'appuie sur trois postulats implicites : le premier, selon lequel la mise à disposition d'une information implique nécessairement son usage de la part des acteurs économiques ; le second, selon lequel l'usage de cette information débouche assurément sur une meilleure connaissance des marchés ; le troisième selon lequel cet usage a un effet modérateur sur les prix. Les pratiques de documentation des vendeurs de logements nuancent la théorie postulat après postulat⁵.

Des données accessibles à tous mais pas utilisées par tous

Alors que 85% des vendeurs ont recours à des informations sur les prix des logements en général, près d'un vendeur sur quatre consomme spécifiquement des données sur le prix des logements. Parmi eux, 27% en mobilisent « peu » et 11% en mobilisent « beaucoup » (Figure 3).

La donnée est donc bien utilisée par un nombre significatif de vendeurs mais pas par la totalité. Elle n'est pas utilisée par tous les profils de vendeurs non plus, sinon plus ou moins intensément. En l'espèce, la consommation de données varie à la fois avec l'âge et la position sociale (Figure 4). D'un côté, les plus « gros » consommateurs ont tendance à être plus âgés et à occuper un rang plus élevé dans la hiérarchie socioprofessionnelle (cadres, chefs d'entreprise, etc.). De l'autre côté, les plus « petits » consommateurs concentrent les vendeurs moins âgés et les catégories socioprofessionnelles inférieures (ouvriers et employés). Il faut dire notamment que les personnes les moins diplômées et aux revenus modestes font partie des populations les moins bien équipées et les moins compétentes au plan numérique (Leglaye, Rolland 2019).

De façon assez contre-intuitive donc, l'évolution du paysage informationnel en faveur de la donnée contribuerait plutôt à renforcer une forme de « ségrégation informationnelle » (Boeing 2019) ; elle procéderait à une « division par la donnée » (« data divide ») et armerait les mieux armés (« empower the empowered ») (Gurstein 2010). En l'espèce, elle réduirait l'asymétrie

⁵ Cette partie sur les pratiques des vendeurs est tirée de l'enquête par questionnaires présentée dans l'introduction à ce working paper.

d'information en faveur des vendeurs les plus dotés mais renforcerait la rente de professionnels également de mieux en mieux informés face aux vendeurs les moins informés. De la même manière, la donnée n'est pas utilisée partout. 40% des plus grands consommateurs de données évoluent en marché tendu et à peine 15% en marché détendu (Figure 5). Les zones de tension immobilière concentrent davantage de consommateurs de données. On peut certes y voir l'effet des profils des vendeurs, eux-mêmes situés géographiquement, les catégories socio-professionnelles supérieures étant plus nombreuses en zone tendue qu'en zone détendue. On peut également y déceler la tentation plus grande de scruter les valeurs de marché dans des territoires où les prix croissent rapidement.

Enfin, le cas des estimations automatiques fournies par les sites et applications d'immobilier renvoie aux limites d'une démarche tendant à la couverture totale du territoire sans distinction de méthode. Si les outils à disposition des particuliers permettent d'estimer automatiquement un bien quel qu'il soit et où qu'il soit, les résultats qu'ils produisent sont plus ou moins fiables, tel que les fournisseurs s'attachent à le spécifier en accompagnant leurs résultats d'intervalles de confiance.

Ces résultats sont des boîtes noires. En matière d'estimation, sa composante métropolitaine s'en devine pourtant. On sait combien les plus grands marchés locaux sont aussi les plus et les mieux documentées en données. D'une part, les transactions s'y font nombreuses, alimentant par exemple la mécanique du *machine learning*. D'autre part, les acteurs de la PropTech s'y concentrent et avec eux, leurs services et outils.

On sait moins en revanche que les plus grands marchés locaux imprègnent la data jusque dans ses composantes. Le constat se fait pourtant sans peine lorsque l'on prend le temps de déplier les grilles d'indicateurs qui sous-tendent les moteurs d'estimation automatique à disposition des particuliers. Ces grilles, appliquées sans distinction à tous les territoires, comportent de nombreux indicateurs qualifiables de métropolitains, sinon d'urbains. La proximité aux transports collectifs en fait partie. C'est un critère certes pertinent en ville mais invalide à la campagne. Sa prise en compte révèle en quelque sorte l'existence d'une géographie de la donnée. Elle rappelle par ailleurs comment l'expertise de terrain, bien que située, peut se traduire en variable algorithmique généralisée.

Au final, contrairement à ce que sous-entend la théorie, il ne suffit pas que des données de prix soient accessibles à tous et partout pour qu'elles soient utilisées par tous et partout. La mise à disposition de données n'implique pas forcément leur utilisation. La bonne information des acteurs économiques ne tient pas seulement à une question d'offre de données. C'est aussi une question d'usage et de compétences (Gurstein 2010).

		Pourcentage	Dont :	Pourcentage
Score d'usage de la donnée	Beaucoup de données	11%	Beaucoup de supports	9%
			Peu de supports	2%
	Peu de données	27%	Beaucoup de supports	16%
			Peu de supports	11%
	Pas de données	62%		
Total		100%		

Figure 3. Degré d'usage de la donnée parmi les vendeurs de logements
Ensemble des vendeurs de logements (n=993)

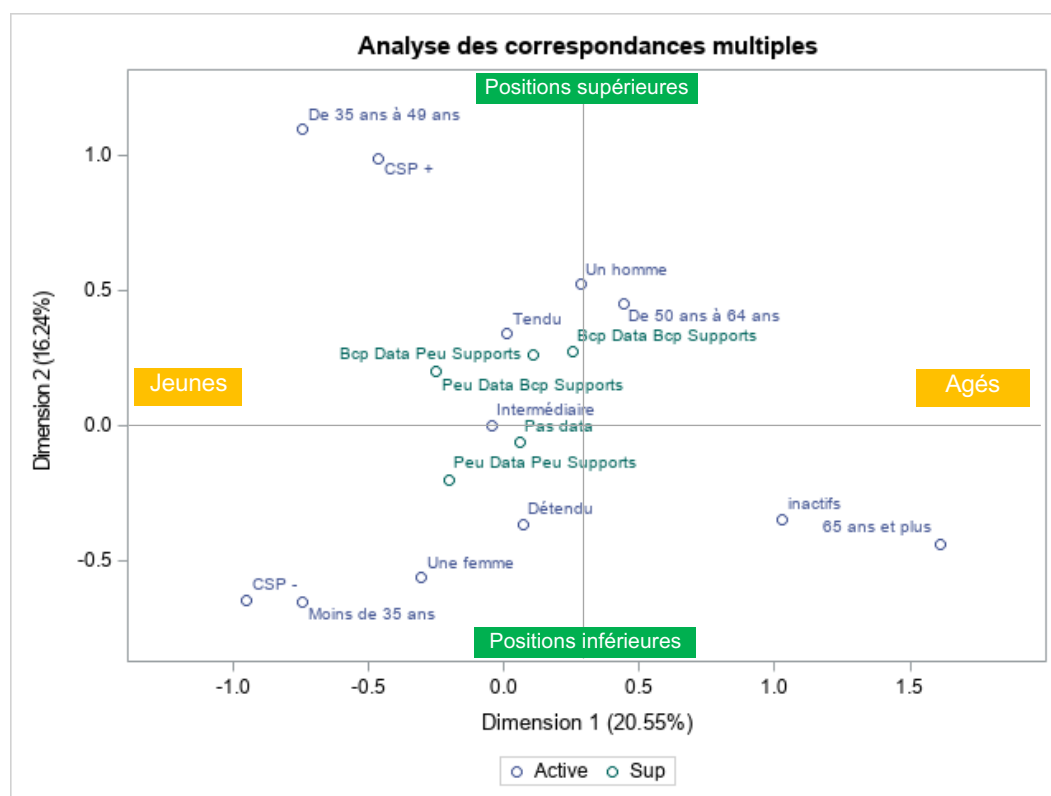


Figure 4. Liens entre consommation de données et profils des vendeurs (analyse factorielle)

Ensemble des vendeurs de logements (n=993)

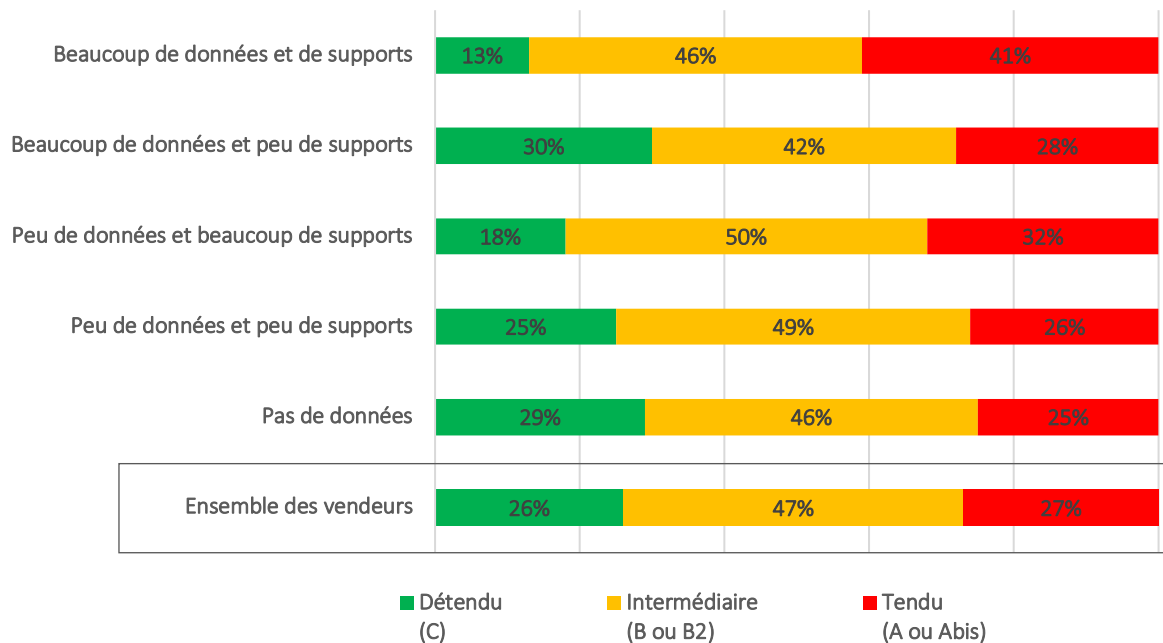


Figure 5. Consommation de données et tension du marché local

Ensemble des vendeurs de logements (n=993)

Pourcentages en ligne - Statistique du Khi-deux : $DDL=8$ / Valeur = 19,87 / Probabilité = 0,0108

Remarque : zonage « ABC » reconstitué par les auteurs à partir de la localisation du bien déclarée dans le questionnaire par le vendeur.

Un effet de transparence conditionnel

Comme l'établit la théorie, la donnée procure globalement un sentiment plus grand de transparence des marchés aux vendeurs qui en consomment. Mais au-delà d'un certain seuil, son utilisation floute la vision du marché. En effet, les plus grands consommateurs de données sont deux fois plus nombreux que les petits consommateurs à entretenir le sentiment que le marché immobilier manque de transparence (Figure 6).

Pourquoi ? D'abord, parce qu'interpréter les données disponibles sur les marchés est un exercice complexe. Ensuite, parce que les données se font de plus en plus nombreuses, notamment sur Internet et qu'elles sont difficiles à comparer. Valeurs estimées, prix de vente, prix moyens : l'ensemble dessine un vaste paysage informationnel dans lequel se repérer s'avère délicat.

Parmi les données, les estimations et les prix moyens sont d'autant plus difficiles à comparer que les différents sites Internet et applications qui en fournissent peuvent afficher des résultats très divergents pour un même secteur géographique ou pour un même bien estimé. En effet, entre deux estimations d'un même bien et entre deux prix moyens pour une même zone géographique, les valeurs proposées par les fournisseurs de données peuvent quasiment varier de un à deux. La confusion ne peut que s'instaurer et elle s'accroît encore lorsqu'à ces

données s'ajoutent les références de prix que les particuliers recueillent auprès de leur entourage et des agents immobiliers, encore centraux dans le paysage informationnel.

D'un point de vue technique, l'éventail parfois très large de valeurs collectées sur la toile même s'explique par l'hétérogénéité des dispositifs méthodologiques qui sous-tendent leur production. Mais du point de vue des particuliers, ces divergences sont difficilement compréhensibles, d'autant que les méthodes sont peu transparentes. L'« opacité méthodologique » (Boulay, Blanke, Casanova, Granié 2020) limite la capacité des particuliers à comparer les valeurs qu'ils compilent dans leur recherche de références de prix. Elle renvoie à la concurrence qui se joue sur le marché même de la donnée. Sur ce nouveau marché on l'a vu, les données (et les méthodes) servent autant à informer qu'à capter des flux d'utilisateurs et d'informations les concernant afin de générer des prospects qualifiés et d'en faire commerce aux professionnels de l'intermédiation immobilière.

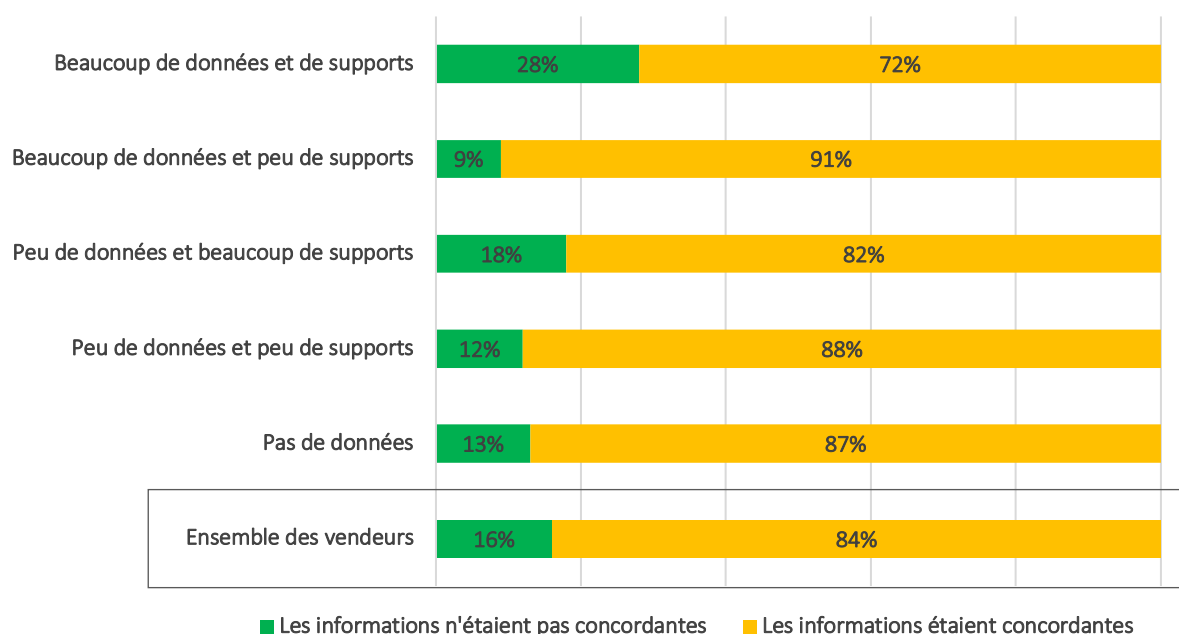


Figure 6. Consommation de données et sentiment de concordance de l'information sur les prix

Vendeurs de logements ayant recours à l'information (n=665, question non-obligatoire)

Pourcentages en ligne - Statistique du Khi-deux : DDL=4 / Valeur = 13,7 / Probabilité = 0,008

Question : Les informations sur les prix étaient-elles concordantes entre les différentes sources ?

Des prix pas forcément modérés

Près de la moitié des plus grands consommateurs de données considèrent avoir positionné leur prix de mise en vente dans le milieu de la fourchette et plus d'un tiers dans la fourchette haute. Cette répartition est spécifique aux vendeurs qui consultent des données en quantité importante pendant leur parcours. Parmi l'ensemble des vendeurs, une plus forte proportion vise le milieu de la fourchette et une plus faible proportion la fourchette haute (Figure 7).

Le positionnement du prix de mise en vente varie donc avec la consommation de données. La tendance supérieure à viser les prix hauts des vendeurs qui consomment beaucoup de données peut être rapprochée d'au moins deux raisons. D'abord, ces vendeurs, on l'a vu, sont plus souvent localisés en zones immobilières tendues, zones qui affichent de fortes augmentations de prix et qui sont propices à des tentatives d'enchère. Ensuite, on l'a vu aussi, ils ont une plus forte propension à entretenir un sentiment de discordance informationnelle. Au fond, les tensions sur les marchés ne seraient pas seules en cause dans les tendances haussières de l'immobilier. Se pencher sur les usages des données indique que le flou qui entourent les marchés pourtant les plus scrutés peut constituer un facteur supplémentaire de croissance des prix.

Ce n'est pas tout : non seulement les plus gros consommateurs de données surfent sur des prix hauts mais ils sont surreprésentés parmi les vendeurs qui ont vendu à un prix final proche de celui qu'ils avaient proposé au départ et dans des délais relativement plus courts. Ce sont donc aussi les plus satisfaits de la transaction conclue (Figure 8). Finalement, la consommation de données semble davantage augmenter la performance économique des vendeurs que minimiser la hausse des prix.

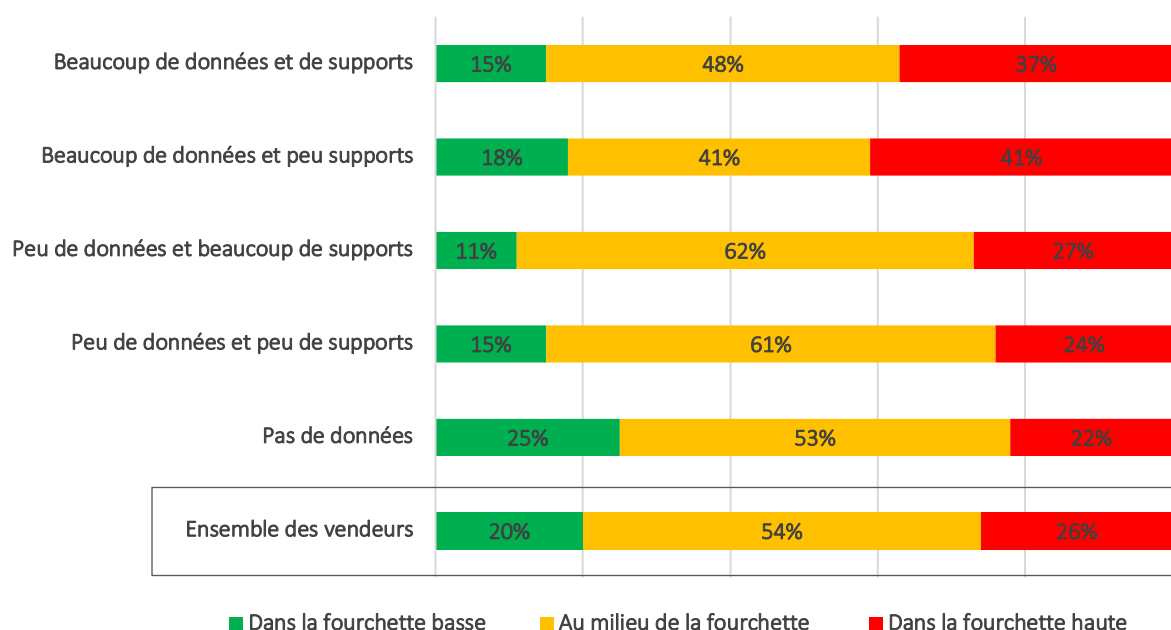


Figure 7. Consommation de données et positionnement du prix de vente (opinion des vendeurs)

Ensemble des vendeurs de logements (n=989, question non-obligatoire)

Pourcentages en ligne - Statistique du Khi-deux : DDL=8 / Valeur = 27,6 / Probabilité = 0,0006

Question : Par rapport aux informations dont vous disposiez, diriez-vous que le prix auquel vous avez mis en vente votre logement était ...

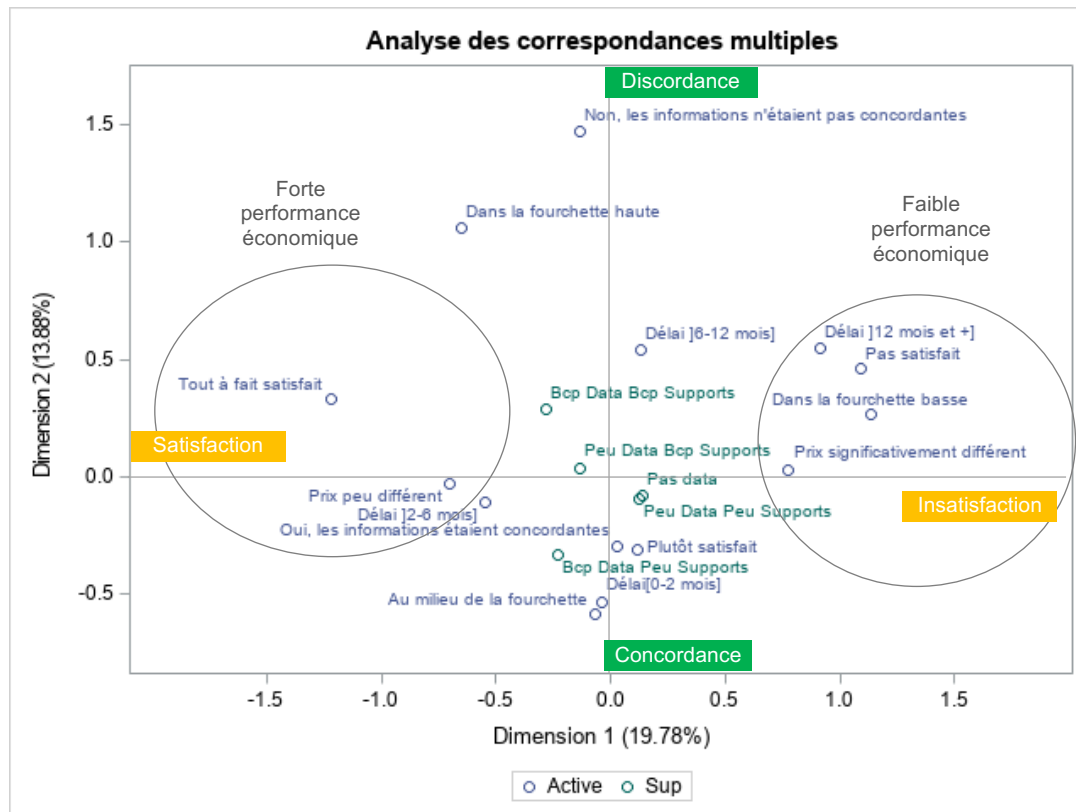


Figure 8. Liens entre consommation de données et pratiques marchandes des vendeurs (analyse factorielle)

Ensemble des vendeurs de logements (n=993)

En somme, l'étude des pratiques concrètes des ménages vendeurs de logements montre que la mise à disposition d'une donnée ne se traduit pas nécessairement par plus de transparence des marchés et par une limitation des tendances haussières sur les prix. Déjà, il ne suffit pas que les données soient accessibles à tous pour qu'elles soient utilisées par tous. La consommation de données est certes une pratique répandue mais elle n'est pas généralisée et s'avère socialement et géographiquement située. Deuxièmement, la donnée a un effet conditionné sur la transparence des marchés. Elle procure un sentiment de transparence aux vendeurs qui en consomment avec modération mais floute la vision de ceux qui en consomment beaucoup. Troisièmement, elle favorise les paris gagnants mais n'a pas forcément un effet modérateur sur les prix et elle peut même favoriser un alignement sur les valeurs hautes du marché.

L'ensemble de ces constats pointe toute l'ambiguïté des effets en termes de régulation des prix d'une politique d'ouverture des données menée à travers la constitution d'un marché de services numériques. Entre la théorie et la réalité s'intercalent les pratiques d'acteurs qui, associant l'enjeu de performance économique à celui de l'intérêt général, s'avèrent plus ou moins aguerris au sein de marchés justement floutés par le jeu de la concurrence par l'information. Alors que la théorie décrit une dynamique qui tend à l'équilibre, le marché immobilier reste dominé par le déséquilibre. Il faut dire que le marché et les usages de la donnée se concentrent dans des marchés locaux haussiers, voire spéculatifs depuis plus de vingt ans. Sur ces marchés, les prix sont interprétés dans la perspective de paris qui

entretiennent la tendance à la hausse. Une question serait de savoir ce qu'il en serait dans le contexte d'une baisse des prix, voire de stagnation tel que les dernières tendances de certaines métropoles s'y confrontent actuellement.

3 Consolidation du marché de la donnée et maintien des acteurs et pratiques historiques

Si environ quatre vendeurs sur dix ont recours à des données de prix pour estimer leur logement, huit sur dix au total utilisent des informations sur les prix, quel que soit leur type et leur canal d'accès. La consommation spécifique de données par les vendeurs de logements s'inscrit donc dans des pratiques de documentation plus larges. Ces pratiques ont ceci de particulier qu'elles maintiennent les professionnels de l'immobilier au centre de l'accès à l'information. Elles révèlent et confortent à la fois l'imbrication et la complémentarité des acteurs historiques du marché immobilier et des starts-ups de la PropTech. On l'a vu, le marché immobilier se digitalise et campe sur une plateformes favorable à la complémentarité des positions. Comme lui, les pratiques de documentation des vendeurs s'hybrident, à l'intersection du canal digital et du canal physique⁶.

Des pratiques de documentation qui dépassent les bornes du marché de la donnée

85% des vendeurs de logements ont recours à des informations sur les prix des logements pour fixer leur propre prix d'offre. Parmi les types d'informations disponibles, leur préférence va, dans des proportions proches et tout canal confondu, pour les estimations, les prix des biens vendus et les prix des biens en vente indiqués dans les annonces (Figure 9). Ces trois catégories sont certes issues de lignes de production différentes (Juillard 2019a) mais elles ont en commun de correspondre à des prix individuels et de renvoyer ainsi à des biens spécifiques.

L'appétence des vendeurs pour les valeurs individuelles traduit d'abord l'importance de l'estimation dans la perspective de vendre. Elle révèle ensuite le principal exercice auquel se prêtent les vendeurs pour fixer le prix de leur logement : la comparaison. De fait, celle-là est établie à partir de deux types de valeurs étalons possibles : les valeurs déjà validées par le marché, à l'instar des prix des biens vendus, une norme en matière de statistiques immobilières et/ou les valeurs proposées à la mise en vente, moins légitimes auprès non seulement des statisticiens mais des professionnels. Entre les deux, le décalage est au moins double et difficile à dépasser par l'interprétation : il est temporel et tient à la fois à la marge de négociation qui s'intercale entre prix d'offre et prix conclu. Le biais d'interprétation qui en découle est d'autant plus fort en période d'évolutions rapides, qu'elles soient haussières comme dans les marchés tendus depuis vingt ans, ou baissières comme dans certains

⁶ Comme précédemment, cette partie emprunte aux résultats de l'enquête par questionnaires présentée en introduction.

marchés plus sensibles aux effets de crise. Pour autant, une pratique courante chez les vendeurs consiste à les combiner dans leur appréciation des prix des logements.

Quant aux estimations, elles ont bien vocation à tenir compte de la spécificité des biens (certaines se font même sur des modèles hédoniques) et des tendances actualisées du marché mais leur « opacité méthodologique » (Boulay, Blanke, Casanova, Granié 2020) est telle qu'il est difficile de les comparer avec toute autre catégorie de références de prix. Comparer les estimations entre elles n'est pas mince affaire non plus. Celles-ci peuvent aller jusqu'à former une fourchette de prix qui va du simple au double pour un même bien, laissant les plus gros consommateurs d'informations, faute de clés d'interprétation on l'a vu, face à un sentiment de discordance informationnelle.

	Vendeurs ayant recours aux types d'informations
Estimations	36%
Prix des biens vendus	35%
Prix des biens en vente (annonces)	32%
Prix au m ² (moyens/médians)	18%

Figure 9. Usage des types d'informations

Ensemble des vendeurs de logements (n=993)

Remarque : la somme est supérieure à 100% car les vendeurs peuvent consulter plusieurs types d'informations à la fois.

Lecture : 32% des vendeurs s'informent sur les valeurs de marché (au moins en partie) en consultant le prix des biens en vente (annonces).

Le maintien des professionnels en position centrale dans l'accès à l'information

Le recours à Internet apparaît majeur dans les pratiques de documentation des vendeurs sur le prix des logements. Cependant, le canal digital n'est mobilisé seul que dans 5% des cas et plus souvent combiné à d'autres canaux, à commencer par celui des professionnels, dans 16% des cas (Figure 10). Dans les moments de recherche d'informations, le canal digital ne remplace donc pas le canal physique et s'offre plutôt en complément. On l'a vu, la complémentarité du digital et du physique est constitutionnelle du modèle de plateforme. Les professionnels ne se bornent d'ailleurs pas à utiliser les sites et applications développées par les acteurs de la PropTech pour publier leurs annonces sur la toile et attirer des prospects à eux. Ils se tournent aussi vers les portails dédiés créés par leurs propres réseaux et instances de représentation. De manière plus générale, leurs pratiques et leurs outils mêmes campent

de plus en plus souvent sur la combinaison des registres physique et digital : le « phygital » (Gardes 2019).

La complémentarité entre le digital et le physique ne s'illustre pas seulement dans le champ de l'information mais aussi sur le front de la transaction. La stabilité de la part de marché de l'intermédiation professionnelle depuis que l'immobilier a engagé son tournant numérique en constitue le principal signal (Vorms et Cusset 2016). Sur ce front, la bataille la plus déterminante à ce stade de la digitalisation du marché se joue plutôt sur d'autres terrains. Elle oppose classiquement l'intermédiation et la désintermédiation et de manière plus récente et plus forte encore, les agents immobiliers et les mandataires parmi les intermédiaires.

Autrement dit, la démultiplication et la démocratisation de l'information immobilière ont beau éroder leur rente informationnelle (Tucillo 1997), les professionnels continuent de jouer un rôle central dans l'accès à l'information. Ceux-là captent 65% des vendeurs en quête de références de prix, un record parmi les canaux d'accès à l'information (Figure 10). Ils constituent leur première source de confiance, loin devant les acteurs du numérique, l'Etat (on le retrouve ici parmi les sources des vendeurs) et l'entourage, au coude à coude dans la hiérarchie (Figure 11).

Aujourd'hui encore, il semblerait que la valeur fonctionnelle des intermédiaires immobiliers continue de s'apprécier en regard de leur avantage informationnel (S. Levitt et C. Syverson 2008) et la première attente des vendeurs à l'encontre des agents immobiliers en particulier reste la détermination du prix de leur logement, devant la recherche d'acquéreurs et la négociation, le soutien logistique et juridique (Larceneux, Lefebvre, Simon 2014). La connaissance par la data reste loin de supplanter l'expertise traditionnelle.

La relation de proximité des professionnels (aux vendeurs et/ou au marché local) constitue aussi un avantage concurrentiel. Elle explique à son tour la permanence du recours à l'entourage dans les pratiques de documentation des vendeurs. Près de 40% des vendeurs se tournent vers leurs proches et voisins pour approcher les prix des biens dans leur secteur (Figure 12). L'ensemble décrit des pratiques composites, entre canal digital et physique, relations formelles et informelles, informations structurées (sous forme de data par exemple) et spontanées. Le marché du logement et de son information n'est apparemment pas près d'être reconfiguré. Il s'hybride mais reste assis sur ses fondamentaux.

Vendeurs ayant recours aux canaux d'informations		
Canal physique	Professionnels	65%
	Entourage	39%
Canal digital	Internet	52%
Canal papier	Presse	20%

Figure 10. Usage des canaux d'accès à l'information sur les prix des logements

Ensemble des vendeurs de logements (n=993)

Remarque : La somme est supérieure à 100% car les vendeurs peuvent consulter plusieurs canaux à la fois.

Lecture : 65% des vendeurs s'informent sur les prix des logements (au moins en partie) via des professionnels (agents immobiliers, mandataires, notaires).

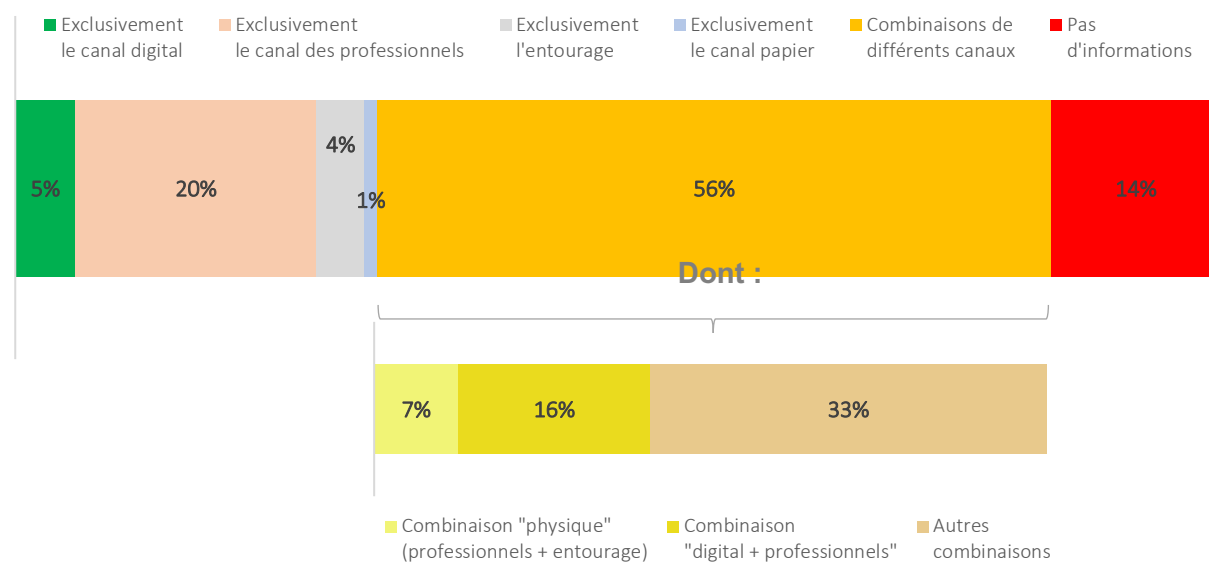


Figure 11. Usages exclusifs et combinaisons des canaux d'accès à l'information

Ensemble des vendeurs de logements (n=993)

Lecture : 20% des vendeurs mobilisent uniquement des professionnels pour s'informer sur les valeurs de marché

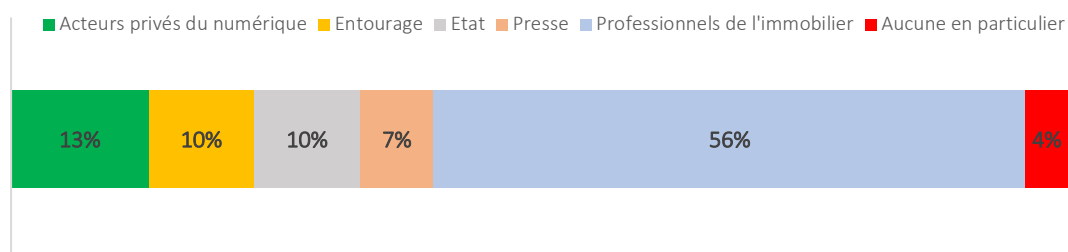


Figure 12. Source d'information la plus influente citées par les vendeurs pendant leur parcours de vente

Vendeurs de logements ayant recours à l'information (n=653, question non-obligatoire)

Question : Parmi les différentes sources d'informations, laquelle vous a le plus influencée ?

Lecture : 10% des vendeurs informés considèrent l'entourage comme la source d'informations la plus décisive.

La recherche d'informations sur les prix, un parcours extensif

Les pratiques de documentation des vendeurs rappellent leur opportunisme (Larceneux, Lefebvre, Simon 2014) et le caractère extensif de parcours de vente tendus vers le meilleur compromis possible (Bonneval 2012). La concurrence qui sévit sur le marché de l'information immobilière n'oppose pas seulement les différentes catégories de fournisseurs entre elles. Le digital et le physique, on vient de le voir, peuvent même tendre à se combiner et se compléter. Les plateformes productrices de données le revendiquent même explicitement et renvoient leurs utilisateurs aux agences physiques pour l'établissement du prix d'offre des biens à vendre.

La concurrence sur le marché de l'information immobilière se joue également au sein de chaque catégorie de fournisseurs. Son terrain privilégié est le canal digital : 37% des vendeurs consultent à la fois plusieurs sites et application pour apprécier les prix des logements (Figure 13). La concurrence est également vive sur le canal physique : dans 23% des cas, les vendeurs mettent en compétition les professionnels sur la question des prix, du moins, ils mobilisent les différentes informations sur les prix qu'ils ont recueillies auprès de plusieurs d'entre eux pour converger vers une valeur par comparaison.

Les canaux d'accès à l'information recoupent en partie les canaux d'accès au marché et les vendeurs dispersent leurs recherches d'informations comme ils dispersent leurs recherches d'acquéreurs potentiels : 47% des vendeurs qui ont recours à l'intermédiation sollicitent plusieurs professionnels pour mettre en vente leur logement (Figure 14). Sur le terrain de la désintermédiation, la tendance s'accroît encore, avec 57% des vendeurs qui décuplent le potentiel du canal de mise vente entre particuliers en multipliant les points de contact à partir de supports dédiés. Ceux-là sont d'ailleurs les plus gros consommateurs d'informations sur les prix sur Internet et assument leur choix de désintermédiation en s'adressant moins souvent que les autres aux professionnels pour se tenir informés sur les prix sur le marché.

		Vendeurs ayant recours à plusieurs supports selon les canaux d'informations
Canal physique	Professionnels	23%
	Entourage	16%
Canal digital	Internet	37%
Canal papier	Presse	6%

Figure 13. Propension des vendeurs à multiplier les supports selon le canal d'informations

Ensemble des vendeurs de logements (n=993)

Remarque : La somme est différente de 100% car les vendeurs peuvent consulter plusieurs supports et plusieurs canaux à la fois.

Lecture : 23% des vendeurs s'informent en consultant plusieurs professionnels à la fois (agents immobiliers, mandataires, notaires).

		Vendeurs ayant recours à plusieurs points de contact par canal de mise en vente
Professionnels	Agents immobiliers, mandataires, notaires	47%
Entre particuliers	Supports dédiés (sites Internet, journaux)	57%
	Bouche-à-oreille	39%

Figure 14. Propension des vendeurs à multiplier les points de contact selon le canal de mise en vente

Ensemble des vendeurs de logements (n=993)

Remarque : La somme est différente de 100% car les vendeurs peuvent mobiliser plusieurs canaux et plusieurs points de contact à la fois.

Lecture : 47% des vendeurs mettent en vente leur logement en mobilisant plusieurs professionnels à la fois (agents immobiliers, mandataires, notaires).

En mobilisant à leur guise une ou plusieurs catégories d'informations sur un ou plusieurs canaux et via un ou plusieurs fournisseurs, les vendeurs de logements affichent une grande hétérogénéité et une forme d'opportunisme dans leurs pratiques de documentation sur les marchés du logement. Cette tendance profite de la diversification et de l'extension du paysage informationnel. Elle consacre le canal digital mais continue d'accorder la primeur au canal physique, à commencer par le professionnel.

4 Conclusion

Les données immobilières et notamment leur progressive diffusion en France forment un cas d'étude assez emblématique des transformations de l'action publique et des formes de régulation associées aux marchés. L'ouverture des données immobilières de l'Etat amorcée dans les années 2010 avec DVF, et encore à l'œuvre aujourd'hui avec l'ouverture récente des données MAJIC Personnes Morales, contribuent à consolider et à légitimer un marché de la donnée. Centré sur le modèle de plateforme, ce marché regroupe des entreprises qui proposent des outils de connaissance aux particuliers acheteurs et vendeurs de logements et produisent de la valeur en monnayant la mise en relation de ces particuliers avec d'autres acteurs économiques, à commencer par les entreprises de la transaction immobilière.

Les politiques d'Open Data visent un objectif d'intérêt général, celui de la transparence des marchés du logement et de la limitation des tendances inflationnistes sur les prix. Elles reposent sur la croyance en l'efficacité des marchés telle que la théorie économique en définit les conditions. La réalité des dispositifs marchands et des pratiques des particuliers qu'elles accompagnent rend leurs effets moins univoques que la théorie ne le prévoit. Gouverner par le marché n'est pas simple.

Empiriquement, l'analyse des pratiques des ménages montrent en effet que la donnée apporte davantage de transparence dans certains cas seulement. Les ménages qui consomment des données ont un profil sociologique et géographique spécifique. De plus, lorsqu'ils consomment trop de données, un sentiment de flou remplace le sentiment de transparence. Par ailleurs, l'effet de la transparence des marchés sur la formation des prix n'est pas si univoque que ce que les discours officiels (la théorie) ne le laissent entendre. Les plus grands consommateurs de données sont davantage soucieux de renforcer leur performance économique que d'assagir les prix de marché. Constatons enfin que ces nouvelles pratiques de documentation sur les marchés s'articulent au maintien des pratiques historiques dans lesquelles les agents immobiliers notamment gardent un rôle central. L'hybridation des pratiques des vendeurs en matière d'information immobilière conforte les fondamentaux du marché du logement.

Ainsi, malgré la diffusion massive de données, un enjeu commun à l'ensemble des acteurs demeure : accompagner les particuliers dans la compréhension du marché. Pour l'heure, le marché de la donnée n'a pas de régulateur. La puissance publique se limite à l'alimenter de données légitimes particulièrement structurantes sans suivi de leurs usages ni pratiques d'étiquetage, tel que Bernard Vorms, Alain Jacquot et Jean-Louis Lhéritier l'appelaient de leurs vœux dans un rapport au gouvernement en 2010 (Vorms, Jacquot, Lhéritier 2010). Construire un marché (de la donnée) pour en réguler un autre (celui du logement) pose pourtant une question de régulation. Cela renvoie aussi à un enjeu révélé ici mais déjà identifié il y a plus de dix ans déjà dans le contexte nord-américain, à l'heure où le marché de la donnée n'était pas encore vraiment constitué en France : un enjeu non pas d'« opportunité » (d'accès à la donnée) mais d'« utilisation effective » (« effective use ») (Gurstein 2010) ainsi que d'équipement et de compétence numériques plus partagés.

Références citées

- Ansaloni M., Trompette P., Zalio P.P. (2017), « Le marché comme forme de régulation politique », *Revue française de sociologie*, 2017/3, vol. 58
- Bergeron H, Castel P, et Dubuisson-Quellier S (2014), « Gouverner par les labels. Une comparaison des politiques de l'obésité et de la consommation durable », *Gouvernement et action publique*, n°3, p. 7-31.
- Boeing G. (2019), « Online rental housing market representation and the digital reproduction of urban inequality », *Environment and Planning A: Economy and space*
- Bonneval L. (2012), « Évolutions du métier d'agent immobilier et marchés du logement », in Bergé M., Rougé L. (dir.), *Être logés, se loger, habiter*, Paris, L'Harmattan, coll. « Habitat et Société »
- Boulay G., Blanke D., Casanova Enault L., Granié A. (2020), « Moving from Market Opacity to Methodological Opacity: Are Web Data Good Enough for French Property Market Monitoring? », *The Professional Geographer*
- Brailly J., Coulondre A. (2016), « Le travail de construction d'une place de marché : centres commerciaux et salons de professionnels », *Revue Française de Socio-Économie*, 2016/1 (n° 16)
- Callon M. (2017), *L'emprise des marchés*, Paris, La Découverte, 501 p.
- Muniesa F et Callon M (2009), « La performativité des sciences économiques », dans Steiner P et Vatin F, *Traité de sociologie économique*, Presses Universitaires de France, p.289-324
- Cavaillès J. (2005), « Le prix des attributs du logement », *Economie et statistique*, n°381-382
- Coulondre A., Juillard C. (2021), Ce que la data fait au marché. Données immobilières, transparence des marchés et fixation des prix des logements en France, étude publiée avec le soutien de Crédit Logement, Groupe BPCE, LIFTI, PUCA
- Cusin F. (2012), « La ville à l'épreuve du marché. Contribution à une sociologie économique de la ville et des mobilités résidentielles », HDR en Sociologie, Université Paris Dauphine, 401 p.
- Desrozières A. (2011), « Les origines statisticiennes de l'économie des conventions : réflexivité et expertise », *Oeconomia*, 1-2
- Dubuisson-Quellier S. (2016), « Introduction. Le Gouvernement des conduites comme modalité d'intervention de l'Etat sur les marchés », *Gouverner les conduites*, Paris, Presse de Sciences Po, coll. « Académique »
- François P. (2008), *Sociologie des marchés*, Armand Colin, 331 p.
- Gardes N. (2019), « Digitalisation du secteur immobilier : la proposition de valeur phytale au cœur de la performance ». *La Revue des Sciences de Gestion*, 299-300 (5)
- Garmaise M.J. et Moskowitz T.J. (2004), « Confronting Information Asymmetries: Evidence from Real Estate Markets », *The Review of Financial Studies*, Vol.17, n°2
- Gurstein M. (2011), « Open data: Empowering the empowered or effective data use for everyone ? », *First Monday*, vol. 16, n. 2
- Juillard C. (2019a), *Le tournant numérique des données immobilières : permanences et recompositions*, Etude publiée par Iread avec le soutien d'Iread, du LIFTI, du PUCA et d'Urbanics
- Juillard C. (2020), « Produire des données de prix et de loyers à l'heure de la PropTech : quel rôle pour l'Etat ? », *Working paper pour la Cities and Digital Technology Chair*, Sciences Po, Paris, janvier 2020
- Kitchin R., Lauriault T. (2014). *Towards critical data studies: Charting and unpacking data assemblages and their work*. Lincoln: University of Nebraska Press
- Larceneux F., Lefebvre T., Simon A. (2014), « La perspective des coûts de transaction perçus : une explication de l'intermédiation immobilière », *Revue d'économie régionale et urbaine*, 2014/3, octobre
- Leglaye S., Rolland A. (2019), « Une personne sur six n'utilise pas Internet, plus d'un usager sur trois manque de compétences numériques de base », Insee Première, n°1780
- Levitt S. et Syverson C. (2008), « Market distortions when agents are better informed : The value of information in real estate transactions », *Review of Economics and Statistics*, Vol. 90, n°4
- Migozzi J. (2020), *Une ville à vendre : numérisation et financiarisation du marché du logement au Cap: stratification et ségrégation de la métropole émergente*, Thèse de doctorat en géographie, Université Grenoble Alpes
- Pumain D., Thisse J-F., Walliser B., Thomas I., Chamussy H., Taillard C., Durand-Dastès F., Marie-Claire Robic M-C., et Saint-Julien T. (2007), « Nouvelle économie géographique et géographie : quel dialogue ? », *L'Espace Géographique*, Vol. 3, n° 3, p.193-214
- Rochet J.-C., Tirole J. (2004), « Two-sided markets: an overview », Institut d'économie industrielle, document de travail, https://frbatlanta.org/filelegacydocs/ep_rochetover.pdf

Sheppard S. (1999), « Hedonic Analysis of Housing Markets », in Mills E.S. and Cheshire P. (dir.), *Handbook of Regional and Urban Economics*, vol. 3, *Applied Urban Economics*.

Simon A. (2008), « Les déterminants de la valeur des appartements : une approche hédoniste », *Etudes Foncières*, Compagnie d'édition foncière

Tuccillo J.-A. (1997), « Technology and the housing markets », *Business Economics*, 32

Vatin F. (dir). (2009), *Evaluer et valoriser. Une sociologie économique de la mesure*, Toulouse, Presses universitaires du Mirail

Vorms B., Jacquot A., Lhéritier J.-L. (2010), *L'information statistique sur le logement et la construction. Rapport d'un groupe de travail au CNIS*, Paris, CNIS

Vorms B., avec Cusset P-Y (2016), *La Révolution numérique et le marché du logement. Nouveaux usages, nouveaux acteurs, nouveaux enjeux*, Paris, France Stratégie