

Modèle économétrique sur le choix de véhicules des ménages

ÉCONOMIE ET ÉVALUATION



Ministère
de l'Écologie,
du Développement
durable,
des Transports
et du Logement

**Collection « Études et documents »
du Service de l'Économie, de l'Évaluation et de l'Intégration du Développement Durable
(SEEIDD) du Commissariat Général au Développement Durable (CGDD)**

Titre du document : « Modèle économétrique sur le choix de véhicules des ménages »

Directeur de la publication : Françoise Maurel

Auteurs : Equipe projet NERA Economic Consulting : Paul Metcalfe, Kenneth Train, Helen Webb, Robin Brejnholt

Contact : Olivier Tessier, olivier.teissier@developpement-durable.gouv.fr

Date de publication : Janvier 2011

Ce document n'engage que son ses auteurs et non les institutions auxquelles ils appartiennent.
L'objet de cette diffusion est de stimuler le débat et d'appeler des commentaires et des critiques.

SOMMAIRE

1. Résumé.....	3
2. Introduction.....	5
3. Revue de Littérature.....	5
3.1. Remarques introductives.....	5
3.2. Présentation des études.....	9
3.3. Récapitulatif et enseignements pour cette étude	19
4. Méthodologie.....	21
4.1. Spécification du Modèle	23
4.2. Agrégation de la Demande de Véhicules, Prévision et Simulation	23
5. Données.....	25
5.1. Exigences Relatives aux Données et Disponibilité	25
5.1.1. Données Relatives aux Ménages.....	25
5.1.2. Données relatives au véhicule.....	27
5.1.3. Données relatives au marché	28
5.2. Traitement des données.....	28
5.2.1. Appréciation Globale de la Qualité des Données.....	35
6. Résultats Descriptifs.....	37
7. Résultats Econométriques.....	45
7.1. Sous-Modèle de Choix de Véhicules pour les Ménages à Un Véhicule	45
7.2. Sous Modèle de Choix de Véhicules pour les Ménages à Deux Véhicules.....	48
7.3. Modèle de Choix du Nombre de Véhicule	50
7.4. Elasticités.....	52
8. Résultats des Simulations.....	56
8.1. Données à renseigner	56
8.2. Résultats des Scenarii	59
9. Conclusions.....	64
9.1. Résumé des Principales Conclusions	64
9.2. Discussion	65
9.3. Recommandations pour des Recherches Futures.....	66
9.4. Ajout de la demande de Carburant des Véhicules comme Sous-Modèle	66
9.5. Ajout de la Détention de Trois Véhicules Comme une Option de Choix.....	67

<i>Références.....</i>	<i>68</i>
<i>ANNEXES.....</i>	<i>69</i>
<i>Annexe A Spécification du Modèle Technique.....</i>	<i>69</i>
<i>Annexe B Variables Pertinentes du Série de données de l'INSEE.....</i>	<i>72</i>
<i>Annexe C Caractéristiques de l'échantillon de l'INSEE.....</i>	<i>74</i>
<i>Annexe D Variables du série de données JATO.....</i>	<i>78</i>
<i>Liste des tableaux</i>	<i>80</i>
<i>Liste des graphiques</i>	<i>81</i>

1. Résumé

Cette étude réalisée par la société NERA pour le compte du CGDD/MEDDTL vise à développer un modèle permettant de simuler les effets de différentes politiques telles que le « bonus écologique » octroyé aux véhicules à faibles émissions de CO₂, sur le choix de véhicule des ménages et sur la demande globale de véhicules.

Il s'agit d'un modèle de choix discret (les ménages ont le choix entre un nombre fini de classes de véhicules), de type « logit hiérarchisé ». Il permet de représenter les différentes phases de choix de possession (et pas seulement d'achat) de véhicules par les ménages : (i) le nombre de véhicules à posséder, et (ii) la catégorie, le millésime, le type de carburant et les émissions de CO₂ de chaque véhicule possédé. Le modèle a été calibré principalement sur la base de données de l'INSEE « Budget des Familles 2006 » (BdF 2006), ainsi que sur la base JATO pour les caractéristiques techniques des véhicules et l'Argus pour les prix des véhicules d'occasion.

Les variables supposées influencer chaque décision apparaissent avec des coefficients de signes anticipés, statistiquement significatifs, et d'un ordre de grandeur plausible. Ainsi, il ressort que les ménages sont sensibles aux prix d'achat et aux coûts variables des véhicules, et d'autant plus sensibles que leurs revenus sont faibles ; ils sont également d'autant plus sensibles à la taille du véhicule que la taille du ménage est grande. Concernant le nombre de véhicules possédés, une hausse des revenus augmente la probabilité qu'un ménage choisisse deux véhicules plutôt qu'un, et un véhicule plutôt qu'aucun ; déménager d'une zone rurale à une ville diminue la probabilité de détenir deux véhicules plutôt qu'un, et de détenir un véhicule plutôt qu'aucun ; enfin, plus la taille de la zone urbaine est importante, plus la probabilité de détenir un véhicule est faible.

Les élasticités au prix (directes et croisées) ont été évaluées par catégorie de véhicule, millésime et tranches d'émissions de CO₂. Ainsi, les élasticités-prix augmentent, en valeur absolue, pour les catégories de A (économique, élasticité de -1,15) à E (luxe, -4,07), avec une élasticité particulièrement élevée pour les voitures de sports (-7,67). Les élasticités-prix propres de la demande diminuent avec l'âge du véhicule ; la demande de nouveaux véhicules étant fortement élastique au prix (de l'ordre de -5,32) alors que celle des véhicules de plus de 5 ans est assez inélastique (-0,42). Enfin, les élasticités-prix propres ont tendance à augmenter avec la quantité de CO₂ émise, de l'ordre de -4 pour les véhicules faiblement émetteurs, elles dépassent -20 pour les véhicules émettant 201 g/km de CO₂ ou plus.

Une fois calibré, le modèle peut être utilisé en projection, notamment pour évaluer l'impact d'une politique incitative agissant sur les prix des véhicules neufs. Les données nécessaires sont les projections des caractéristiques de chaque véhicule (notamment de prix et de performances) et les projections des variables socio-économiques telles que la distribution de la population, des revenus et de la taille des ménages. La principale limite du modèle est qu'il n'intègre pas de façon endogène l'évolution des performances des véhicules proposés aux ménages, autrement dit l'offre des constructeurs automobiles dont la sensibilité à une politique de type bonus écologique est réelle.

2. Introduction

Le MEDDTL a mandaté NERA afin de mener une étude économétrique permettant d'examiner les achats de véhicules des ménages. L'objectif principal de cette étude est de développer un modèle permettant de simuler les effets de différentes politiques telles que le « bonus écologique » octroyées aux véhicules à faibles émissions de CO₂, sur le choix de véhicule des ménages et sur la demande globale de véhicules. Ce document est le rapport final sur l'étude. Un outil Excel est communiqué en annexe de ce rapport dans le but de permettre à l'utilisateur de simuler les effets de différentes politiques d'incitation sur l'acquisition de véhicule ainsi qu'un guide d'utilisation pour cet outil.

Le rapport est composé comme suit : La section 2 contient une revue de la littérature. La section 4 explique la méthodologie suivie. La section 5 résume les données utilisées, et présente des statistiques descriptives de l'échantillon des ménages et des types de véhicules utilisés dans l'analyse. La section 6 présente les résultats descriptifs de l'échantillon concernant le choix de véhicules des ménages. La section 7 présente les principaux résultats économétriques. La section 7 présente les résultats de simulations de certaines politiques identifiant l'effet du « bonus écologique » sur les choix d'acquisition de véhicules. La section 8 présente les observations/résultats finales et des recommandations pour des recherches futures.

Les détails techniques de la spécification du modèle sont décrits dans l'Annexe A. Les détails liés aux données, dont la liste des variables pertinentes et des statistiques descriptives sur les variables sélectionnées, sont exposés en Annexes B, C and D.

3. Revue de Littérature

Ce chapitre a pour objet de présenter un aperçu de la très vaste littérature relative à la modélisation de la demande d'automobiles. Cette revue se concentre sur les modèles de choix discrets, ayant pour objet d'analyser les *choix entre différents types de véhicules*. Pour une vue plus globale des travaux de recherche sur le sujet, nous conseillons notamment la lecture de De Jong et al. (2004), qui offre un panorama des approches et techniques de modélisation disponibles.

La littérature relative aux choix d'automobiles a connu un regain d'intérêt au cours des dernières années, suite :

- *aux efforts et politiques de réduction des émissions de gaz à effet de serre*, l'élaboration et la mise en place de ces politiques nécessitant une bonne compréhension des décisions d'achat et de détention de véhicules¹ ; et
- *à l'avancée des méthodes de modélisation des choix, et notamment des techniques de simulation*, permettant de s'affranchir des contraintes imposées par les modèles classiques, « de première génération. »²
- Les récentes études commanditées par le ministère des transports britanniques (Department for Transport), sur lesquelles nous revenons en détail par la suite, s'inscrivent tout à fait dans ce contexte.

Mais avant de présenter les travaux de recherche que nous avons sélectionnés, il convient de préciser quelques points méthodologiques, afférant à la spécification et l'estimation des modèles de choix discrets.³

3.1. Remarques introductives

Les modèles de choix discrets ont pour objet d'expliquer le comportement d'un agent lorsque celui-ci fait un choix – ou une série de choix – entre un ensemble fini d'options (le « choice set »).⁴ C'est une approche causale, par laquelle on cherche à identifier la totalité des facteurs expliquant un choix – ou une série de choix.

¹ Ainsi aux Etats-Unis, l'introduction des réglementations CAFE (« Corporate Average Fuel Economy ») et des systèmes de « feebates » (comparable au bonus-malus écologique français) ont suscité de nombreuses publications.

² L'ouvrage en ligne de Kenneth Train offre une très bonne description de cette mutation ; « Discrete Choice Methods with Simulation, » 2003, <http://elsa.berkeley.edu/books/choice2.html>

³ Cette section est basée sur Train (2003).

Parmi ces facteurs explicatifs, certains seront observés par l'analyste, d'autres ne le seront pas. Dès lors, le choix de l'agent ne pourra être « prédit » de manière exacte ; on cherchera donc à déterminer la *probabilité* d'un choix donné.

Les facteurs non-observés sont traités comme des variables aléatoires. Et la probabilité qu'un agent choisisse une option donnée est simplement la probabilité que ces variables aléatoires prennent une valeur telle que le modèle, la fonction de détermination des choix, produise l'option choisie.⁵ La probabilité d'un choix donné peut alors être exprimée comme l'*intégrale* d'un indicateur de choix sur l'ensemble des valeurs possibles des facteurs non-observés, aléatoires.⁶ Cette intégrale peut être évaluée selon trois approches:

- *estimation en forme fermée* (« closed-form »), où la fonction de détermination des choix et la distribution des termes aléatoires peuvent être exprimées analytiquement, et où la probabilité d'un choix (en fonction d'un ensemble de facteurs observés) peut être calculée de manière exacte, par une formule. Les modèles logistique multinomial, logistique multinomial emboîté et logistique ordonné appartiennent tous à cette catégorie.
- *approximation par simulation complète*, où les probabilités de choix ne peuvent être définies analytiquement, mais doivent être estimées par approximation numérique. Le modèle Probit en est l'exemple le plus connu.
- *approche mixte (forme fermée et simulation)*, où les facteurs non-observés, aléatoires peuvent être répartis en sous groupes et où une solution analytique existe pour une partie de l'intégrale uniquement. Ce principe de décomposition est utilisé notamment dans le modèle logistique mixte (« mixed logit »).

Les modèles de choix discrets sont généralement estimés en faisant l'hypothèse que l'agent maximise son bien être, son utilité. De tels modèles sont appelés « modèles d'utilité stochastique »⁷. Stochastique parce que l'analyste ne peut observer directement l'utilité des agents : il n'a d'information que sur les options offertes aux agents, et sur certaines de leurs caractéristiques.⁸

L'utilité des agents peut alors être exprimée comme la somme d'une utilité dite « représentative » (elle-même fonction des attributs des options et des caractéristiques des agents, et dont l'analyste cherche à estimer les paramètres), et d'éléments non-observés qui impactent l'utilité des agents, mais ne sont pour autant pris en compte dans l'utilité représentative. La distribution des ces éléments non-observés, que l'analyste traite comme aléatoires, peut prêter à plusieurs interprétations.⁹ Elle donne lieu à cinq catégories de modèles :

- *le modèle logistique traditionnel*, dont l'hypothèse principale est que chaque terme aléatoire – pour chacun des agents et alternatives – est indépendamment et identiquement distribué (iid).¹⁰ Ce qui implique notamment que la portion d'utilité non-observée pour une alternative donnée est indépendante de la portion d'utilité non-observée pour une autre alternative ; ou encore que la probabilité relative de choisir une alternative *i* contre une alternative *j*, est indépendante de l'existence et des caractéristiques de toute autre alternative. C'est l'hypothèse dite d'indépendance des alternatives non pertinentes.¹¹

⁴ Les options, ou alternatives, au sein du « choice set » doivent être mutuellement exclusives et en nombre fini. Le « choice set » doit être exhaustif : il doit comprendre toutes les options possibles (à l'exception du « bien externe, » sur lequel nous revenons par la suite).

⁵ En d'autres termes, $P(y | x) = \text{Prob}(\text{tel que } h(x, \epsilon) = y)$; où *y* est l'option choisie, *x* un ensemble de facteurs observés, et un ensemble de facteurs non-observés, aléatoires. *h*(*x*, ϵ) est une fonction quelconque, décrivant la formation des choix (« behavioural process ») ; Train (2003), Chapitre I, page 3

⁶ Ou encore : $P(y | x) = \text{Prob}(I[h(x, \epsilon) = y] = 1) = \int I[h(x, \epsilon) = y] f(\epsilon) d\epsilon$ où *f*(ϵ) est la fonction de densité de la variable ϵ et *I*[.] est un indicateur booléen de choix ; Train (2003), Chapitre I, page 4

⁷ Ou « random utility models, » dans la littérature de langue anglaise

⁸ Caractéristiques des agents

⁹ Train (2003), Chapitre II, page 21

¹⁰ Selon une distribution dite Gumbel ou « extreme value » ; la différence entre deux termes suivant une distribution logistique.

¹¹ Cette hypothèse, et les problèmes qui lui sont associés, ont été illustrés par l'économétricien américain McFadden avec l'exemple du « bus bleu / bus rouge. » Considérant une situation initiale où un voyageur n'a le choix qu'entre deux moyens de transport : la voiture et un bus bleu ; et où ce voyageur choisit entre ces deux options avec la même probabilité, 1/2. Si une

- *le modèle logistique multinomial emboîté* appartient à une classe de modèles développés en grande partie pour contourner les restrictions imposées par le modèle logistique classique.¹² Il consiste à assembler les différentes alternatives en sous-groupes, de sorte que l'hypothèse d'indépendance des alternatives non pertinentes puisse être maintenue entre les alternatives appartenant au même groupe, mais relaxée entre les alternatives appartenant à des groupes différents.
- *le modèle logistique ordonné* s'applique aux situations où les options peuvent être classées¹³ et, dès lors, où l'hypothèse d'indépendance des termes aléatoires ne tient pas (puisque en présence d'alternatives ordonnées, une alternative ressemble davantage à celles dont elle est proche).¹⁴
- *le modèle probit* repose sur l'hypothèse que les facteurs non observés sont distribués conjointement selon la loi normale. Il offre une très grande flexibilité de modélisation des corrélations entre alternatives et, si besoin, périodes. Son estimation, cependant, reste complexe.
- *le modèle logistique mixte* permet aux coefficients de la fonction d'utilité – qui sont fixes dans la spécification traditionnelle – de suivre toute distribution. La distribution des (autres) facteurs non-observés, elle, reste logistique.¹⁵ Ce type de modèle autorise la prise en compte des variations aléatoires de préférence entre consommateurs (« random taste variations »). Il permet également aux facteurs non-observés d'être corrélés entre eux, dans le temps et entre différentes alternatives. Son estimation, reposant sur des techniques de simulation, est difficile ; et nécessite une base de données très fournie.¹⁶

De manière générale, l'estimation des modèles de choix discrets repose sur l'utilisation de données sur les attributs des options (au sein d'un « choice set » déterminé) et sur les caractéristiques des agents.

Ces données peuvent être associées à : i) des choix effectivement réalisés (et reflétés dans des indicateurs de vente, des bases de données commerciales ou enquêtes de consommation) ; ii) des intentions ou déclarations de choix, dans le cadre de procédures expérimentales ; ou iii) une combinaison des deux.

Le niveau d'agrégation des choix utilisés dans l'estimation des paramètres du modèle donne lieu à deux grandes catégories d'approches : ¹⁷

- *l'approche agrégée* considère l'intégralité des décisions d'achat sur une période et dans un périmètre géographique donnés pour déterminer des « parts de marchés » (proportion des ménages ayant choisi chacun des modèles disponibles sur le marché, ou ayant choisi de ne pas acquérir de véhicule). Ces parts de marché sont ensuite mises en relation avec les caractéristiques techniques et tarifaires des véhicules disponibles pour inférer les paramètres de la fonction de choix sous-jacente ; et ce sans qu'il soit nécessaire de recourir à des informations (supplémentaires) sur les agents.¹⁸
- *l'approche désagrégée* repose sur l'analyse directe de décisions individuelles. Elle met en relation un ensemble de choix observés sur une période donnée avec les caractéristiques des individus (ou ménages) qui leur sont associés, et les attributs des options (véhicules) disponibles.

troisième option est introduite sous la forme d'un bus rouge, l'hypothèse d'indépendance des alternatives non pertinentes implique que la probabilité associée à chacun des trois modes devienne 1/3.

12 Modèles appelés « Generalised Extreme Value » (GEV) où les portions d'utilité non-observées suivent une distribution Gumbel, mais peuvent être corrélées entre elles.

13 Par exemple, dans le cadre d'un sondage d'opinion, où les options seraient « mauvais », « bon », « très bon », « excellent. »

14 Le modèle logistique emboîté est parfois appelé, à tort, modèle logistique ordonné. Il s'agit de deux procédures différentes.

15 Distribution conditionnelle à la valeur des coefficients aléatoires.

16 Train, K., «Discrete Choice Methods with Simulation», Chapter 6, 2003, p 138-153

¹⁷ Il est important de souligner que les modèles agrégés et désagrégés ont tous deux pour fondation une fonction d'utilité individuelle; ils ne diffèrent que par la nature des informations utilisées pour estimer les paramètres de cette fonction. Voir Eftec (2008) pour une discussion détaillée sur les deux approches (pages 7 à 10).

18 Cependant, comme le démontrent Berry et al. (1995) ou Berry et al. (2004), les caractéristiques socio-économiques des agents peuvent être prises en compte dans le cadre d'un modèle agrégé.

Alors que le modèle agrégé peut être estimé en utilisant des données de marché, d'accès relativement facile (immatriculations de voitures neuves, volumes de vente, etc.), le modèle désagrégé nécessite l'utilisation de données détaillées sur les consommateurs : données existantes ou, plus fréquemment, données collectées pour les besoins spécifiques de l'estimation.

La littérature aborde également le processus de prise de décision d'achat d'un véhicule automobile au sein d'un ménage.

Il est généralement admis que cette décision est prise, collectivement, par les membres d'un ménage. Ainsi, la plupart des études identifiées utilisent les caractéristiques du *ménage* comme variables explicatives.¹⁹ Cependant, afin de simplifier l'analyse, les modèles théoriques sur lesquels reposent ces études ne spécifient la fonction d'utilité que pour un seul individu. Bien que les préférences de cet individu puissent être considérées comme étant représentatives de l'ensemble du ménage, une telle simplification néglige le processus de prise de décision sous-jacent.

Récemment, l'économiste David Hensher, dans une série d'articles, a proposé une approche alternative, consistant à modéliser explicitement la formation des choix au sein d'un groupe.²⁰ Nous estimons cependant que les possibilités d'utilisation de ces développements théoriques restent à ce jour limitées.²¹

L'un des résultats principaux d'une analyse de choix de véhicule (en particulier dans une optique de simulation de politiques incitatives, telles que le bonus-malus écologique) est la magnitude des coefficients d'élasticité-prix. Les études que nous avons identifiées ne définissent pas ce concept de manière uniforme. Il est donc nécessaire de fournir ici quelques éléments explicatifs :

- *l'élasticité-prix directe* est le rapport entre la variation relative de la demande d'un véhicule et la variation relative du prix de ce véhicule. Bien que cette élasticité puisse être calculée pour chaque véhicule considéré dans l'analyse, elle est fréquemment présentée comme une *moyenne* estimée pour tous les véhicules appartenant à une même catégorie.²²
- *l'élasticité-prix croisée* est le rapport entre le pourcentage de variation de la quantité demandée d'un véhicule donné et le pourcentage de variation du prix d'un autre véhicule. Une élasticité-prix croisée peut être estimée pour toutes les paires de véhicules considérés dans l'analyse, mais pour faciliter la présentation, ces élasticités sont souvent agrégées.²³
- une *semi-élasticité* mesure l'effet (relatif) sur la demande d'un changement de prix *absolu*; cette statistique est retenue notamment par Berry et al. (1995), et Berry et al. (2004).
- enfin, l'élasticité de la demande peut être exprimée en termes de quantité demandée ou en termes de *part de marché*; la plupart des études identifiées retiennent la seconde approche.²⁴

¹⁹ Mohammadian et Miller (2003), que nous présentons en Section 2, inclut une combinaison de données sur les ménages et sur les utilisateurs principaux (conducteurs) des véhicules. Il est également fréquent d'utiliser l'âge du seul chef de famille comme variable explicative.

²⁰ Hensher et al. (2008), par exemple, propose un modèle qui décrit la décision d'achat de véhicule automobile comme une séquence d'étapes dans lesquelles chaque membre d'un ménage, l'un après l'autre, choisit son modèle préféré, puis ajuste son choix au cours du processus. Le modèle est estimé à l'aide de données expérimentales (lors de l'expérience, il est demandé à un couple de suivre chacune des étapes du modèle : choix initial, discussion, ajustements, etc.). Parmi les résultats de l'étude, on notera que l'effet positif de certains attributs du véhicule sur l'utilité des acteurs diminue lorsque ceux-ci ne s'entendent pas lors de la première étape. On notera également que la probabilité d'entente est plus élevée pour un couple marié que pour un couple non-marié.

²¹ Notamment dans le cadre de notre étude, qui ne prévoit pas de collecte de données supplémentaires.

²² Eftec (2008), par exemple, présente une élasticité-prix directe moyenne pour neuf catégories.

²³ Ainsi Eftec (2008) présente i) l'élasticité prix-croisée moyenne entre tous les véhicules d'une même catégorie, et ii) l'élasticité moyenne de la demande de tous les véhicules d'une certaine catégorie par rapport à un changement de prix dans une autre catégorie.

²⁴ C'est le cas notamment d'Eftec (2008), CE et BMRB (2008), Berry et al. (2004), Berry et al. (1995), et Train et Winston (2007).

Pour illustrer les approches, techniques économétriques et données utilisées dans l'estimation de modèles de choix de véhicule automobile, nous avons sélectionné une dizaine d'articles. Ces études présentent les particularités suivantes :

- elles ont été réalisées afin d'aborder des problématiques comparables à celles qui nous intéressent : Eftec (2008), CE et BMRB (2008), Bento et al. (2006), Greene et al. (2005) ;
- elles concernent le marché automobile français : Prieto (2006), Cecchini (2001) ;
- elles utilisent une approche similaire à celle que nous envisageons (modèle logistique multinomial emboîté) : Mohammadian et Miller (2003), McCarthy et Tay (1998), Berkovec et Rust (1984) ; ou
- elles présentent un apport significatif en termes de méthodologie ou résultats : Train et Winston (2007), ou Berry et al. (1995) et (2004).

Nous présentons ci-dessous chacune de ces études en détail. Les études sont introduites par ordre chronologique inversé.²⁵

3.2. Présentation des études

Eftec (2008) a pour objet d'expliquer les décisions d'achat de véhicules *neufs* des ménages britanniques. L'étude cherche en particulier à évaluer l'importance relative d'un ensemble de caractéristiques (gamme, cylindrée, nombre de portières, système de freinage, coût d'achat, coût d'utilisation, etc.) dans les processus de choix.

Trois options ont été considérées à cette fin : la méthode des prix hédoniques,²⁶ un modèle de choix discrets agrégé, et un modèle de choix discrets désagrégé.²⁷ L'approche retenue est un *modèle de choix discrets agrégé* puisque, selon les auteurs, cette option permet : i) une analyse fine des parts de marché; ii) l'identification des comportements de substitution entre véhicules et groupes de véhicules; et iii) la modélisation d'un « bien externe » (défini ici comme « ne pas acheter de voiture neuve ») permettant de tenir compte des gains de bien-être associés à la décision de ne pas acheter de véhicule.²⁸ Enfin, les modèles agrégés ne nécessitant pas de données d'enquête, ils peuvent être entièrement estimés à l'aide de données secondaires, existantes.

Un modèle logistique de *choix emboîtés* a été préféré au modèle logistique traditionnel afin de s'affranchir de l'hypothèse d'indépendance des alternatives non pertinentes et d'affiner l'analyse des comportements de substitution par catégorie de véhicule.

Les informations relatives aux achats de voitures nouvelles proviennent d'une base de données sur les immatriculations (« new car registrations ») alimentée et maintenue par la « Driver and Vehicle Licensing

25 A l'exception de Berry et al. (2004), présenté après Berry et al. (1995).

26 « Hedonic pricing model, » où l'on cherche à déterminer la valeur implicite de certains des attributs d'un produit en modélisant son prix de vente.

27 Comme nous l'expliquons en introduction, un modèle de choix agrégé examine comment les parts de marché varient en fonction des caractéristiques physiques et du prix des véhicules disponibles sur le marché. Il utilise cette information pour estimer les paramètres de la fonction d'utilité représentative. En contraste, dans un modèle désagrégé, les paramètres de la fonction d'utilité sont déterminés de manière à optimiser la prédiction des choix individuels, faits par chacun des ménages (Eftec (2008) pages 6 et 7).

28 Un bien externe (« outside good ») est une option générique, pouvant être choisie par le consommateur au lieu d'une des options du « choice set. » La définition du bien externe varie d'une étude à l'autre, et dépend de la définition du « choice set. » Si l'analyse ne considère que l'achat de véhicules neufs, alors le bien externe peut être défini comme l'absence d'achat de véhicule neuf (c'est ce que font Eftec (2008), Greene et al. (2005), ou Berry et al. (1995)). Si l'analyse considère l'achat de véhicules neufs et d'occasion, alors le bien externe est défini comme l'absence d'achat de véhicule. Le bien externe peut également être défini comme l'absence de véhicule, comme il est fait dans Bento et al. (2006).

Agency. » Cette base de données comprend près de 850 000 observations sur la période 2001 – 2006,²⁹ représentant 2 190 modèles de voiture différents. Ces immatriculations ont ensuite été recoupées avec les informations disponibles dans la base de données JATO,³⁰ comprenant les caractéristiques techniques détaillées de 2 062 des 2 190 modèles identifiés.

À l'issue des analyses menées, l'impact d'une série de facteurs (relatifs aux coûts et autres caractéristiques des véhicules) a pu être mesuré :

- l'effet du prix d'achat sur la probabilité d'achat d'un véhicule est négatif et significatif. D'autres facteurs relatifs aux coûts (comme le prix de revente, les coûts fixes et variables d'utilisation) ont un effet statistiquement significatif. Ces résultats se conforment globalement aux attentes théoriques, même si l'effet estimé varie d'une catégorie de coûts à l'autre³¹ ;
- en général, la valeur absolue de l'élasticité-prix directe moyenne est supérieure à 1,0. L'élasticité-prix est plus élevée (en valeur absolue) pour les petites voitures (comme la Ford Ka) que pour les voitures de sport ou 4X4 (« Sport Utility Vehicles ») ;
- les élasticités-prix directes moyennes sont comprises entre -1,5 (catégorie « Renault Scenic ») et -5,3 (catégorie « Honda Civic ») ;
- une analyse des élasticités-prix croisées entre voitures appartenant à la même catégorie révèle que l'élasticité moyenne est plus élevée pour les petites voitures (catégorie Ford Ka) pour lesquelles une augmentation de 1 pourcent du prix entraîne une augmentation de la demande de 0,105 pourcent pour d'autres automobiles appartenant au même segment ;
- enfin, l'étude s'est attachée à analyser les variations de la demande suite à un changement de 1 pourcent du prix de vente des véhicules, tout en classant ces véhicules en fonction de la quantité de CO2 qu'ils rejettent. Les auteurs observent que l'augmentation du prix des véhicules dans une certaine catégorie entraîne une diminution de la demande globale pour cette catégorie. Par ailleurs, la substitution de voitures (suite à un changement de prix) a tendance à se limiter à des voitures ayant un rejet de CO2 comparable. Enfin, l'importance du bien externe est confirmée, notamment dans la catégorie 111 – 120 grammes de CO2 par kilomètre, puisque plus de 7 pourcent des ménages décidant de ne plus acheter dans cette catégorie (suite à une augmentation des prix) choisissent de ne pas acheter de véhicule du tout.

L'objectif de l'étude **CE et BMRB (2008)** est d'identifier les facteurs « décisifs » dans le choix d'un véhicule, de manière à définir les mesures adéquates permettant de diminuer les émissions de CO2. Pour ce faire, l'accent est mis sur deux éléments : les élasticités-prix propres ou directes, et les élasticités-prix croisées.

L'approche retenue consiste à estimer un modèle de choix discrets (désagrégé) selon une spécification logistique mixte (« mixed logit »), dans laquelle chaque véhicule est considéré comme un ensemble de caractéristiques, et les consommateurs expriment leurs préférences par rapport à ces seules caractéristiques. En outre, chaque consommateur valorise personnellement l'utilité (marginale) que lui apporte chaque caractéristique.

Le modèle exprime l'utilité procurée par un véhicule donné à un ménage donné en fonction de i) l'utilité *moyenne* procurée par ce véhicule à *tous* les ménages, et ii) l'utilité *supplémentaire* procurée par le véhicule donné au ménage donné, compte tenu des caractéristiques du ménage.

Les données relatives aux caractéristiques des ménages proviennent d'un sondage réalisé pour les besoins spécifiques de l'étude.³² Les caractéristiques techniques des véhicules ont été établies grâce au sondage et

²⁹ Chaque observation correspond à un ensemble d'immatriculations (les données reçues de la DVLA étant partiellement agrégées). Au total, entre 2001 et 2006, plus de 15 millions de véhicules automobiles neufs ont été immatriculés au Royaume Uni (sans l'Irlande du Nord).

³⁰ <http://www.jato.com>

³¹ Ce qui reflète les poids implicitement attribués par le consommateur à ces différentes catégories de coûts. L'effet du prix d'achat, par exemple, est plus fort que l'effet du prix de revente.

³² Elles incluent : le revenu du ménage (brut et net), la valeur de son patrimoine, le type de logement, l'âge et le sexe de la personne de référence, le nombre d'enfants et le nombre d'automobiles détenues.

complétées par l'utilisation d'une base de données JATO. Les données sur les prix, coûts fixes et coûts variables ont également été fournies par JATO.³³

À l'issue de l'estimation, il apparaît que les caractéristiques suivantes jouent un rôle important dans la décision d'achat : la consommation de carburant, le type de boîte à vitesses (manuelle ou automatique), le type de carburant (essence ou diesel), le prix d'achat, et la marque du véhicule.

Les auteurs estiment les élasticités-prix directes et croisées pour des changements de prix, des changements de coûts variables et des changements de coûts fixes. Ces résultats sont présentés par catégorie de véhicule et par niveau de rejet de CO₂. Les résultats les plus importants peuvent être résumés ainsi :

- les élasticités-prix moyennes (directes et indirectes) dans une même catégorie sont en général supérieures à l'unité ;
- la demande pour les voitures luxueuses est la plus sensible aux variations de prix (-3,5) ; la demande pour les 4X4 est la moins élastique (-0,96) ;
- les élasticités-prix croisées moyennes (entre catégories) ont tendance à être plus élevées pour des catégories de véhicules similaires. Les consommateurs de voitures luxueuses, supérieures (« executive ») et citadines (« mini ») sont particulièrement loyaux à leur catégorie. Ainsi, la sensibilité de la demande pour d'autres catégories à une variation de prix dans la catégorie luxueuse est inférieure à 0,05 ;
- en groupant les véhicules par niveau de rejet de CO₂, les élasticités-prix moyennes révèlent que l'effet de substitution est le plus fort dans les catégories centrales (rejet de CO₂ compris entre 121 et 165 grammes par kilomètre) ;
- les élasticités-coûts-variables croisées moyennes sont plus élevées (en valeur absolue) pour les véhicules appartenant à des catégories de rejet comparables (121 – 135 g/km, 136 – 150 g/km et 151 – 158 g/km) que pour des voitures plus polluantes ; et
- les effets de substitution moyens entre catégories sont les plus réduits pour les voitures les plus polluantes (rejet supérieur ou égal à 300 g/km). Bien que l'effet de substitution entre catégories soit limité, l'impact négatif sur la demande dans la même catégorie est le plus fort pour les voitures à rejet élevé.

Train et Winston (2007) cherche à comprendre l'érosion des parts de marché des fabricants d'automobiles américains sur les dix dernières années par l'intermédiaire d'un modèle économétrique identifiant les principaux facteurs de choix. Outre les facteurs traditionnels que sont les attributs physiques et le prix des véhicules, l'étude s'attache à estimer l'effet d'autres facteurs, tels que la loyauté à une marque ou la qualité du réseau de distribution.

Un modèle *logistique mixte* est utilisé afin d'associer le choix du consommateur à i) l'utilité *moyenne* qu'offre un modèle particulier (moyenne pour l'ensemble des consommateurs) ; ii) l'utilité supplémentaire qu'offre ce modèle aux ménages présentant les caractéristiques du ménage étudié ; et iii) l'utilité supplémentaire qu'offre ce modèle de manière aléatoire (c'est à dire qu'on ne peut attribuer à aucune des caractéristiques observées).

L'estimation du modèle est rendue difficile par la prise en compte des éléments suivants :

- *hétérogénéité des préférences* : à cette fin, le modèle examine non seulement le véhicule choisi, mais considère aussi l'ensemble des véhicules qui représentaient une option « sérieuse » pour le ménage, au moment de son choix (information recueillie lors du sondage) ;
- *préférence moyenne* : l'utilité moyenne associée à chaque véhicule a été estimée de manière à optimiser la prédiction des parts de marché observées, selon une procédure itérative ; et
- *fidélité à une marque* : les phénomènes de fidélité ont été pris en compte par l'inclusion d'une variable mesurant le nombre d'achats consécutifs, pour six groupes de constructeurs.

³³ Les véhicules sont catégorisés en neuf groupes (« mini », « supermini », « lower-medium », « upper-medium », « executive », « luxury saloon », « sports », « 4x4 » et « Multi Purpose Vehicle ») selon une classification de la « Society of Motor Manufacturers. »

Les données relatives aux ménages et à leur consommation proviennent d'un sondage réalisé spécifiquement pour les besoins de l'étude.³⁴ Les informations relatives aux caractéristiques des véhicules proviennent de publications variées, dont le « Market Data Book » publié par « Automotive News, » et le « Wards' Automotive Yearbook. »

Parmi les résultats d'estimation, on retiendra :

- une élasticité-prix moyenne de - 2,32 ;
 - la fiabilité du véhicule, le rapport nombre de chevaux / poids à vide, l'empattement et la longueur du véhicule concourent tous au choix d'un véhicule. A l'inverse, la consommation d'essence (par mile) est négativement corrélée au choix d'un véhicule ;
 - la qualité du réseau de distribution d'une marque donnée influence positivement, et de manière significative, le choix d'un véhicule de cette marque ; et
 - les consommateurs ont tendance à être plus fidèles aux marques européennes qu'aux autres marques.
- Enfin, pour revenir à l'objet principal de l'étude, selon les auteurs, le déclin des parts de marché des constructeurs américains s'explique *avant tout* par un changement relatif des attributs de base des véhicules (prix, taille, puissance, coût d'utilisation, fiabilité, etc.) ; les constructeurs américains n'ayant pas amélioré leurs produits au même rythme que leurs homologues japonais ou européens.

Bento et al. (2006) cherche à analyser l'impact d'une augmentation de la taxe essence (« gasoline tax ») sur la consommation de carburant *et* le choix de véhicules. Pour ce faire, les auteurs développent un outil de simulation, ancré sur un modèle économétrique de choix discrets-continus. Ce modèle permet d'examiner à la fois les choix de véhicule (choix discret) et les distances parcourues (variable continue), tout en tenant compte des interactions entre les facteurs de choix affectant les deux dimensions.

L'hypothèse sous-jacente est qu'une hausse des taxes sur les carburants pousse les automobilistes vers des voitures à consommation réduite, et entraîne donc un renouvellement du parc automobile. Pour faciliter la prise en compte de ce renouvellement, le modèle utilise des données couvrant non seulement le marché du neuf et de l'occasion mais aussi des voitures abandonnées (« scrapped cars »). L'analyse est réalisée sur dix ans.

Les données utilisées proviennent de deux sources principales :

- un sondage sur les comportements d'achat d'automobiles des ménages américains³⁵ regroupant 26 038 entrées (20 429 après traitement) ; et
- un ensemble d'informations sur le prix et les caractéristiques de voitures neuves et d'occasion issues du « Wards Automotive Yearbook », du « National Automobile Dealer's Association Used Car Guide » et du site internet du ministère de l'énergie (« Department of Energy »).

Parmi les résultats saillants de l'étude, on notera :

- *impact de l'augmentation de la taxe sur la consommation d'essence*: chaque cent (par gallon) d'augmentation du prix donne lieu à une baisse de la consommation d'environ 0,2 pourcent (élasticité-prix);
- *effets sur les revenus fiscaux de l'état*: une augmentation de taxes comprise entre 10 et 30 cents génèrent approximativement 15 cents de revenus fiscaux supplémentaires ; une augmentation de 75 cents, environ 25 cents supplémentaires ; et
- *comportements expliquant la baisse de la demande d'essence*: la baisse s'effectue principalement par une diminution des distances parcourues ; l'impact du passage à un véhicule à consommation plus réduite n'étant que résiduel.

Prieto (2006) analyse la mutation des marchés automobiles européens (prédominance du marché de l'occasion et montée en force des véhicules « low cost ») et cherche à tirer les enseignements de cette mutation quant aux

³⁴ Un échantillon de 458 consommateurs ayant acquis un véhicule neuf (modèle de l'année 2000) a été retenu.

³⁵ National Household Travel Survey (NHTS), U.S. Department of Transportation, 2001

facteurs de motorisation des ménages et aux choix du type d'automobile. En d'autres termes, il s'agit de comprendre l'évolution des comportements d'achat et de détention d'automobiles des ménages.

L'étude comprend deux parties et cinq chapitres. La première partie (chapitres 1 et 2) s'appuie sur une analyse de données portant sur la structure du parc automobile français et la consommation d'automobiles (cette partie comprend notamment une analyse factorielle discriminante, cherchant à « expliquer » la motorisation et multi-motorisation des ménages, sur la base de données d'enquêtes). La seconde partie (chapitres 3 à 5) s'attache à analyser les décisions d'achat, en introduisant les méthodes de choix discrets. La présente synthèse se concentre sur les chapitres 4 et 5.

Dans le chapitre 4, l'auteur propose un tour d'horizon des outils statistiques disponibles pour analyser la demande de produits différenciés. Il explique pourquoi les méthodes de choix discrets sont les mieux adaptées à l'analyse des choix individuels de consommation ; et ce tant pour des raisons empiriques que théoriques.

Le chapitre 5 présente une application des méthodes de choix discrets à l'analyse du choix des consommateurs français pour les principaux « segments » automobiles (inférieur, moyen-inférieur, moyen-supérieur, et supérieur). La particularité de cette analyse tient à ce qu'elle considère le choix des segments en tant qu'*alternatives ordonnées*, et à ce qu'elle inclut la demande de produits d'occasion. Il s'agit donc d'estimer l'impact d'un ensemble de variables socio-démographiques sur la demande d'automobiles pour chacun des segments du marché, qu'il s'agisse de véhicules neufs ou d'occasion.

Pour les besoins de l'analyse, l'auteur considère que le choix se fait en deux étapes : tout d'abord le consommateur choisit entre véhicules neufs et d'occasion ; dans un deuxième temps, il choisit un segment.³⁶

Le choix entre neuf et occasion est estimé avec un modèle *Probit*. La variable expliquée est une variable booléenne (prenant la valeur 1 si le véhicule acquis est neuf, 0 s'il est d'occasion). Les variables explicatives retenues incluent l'âge et le niveau d'éducation de la personne de référence, le lieu de résidence du ménage (département ayant une frontière avec l'étranger ou non), le revenu moyen par tête, le rang du véhicule acquis (rang chronologique, en fonction de la date d'acquisition) et le prix du véhicule.

Deux analyses distinctes sont alors menées. Pour un véhicule neuf, un modèle *logistique ordonné* est estimé. La variable expliquée est une variable ordinale égale à 0 si le véhicule acquis appartient au segment inférieur, 1 s'il appartient au segment moyen inférieur, 2 au segment moyen supérieur, et 3 au segment supérieur. Les variables explicatives retenues incluent le revenu par unité de consommation, l'usage du véhicule (professionnel ou non), le rang du véhicule, la taille du ménage et son degré d'urbanisation. Pour un véhicule d'occasion, le modèle et la variable dépendante sont les mêmes (logistique ordonné, variable ordinale entre 0 et 3), mais un indicateur d'usure (différence entre le prix d'achat du véhicule et le prix de l'équivalent neuf) est ajouté à la liste des variables explicatives.

Les données utilisées dans le cadre de cette analyse sont issues de l'enquête «Budget des Familles» de l'INSEE, de 1995.

Trois grands ensembles de conclusions se dégagent des résultats d'estimation :³⁷

- *sur le choix entre véhicules neufs et d'occasion* : la probabilité de choisir un véhicule neuf diminue avec le prix et augmente avec le revenu. Les achats de véhicules neufs sont positivement corrélés avec l'âge de la personne de référence et avec son niveau d'éducation. Les achats de véhicules neufs sont négativement corrélés avec le rang du véhicule (lorsque le véhicule acquis est une voiture principale, la probabilité d'opter pour un véhicule neuf est plus élevée que dans le cas d'un véhicule secondaire). Enfin, un ménage vivant dans un département ayant une frontière terrestre avec un pays étranger aura plus souvent tendance à acquérir un véhicule neuf.
- *sur le choix de segments dans le cas de véhicules neufs* : l'augmentation du revenu ou de la taille du ménage oriente les choix des ménages vers les trois segments supérieurs (par rapport au segment inférieur). L'achat d'un véhicule principal oriente également les choix vers les segments supérieurs. De même, l'utilisation du véhicule pour des déplacements professionnels contribue positivement à l'achat de véhicules de gamme supérieure. Enfin, la préférence des ménages français pour les marques nationales est démontrée, surtout sur le segment inférieur.

³⁶ En pratique, cette modélisation passe par un modèle emboîté, à deux étapes.

³⁷ Prieto (2006), pages 257 et 258

- *sur le choix des segments d'occasion* : le rang du véhicule acquis est un facteur déterminant de choix. Le fait d'acheter un véhicule principal oriente le choix des ménages vers les trois segments supérieurs (par rapport au segment inférieur). Enfin, la probabilité qu'un ménage choisisse un véhicule de gamme supérieure est plus élevée si la personne de référence est un homme.

Greene et al. (2005) a pour objectif d'évaluer le système américain des « feebates »³⁸ et d'analyser les élasticités-prix de la demande de véhicules et de la demande de carburant.

Dans le cadre de cette analyse, un modèle multinomial logistique emboîté a été adopté afin de modéliser le choix des consommateurs (en termes de véhicules). Il s'agit d'un modèle national agrégé, utilisant des données sur les ventes de véhicules sur une année (2000).

Le modèle part du principe que le consommateur oriente ses choix vers la solution (globalement) la moins onéreuse. Dès lors, pour chaque véhicule (c'est-à-dire chaque marque et chaque modèle), on cherchera à prendre en compte : i) les dépenses en termes de carburant ; ii) le surplus payé à l'achat de manière à pouvoir utiliser une technologie nouvelle ; et iii) le « feebate » éventuel.³⁹

Les classes de véhicules ainsi que les données relatives aux majorations de prix suite à l'introduction de technologies nouvelles (permettant de réduire la consommation de carburant) sont issues d'une étude réalisée par le « National Research Council. » Les données permettant l'estimation du « feebate » proviennent de la « National Highway Traffic Safety Administration. »

Le modèle de simulation retenu pour l'analyse repose sur un certain nombre d'hypothèses concernant les distances moyennes parcourues annuellement, le coût du carburant, ou la durée de vie des véhicules. Par ailleurs, les véhicules ont été répartis en 11 classes ; trois niveaux de consommation de carburant sont envisagées (faibles, moyennes et élevées) ; et on considère que les consommateurs, au moment d'acquies un véhicule, estiment les économies de carburant sur trois ans sans actualisation, ou bien les économies à vie, avec actualisation.

Au total, 19 scénarios ont été considérés : deux font l'hypothèse d'absence de politique ; dix sont consacrés à l'évaluation des « feebates » (en variant leur montant et/ou le nombre et la valeur des pivots) ; quatre à l'évaluation de politiques de « rebates » ; et deux à celle des taxes sur les « gas-guzzler. » Les coefficients d'élasticité sont multipliés par deux dans un dernier scénario (analyse de sensibilité).

Les principaux résultats de l'étude peuvent être résumés ainsi :

- *introduction de « feebates »* : l'introduction de feebates a un impact significatif sur la consommation de carburant dans tous les cas de figure ;
- *nombre de pivots mis en place dans le cadre des « feebates »* : l'introduction d'un nombre différent de pivots (1, 2 ou 11 – pour chacune des classes de véhicules) n'a pas d'impact significatif sur la consommation ;
- *élasticités-prix* : dans l'immense majorité des cas, les économies réalisées en carburant suite à l'introduction de « feebates » tiennent à l'adoption de véhicules utilisant des technologies nouvelles, et non pas à un changement de classe ou type de véhicule ;
- *les « rebates » et taxes « gas-guzzler »* : ces politiques ne pouvant s'appliquer qu'à une partie des véhicules affectés par les « feebates », leurs effets sur la consommation de carburant sont logiquement inférieurs. Contrairement à ce qui a été observé sur les « feebates », les pivots mis en place ont un impact réel pour les

³⁸ Aux Etats-Unis, les « feebates » sont un programme par lequel les véhicules à forte consommation (au delà d'une valeur pivot) se voient imposer une surcharge (un malus), alors que les véhicules à faible consommation (en deçà de la valeur pivot) bénéficient d'une ristourne (bonus). Le système des « rebates » ne prévoit qu'un bonus. La « gas-guzzler tax » prévoit un malus pour les automobiles offrant 22.5 « miles per gallon » ou moins.

³⁹ Il est important de noter que Green et al. (2005) n'estime pas de modèle économétrique, mais simule un modèle théorique en utilisant des données et des paramètres existants. L'analyse s'appuie sur un modèle logistique multinomial emboîté qui décrit les choix des ménages par une fonction d'utilité dans laquelle la consommation de carburant est un élément clé. Ce modèle théorique est calibré en utilisant des coefficients d'élasticité existants et des données sur les parts de marchés observées (les constantes du modèle sont déterminées de manière à reproduire au plus près les parts de marché de l'année de base).

« rebates » et les taxes « gas-guzzler ». Par ailleurs, si les taxes « gas-guzzler » découragent les consommateurs d'acheter des véhicules à forte consommation, ils ne les encouragent pas pour autant à acheter des véhicules à consommation réduite.

Mohammadian et Miller (2003) s'inscrit également dans le contexte des efforts de réduction des gaz à effets de serre, en proposant une analyse permettant de mieux appréhender la réaction des consommateurs à certaines actions publiques.

L'étude retient une modélisation des choix en deux étapes : i) le ménage décide d'acquérir un nouveau véhicule, de garder le sien, ou de le remplacer ; et ii) il choisit une marque et un modèle spécifiques. La modélisation présentée ci-dessous ne concerne que la deuxième étape (la première étant abordée dans une autre étude).

L'approche retenue par les auteurs repose sur un modèle multinomial logistique emboîté, qui fonctionne en deux temps. En premier lieu, on fait l'hypothèse que le ménage choisit une classe de véhicule (« sub-compact », « compact », « midsize », « large », « Special Purpose Vehicle » et « van »); puis en second lieu, il choisit parmi un véhicule « brand new », « second hand », « used » ou « old »⁴⁰.

L'utilité retirée par les ménages de l'utilisation de leur véhicule est exprimée en fonction de 29 variables, rassemblées en quatre grandes catégories : les caractéristiques du ménage ; les caractéristiques de son parc automobile ; les caractéristiques du véhicule ; et un ensemble de constantes relatives à la classe et à l'ancienneté du véhicule.

En ce qui concerne les caractéristiques des véhicules, un travail important de traitement des données a été entrepris. Les données (brutes) disponibles incluent : le poids, la taille (espace intérieur), la consommation, la cylindrée, la taille du coffre, l'origine, l'empattement, le type, et l'âge. Certains de ces attributs étant corrélés, les auteurs se sont livrés à une analyse en composante principale⁴¹, ayant permis la définition de deux variables synthétiques : la performance (regroupant poids, consommation et cylindrée) et l'espace (regroupant taille et type du véhicule, empattement, et taille du coffre).

Les informations relatives à la consommation des ménages proviennent du « Toronto Area Car Ownership Study » qui regroupe des informations sur les transactions automobiles effectuées entre 1990 et 1998, par 900 ménages vivant à Toronto.⁴² Les informations relatives aux caractéristiques des véhicules sont issues du « Canadian Vehicle Specifications System » qui rassemble des informations sur les modèles de véhicules fabriqués entre 1971 et 1999. Les informations relatives à la consommation de carburant proviennent de la « Fuel Economy Guide Database. » Enfin, les informations relatives au prix des véhicules (au moment de leur achat) sont issues du « Canadian Red Book. »

Parmi les principaux enseignements de l'étude, on retiendra que :

- les ménages disposant du niveau d'éducation le plus élevé sont plus enclins à acheter des véhicules neufs et moins enclins à acheter des véhicules d'occasion ;
- la durée moyenne de détention des véhicules appartenant aux catégories d'ancienneté « used » et « old » est quasiment identique ;
- l'impact du facteur synthétique « performance » (premier axe factoriel) est systématiquement positif ;
- l'impact du facteur « espace » (deuxième axe factoriel) n'est positif que pour les « Special Purpose Vehicles » et les « vans » ;

⁴⁰ « Neuf », « de deuxième main », « usagé » et « vieux. »

⁴¹ L'analyse en composante principale, ou ACP, est une technique mathématique permettant de réduire les dimensions d'un nuage de points (réduction du nombre de variables). Les composantes principales sont des variables synthétiques (combinaisons linéaires des variables d'origine) définies de telle sorte qu'elles préservent au mieux la dispersion des données. La première composante (le premier axe) est celle qui conserve au mieux la distance entre les points ; la deuxième composante est perpendiculaire au premier axe et de variance maximale; etc. En éliminant les axes de moindre importance (de plus haut degré), on peut ainsi « réduire les données » tout en préservant un maximum d'information.

⁴² Étude réalisée par l'Université de Toronto.

- les ménages disposant du niveau d'éducation le plus élevé, ou ceux dont la moyenne d'âge est avancée, sont moins enclins à acheter un « Special Purpose Vehicle » ;
- les conducteurs hommes sont plus enclins à acheter des véhicules de grosse taille ; et
- la disponibilité d'une classe de véhicules impacte positivement sa demande (ce qui implique que la consommation des ménages varie en fonction de leur situation géographique).

Cecchini (2001) replace le choix de véhicules dans le contexte général de la demande de produits différenciés sur un marché oligopolistique. Elle adopte un modèle national agrégé de type multinomial logistique emboîté afin de modéliser le choix des consommateurs *français* en matière de véhicules.⁴³

Les variables explicatives sont regroupées en trois sous-groupes :

- - caractéristiques du véhicule : prix, originalité⁴⁴, tenue de route, robustesse, esthétique, finition, capacité (5 sièges ou non), positionnement dans la gamme ;
- - caractéristique des ménages (non spécifiés) ; et
- - origine du véhicule : France, Etats-Unis, Grande Bretagne, Japon, Allemagne ou Italie.

Les données utilisées datent de 1994 et concernent 195 consommateurs disposant d'un véhicule de petite taille (l'étude ne repose que sur 12 véhicules différents).⁴⁵

La variable la plus changeante d'une nationalité à l'autre est la robustesse ; elle conduit à privilégier les véhicules français ou allemands. Les caractéristiques qui influencent les décisions d'achat restent le prix, l'originalité, la tenue de route, la robustesse et le positionnement du véhicule dans sa gamme. Enfin, dans 60 pourcent des cas, le modèle parvient à « prédire » correctement le véhicule des ménages et les parts de marchés correspondantes.

McCarthy et Tay (1998) a pour objectif de comprendre les facteurs influant sur la demande de véhicules neufs, et notamment de comprendre en quoi la demande de véhicules à faible consommation (de carburant) diffère de la demande de véhicules à forte consommation.

Afin de mener à bien l'analyse, un modèle multinomial logistique emboîté a été préféré au modèle logistique traditionnel. Trois groupes de véhicules sont créés en fonction de leur consommation : faible, normale (moyenne), et forte.

L'utilité retirée par les ménages est exprimée en fonction i) d'éléments de coûts (prix du véhicule par rapport au revenu annuel, coûts variables par mile) ; ii) des caractéristiques du véhicule (airbag, longueur, coffre, etc.) ; iii) d'éléments socio-économiques (achat de véhicules japonais par les ménages résidant sur la côte ouest, fidélité à une marque, etc.) ; et iv) des caractéristiques du véhicule en relation avec sa consommation de carburant.

Les données utilisées proviennent d'un sondage réalisé en 1989 par JD Power and Associates intitulé « New Car Buyer Competitive Dynamics Survey » rassemblant des observations sur 33 284 acheteurs de véhicules neufs. Ce sondage ne contenant que des informations relatives aux modèles achetés et caractéristiques des acheteurs, il a été complété par des informations sur les attributs des véhicules. Il restait 1 564 observations après traitement des données (c'est-à-dire après élimination des observations pour lesquelles des informations clefs manquaient, et prise en compte d'un panel à l'image de la répartition réelle des véhicules dans la population américaine).

L'estimation du modèle introduit ci-dessus a permis de mettre en avant les éléments suivants :

- il est confirmé que l'amélioration de la taille, de la sécurité et de la qualité d'un véhicule accroît sa demande ;
- les marques les plus appréciées (aux Etats-Unis) sont, dans l'ordre décroissant : Nissan, Toyota, Honda, Ford, General Motors, et Mazda ;

⁴³ Le choix d'un véhicule est réalisé en deux étapes : choix d'un pays d'origine ; puis choix d'une marque et modèle.

⁴⁴ La définition de cette variable qualitative n'est malheureusement pas précisée.

⁴⁵ La source exacte des données n'est pas identifiée.

- les femmes, les ménages les plus défavorisés, les jeunes, les ménages « non blancs » ainsi que les ménages résidant dans les zones les plus densément peuplées sont plus demandeurs de véhicules à faible consommation ; et
- l'analyse révèle également que le sous-groupe disposant de véhicules à faible consommation est le moins sensible aux variations de prix au sein de son sous-groupe, alors que le sous-groupe disposant de véhicules à forte consommation est le moins sensible aux variations de prix dans les autres sous-groupes.

Berry et al. (1995) estime un modèle de choix discret *agrégé*, en utilisant une variante de l'approche logistique mixte. Dans le modèle, les ménages choisissent le véhicule automobile dont les caractéristiques observées et non-observées maximisent leur utilité. Les parts de marché sont déterminées en considérant conjointement l'ensemble des caractéristiques individuelles (observées et aléatoires) conduisant au choix d'un certain modèle.⁴⁶ La fonction d'utilité associée au choix du véhicule, spécifiée selon un modèle à coefficients aléatoires, est insérée (« nested ») dans une fonction d'utilité plus globale⁴⁷, qui a pour deuxième composante, les dépenses sur un bien externe (défini ici comme la différence entre le revenu des ménages et le prix d'achat du véhicule). Cette spécification autorise la prise en compte des interactions entre les caractéristiques des agents d'une part et les caractéristiques des véhicules d'autre part (et ce dans le cadre d'un modèle agrégé); et permet selon les auteurs une modélisation plus réaliste des effets de substitution.

Les données sur les attributs des véhicules sont extraites du « Automotive News Market Data Book ». Un total de 2 217 modèles sont considérés⁴⁸. La base de données utilisée pour l'analyse ne comprend que des véhicules neufs, mais couvre une période de vingt ans (de 1971 à 1990).

Les résultats principaux de l'étude peuvent se résumer ainsi :

- les élasticités de la demande par rapport aux attributs des véhicules révèlent que la demande de véhicules à faible consommation est sensible aux changements d'efficacité énergétique; cette élasticité étant plus réduite pour les véhicules à forte consommation. Par exemple, l'élasticité des parts de marchés par rapport à l'efficacité énergétique est de 1,01 pour une Mazda 323 et -0,02 pour une BMW 735i.
- les semi-élasticités croisées sont généralement élevées pour les véhicules ayant des caractéristiques comparables. Par exemple, une augmentation de \$1.000 du prix d'une Mazda 323 augmente la part de marché d'une Nissan Sentra de 0,75 pourcent, mais n'affecte pas la part de marché de modèles haut-de-gamme. De même, la BMW 735i a la plus forte semi-élasticité croisée avec la Lexus LS400, toutes deux étant considérées comme des voitures de luxe.
- la probabilité de substitution du bien externe (défini comme l'absence d'achat) à l'acquisition d'une voiture neuve est plus élevée pour les véhicules bas-de-gamme.

Berry et al. (2004) utilise une approche comparable à Berry et al (1995) et estime un modèle de choix discrets *agrégé*, dont la fonction d'utilité sous-jacente est à coefficients aléatoires. L'innovation principale de cet article est l'utilisation de données sur un hypothétique « second choix ». Ces données sont issues d'une enquête commanditée par la société General Motors. Pour chacun des 203 modèles de l'année 1993 retenus, un échantillon aléatoire d'acheteurs a été déterminé (pour un total de 37 500 ménages interrogés). Outre l'identification du second choix, le questionnaire comprenait des informations relatives aux ménages (revenu, âge du chef de famille, lieu de résidence (urbain ou rural), et taille du ménage) et aux véhicules achetés (y compris le prix d'achat effectif). Parmi les résultats de l'étude, on retiendra que :

- les véhicules retenus comme second choix ont des caractéristiques très proches des véhicules choisis pour l'achat ;
- les semi-élasticités prix (élasticité des parts de marché par rapport à un changement absolu du prix d'achat) décroissent avec le prix, et à niveau de prix constant, sont inférieurs pour les vans et mini-vans, les « Sport

⁴⁶ En termes plus techniques, la part de marché – ou demande agrégée – du véhicule k (portion des ménages choisissant le véhicule k), est obtenue en prenant l'intégrale d'une fonction de choix individuelle sur l'ensemble des attributs des ménages qui impliquent une préférence pour le véhicule k.

⁴⁷ De type "Cobb-Douglas."

⁴⁸ Chaque "modèle" étant défini par un modèle et une année.

Utility Vehicles » et – dans une moindre mesure – les voitures de sport que pour les autres catégories de véhicules ;

- il est possible de déterminer des « best price substitutes » pour chacun des modèles analysés. Par exemple, le meilleur substitut prix pour la Toyota Corolla est la Honda Civic puis, en second lieu, la Ford Escort (ensemble, ces deux modèles représentent 25 pourcent des substitutions à la Toyota Corolla lorsque le prix de celle-ci augmente) ; et
- le modèle prédit le meilleur substitut (c'est-à-dire le second choix) 9 fois sur 17 ; il prédit l'un des trois meilleurs substituts 15 fois sur 17.

Berkovec et Rust (1984) cherche à comprendre le processus de décision qui sous-tend le choix d'une automobile dans les ménages ne disposant que d'un seul véhicule. Les auteurs adoptent une méthode de choix discrets et ont privilégié l'utilisation d'un modèle logistique multinomial emboîté.

L'utilité retirée d'un véhicule donné est formalisée en fonction : i) des coûts associés au véhicule (prix d'achat et coûts opérationnels) ; ii) de sa performance et de sa capacité (taille du véhicule compte tenu de la taille du ménage, rapport nombre de chevaux / poids, autonomie) ; et iii) d'autres éléments (dont une variable booléenne appelée « transaction » permettant de différencier les cas où les véhicules avaient été acquis l'année précédente l'enquête (« transaction » = 0) des autres cas (« transaction » = 1)).

Les données utilisées pour estimer le modèle proviennent d'un sondage réalisé en 1978 par Cambridge Systematics Inc. pour la « National Science Foundation » auprès de 1.095 ménages américains.

De cet échantillon, les auteurs ont exclu : i) les ménages ne possédant pas de véhicule ou, à l'inverse, possédant plus de deux véhicules ; ii) les ménages dont le véhicule était un van, un pick-up, un camion ou un utilitaire ; iii) les ménages dont le véhicule datait d'avant 1967 ; et iv) les ménages dont les données n'étaient pas disponibles pour au moins l'un des critères retenus. A l'issue de ce traitement, 237 observations restaient utilisables.

Les données sur les caractéristiques des véhicules comportent les éléments suivants : prix d'achat, poids, consommation (en miles par gallon), nombre de places, autonomie, rayon de braquage, et nombre de chevaux. Ces caractéristiques ont permis de grouper les véhicules en quinze catégories.

Les résultats d'estimation se conforment aux attentes théoriques. Les facteurs relatifs aux coûts (prix d'achat, coût du carburant, etc.) ont un effet négatif sur la demande.⁴⁹ Les caractéristiques comme le rayon de braquage ont également l'effet attendu. La variable « transaction » est positive et significative. Il s'en suit que la possession de véhicules est caractérisée par une forte inertie: sur chaque période, il est plus probable qu'un ménage conserve son véhicule plutôt qu'il en achète un nouveau. Cette inertie peut s'expliquer, selon les auteurs, par l'importance des coûts de transactions caractérisant le marché de l'automobile.

⁴⁹ L'effet du prix, cependant, n'est pas statistiquement significatif pour les ménages dont le revenu excède \$15,000 (environ 50,000 euros d'aujourd'hui).

3.3. Récapitulatif et enseignements pour cette étude

Un récapitulatif des principales informations recueillies au cours de notre revue de la littérature est produit dans le Tableau 3.1, ci-dessous.

Tableau 3.1
Résumé des Principales Études

Études	Marché	Méthodologie	Données	Résultats et Contributions
Eftec (2008)	Royaume Uni	Modèle agrégé, multinomial logistique emboîté	Immatriculations (2001-2006) et base JATO	Décisions d'achat de véhicules neufs ; élasticités-prix directes et croisées, par classe de véhicules et niveaux d'émission ; outil de simulation
CE et BMRB (2008)	Royaume Uni	Modèle désagrégé, multinomial logistique mixte	Sondage auprès de ménages britanniques, et base JATO	
Train et Winston (2007)	États Unis	Modèle désagrégé, multinomial logistique mixte	Sondage auprès de ménages américains, Market Data Book, Wards' Automotive Yearbook	Évolution des parts de marché des constructeurs américains ; caractéristiques des véhicules et autres facteurs (dont la qualité des réseaux de distribution)
Bento et al. (2006)	États Unis	Modèle de choix discrets-continus (choix de véhicule et distance parcourue)	National Household Travel Survey (2001), Wards Automotive Yearbook, National Automobile Dealer's Association Used Car Guide, site Internet du Department of Energy	Effets d'une augmentation des taxes sur la consommation de carburant et le choix de véhicules
Prieto (2006)	France	Modèle emboîté à deux étapes : neuf ou occasion (Probit) ; choix de segment (logistique ordonné)	Enquête «Budget des Familles» de l'INSEE (1995)	Déterminants des choix entre véhicules neufs et d'occasion, et des choix de segment (inférieur, moyen inférieur, moyen supérieur, et supérieur)
Greene et al. (2005)	États Unis	Modèle de simulation (sur la base d'un modèle national agrégé, multinomial logistique emboîté)	Étude réalisée par le National Research Council (2002) ; modèle de la National Highway Traffic Safety Administration (2000)	Analyse du système des « feebates » et autres mesures incitatives

Études	Marché	Méthodologie	Données	Résultats et Contributions
Mohammadian et Miller (2003)	Canada	Modèle multinomial logistique emboîté à deux étapes : i) choix de classe et ii) neuf ou occasion (deuxième main, usagé, vieux)	Toronto Area Car Ownership Study (1990 – 1998) : sondage après de 900 ménages de Toronto ; Canadian Vehicle Specifications System ; Fuel Economy Guide Database ; Canadian Red Book	Déterminants des comportements d'achat et de détention des véhicules ; définition de facteurs « performance » et « espace » (ACP)
Cecchini (2001)	France	Modèle national agrégé, multinomial logistique emboîté	195 consommateurs disposant d'un véhicule de petite taille ; 12 véhicules différents (source exacte non spécifiée)	Déterminants des comportements d'achat ; utilisation de la nationalité comme une étape de choix (« nest »)
McCarthy et Tay (1998)	États Unis	Modèle multinomial logistique emboîté	New Car Buyer Competitive Dynamics Survey (1989) : sondage auprès de 33.000+ acheteurs ; 1.500 observations après traitement	Analyse de la demande de véhicules neufs, par niveau de consommation de carburant (faible, moyenne, forte)
Berry et al. (1995)	États Unis	Modèle agrégé ; modèle à coefficients aléatoires augmenté	Automotive News Market Data Book ; données de ventes et caractéristiques de 2.217 combinaisons modèle-année sur la période 1971-1990	Analyse des déterminants de la demande de véhicules ; analyse des effets de substitution
Berry et al. (2004)	États Unis	Modèle agrégé ; modèle à coefficients aléatoires augmenté	Sondage commandité par General Motors ; modèles de l'année 1993 ; 203 véhicules ; 37.500 observations ; information sur les seconds choix	Analyse des déterminants de la demande de véhicules avec prise en compte des seconds choix ; détermination des meilleurs substituts prix
Berkovec et Rust (1984)	États Unis	Modèle multinomial logistique emboîté	Sondage auprès de 1 095 ménages britanniques (1978)	Déterminants des comportements d'achat et de détention des véhicules ; variable « transaction » et inertie

Source : NERA

Les principaux enseignements de cette revue de la littérature peuvent être formulés ainsi :

- le modèle logistique emboîté est une alternative communément utilisée pour la modélisation des choix de véhicule ; sa spécification exacte (nombre et nature des étapes) varie d'une étude à l'autre, même si dans la plupart des cas, il s'agit de spécifications relativement simples (deux étapes) ;
- le modèle logistique mixte est également très répandu, mais i) il est d'application plus complexe, et ii) il nécessite des informations sur les seconds choix (qui souvent ne sont accessibles que par l'intermédiaire de sondages réalisés spécifiquement pour les besoins de l'étude) ;
- la richesse des données utilisées pour l'estimation des modèles de choix est variable ; dans le cadre de modèles agrégés (analyse des parts de marché), l'utilisation de base de données sur les immatriculations de véhicules neufs est assez fréquente ; dans le cadre de modèles désagrégés (modélisation des choix individuels des ménages), l'analyste doit recourir à des données d'enquête. Dans les deux cas, la qualité de l'estimation peut être considérablement accrue en introduisant des variations temporelles ;
- les indicateurs de marché (immatriculations) ou données d'enquête sont souvent complétés par des informations techniques provenant de bases de données ou publications spécialisées (de type JATO) ;
- la littérature offre un très grand nombre de suggestions relatives à la sélection des attributs des véhicules (caractéristiques physiques, prix, coût fixe, coût variable, etc.), ou à la modélisation de certains facteurs de choix (bien externe, fidélité à une marque, etc.). Elle propose également différents systèmes de classification pour les véhicules (par gamme, pays d'origine, niveau de consommation, et autres) qu'il conviendra de revisiter ;
- un certain nombre d'outils de simulation ont été développés sur la base de modèles de choix discrets (voir Eftec (2008) ou Greene et al. (2005) par exemple) ; et
- les deux études commanditées récemment par le ministère des transports britannique offriront une excellente référence pour le développement de notre méthodologie. Par ailleurs, il peut être intéressant de comparer – et si possible valider – les résultats obtenus dans le cadre de notre étude, en termes notamment d'élasticités-prix directes ou croisées.⁵⁰

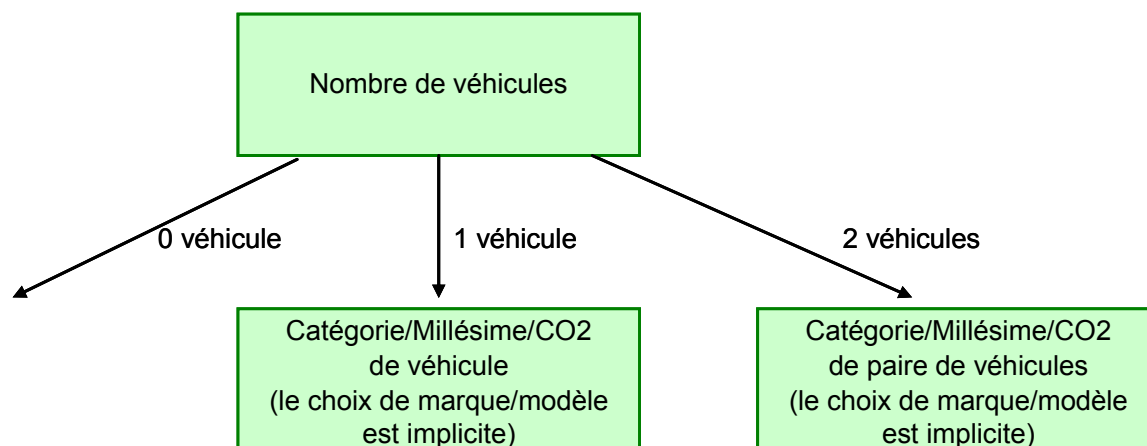
4. Méthodologie

Le modèle que nous développons concerne le choix des ménages pour l'acquisition de véhicules. Cette décision se subdivise en deux sous-options: (i) le nombre de véhicules à posséder, et (ii) la catégorie, le millésime, le type de carburant et les émissions de CO₂ de chaque véhicule possédé. Cette approche reflète, en partie, le modèle du choix d'un véhicule développé par Train (1986), et elle est bien adaptée pour répondre aux objectifs de la présente étude. Elle est capable de générer des estimations fiables des élasticités de substitution entre types de véhicules lors de changements dans les coûts fixes et variables des véhicules. Ces estimations permettront de simuler les effets des politiques qui affectent les coûts des véhicules, et particulièrement le «bonus écologique», sur la demande de véhicules selon la catégorie, le millésime et / ou les émissions de CO₂.

⁵⁰ Une comparaison des élasticités françaises et britanniques est disponible dans Goldberg et Verboven (2001). Les auteurs analysent la dispersion des prix sur le marché automobile européen et estiment un ensemble d'élasticités-prix par pays, à l'aide d'un modèle multinomial logistique emboîté sur données agrégées. On notera en particulier que i) l'élasticité-prix directe moyenne est plus élevée au Royaume Uni qu'en France (-6.21 contre -5.60); et ii) les élasticités-prix croisées des deux pays sont généralement assez proches (à l'exception de l'élasticité-prix croisée moyenne pour les voitures domestiques – suite à un changement de prix d'une voiture domestique – sensiblement plus élevée en France : 0.170 contre 0.076). Les données utilisées dans cette étude sont assez anciennes (elles sont antérieures à l'abolition de nombreuses restrictions à l'importation), et les conclusions établies ne sont sans doute pas transposables à l'époque actuelle. De manière plus générale, il faudra rester prudent lors de la comparaison de nos résultats avec ceux obtenus dans d'autres pays (en raison notamment de différentes structures de marché et/ou caractéristiques de l'offre).

Le schéma 3.1 illustre la structure du modèle.

Graphique 4.1
Modèle de Choix de Véhicule des Ménages*



* Chaque rectangle représente un sous modèle spécifique

Le modèle est formulé à travers un logit hiérarchique, qui traite les choix comme ayant été effectués simultanément.

La structure du modèle adopté est liée, de manière assez importante, à de récentes études de premier ordre, notamment de Cambridge Econometrics / BMRB (2008), Train et Winston (2007) et Eftec (2008). Cambridge Econometrics / BMRB (2008) et Train et Winston (2007), tout comme la présente étude, utilisent des données non globales d'enquête portant sur les choix individuels des ménages. Cependant, ces deux études se concentrent sur les achats de nouveaux véhicules plutôt que sur tous les types de véhicules, et utilisant la marque et le modèle plutôt que la catégorie/millésime/type de carburant/émissions de CO2 comme base du choix. Les modèles de substitution entre nouveaux véhicules estimés par ces études sont dépendants, par exemple, des changements de prix du véhicule, mais ne tirent pas de conclusions sur la demande globale de véhicules neufs, ni sur la substitution entre véhicules de différents millésimes. En revanche, notre modèle, basé sur les choix d'acquisition du véhicule, est capable de faire ce lien. Eftec (2008) estime un modèle de choix en utilisant les données globales sur les parts de marché de chaque marque / modèle toujours en production. Cette étude est également axée sur le marché des véhicules neufs, mais est capable de prédire les niveaux de la demande globale ainsi que les parts de marché des différentes marques / modèles. Cependant, comme les deux autres études, le modèle de Eftec ne permet pas de prédire la substitution entre véhicules de différents millésimes. Le modèle de Eftec (2008) ne peut aussi que faiblement identifier l'effet des caractéristiques des ménages sur les choix de véhicules en raison de sa dépendance à la variation limitée des caractéristiques moyennes des ménages au fil du temps, et à travers les régions au Royaume-Uni, pour estimer ses paramètres. L'approche globale de Eftec est utilisable pour la présente étude, mais entraînerait le rejet de toutes les données sur le choix de véhicule fournies par l'INSEE en faveur de données globales sur les parts de marché. Nous pensons que cela reviendrait à perdre une source de données très importante. Notre choix de structure du modèle, en revanche, permet l'usage des données de l'INSEE dans son intégralité.

Dans le reste de cette section, nous exposons la spécification du modèle. Nous décrivons comment il peut être utilisé dans un outil de simulation pour prévoir la demande globale, par type de véhicule, selon des scénarios politiques utilisables. L'exposé qui suit se veut non technique. Les détails mathématiques de la spécification du modèle sont présentés en Annexe A.

4.1. Spécification du Modèle

Notre approche pour modéliser le choix d'un véhicule s'inscrit dans la théorie économique du comportement des consommateurs, dans lequel les ménages sont censés choisir l'option qui maximise leur utilité, dans une situation de choix donnée. Dans cette approche, l'utilité d'un ménage est définie comme étant la somme d'une fonction déterministe de l'option et des caractéristiques du ménage, ainsi que d'un terme d'erreur. L'approche génère des estimations sur les arbitrages que les ménages font entre les caractéristiques des options du choix, et permet ainsi la simulation des effets de substitution dans le marché dans son ensemble, compte tenu d'un changement dans l'une de ces caractéristiques.

Le type spécifique du modèle que nous proposons d'adopter est le logit hiérarchisé, un type très répandu des modèles de la famille de la Valeur Extrême Généralisée (VEG). En comparaison avec le logit standard, traditionnellement utilisé pour la modélisation des choix discrets, les logit hiérarchisés permettent une modélisation plus réalistes de substitution entre les solutions alternatives. Le type de substitution permis dépend de la hiérarchisation adoptée. Dans notre cas, suivant Train (1986) nous adoptons un modèle logit hiérarchisé d'acquisition de véhicules ayant trois niveaux d'imbrication, l'un étant implicite et non estimé. Comme illustré dans le schéma 3.1, la limite supérieure de l'arbre décisionnel est le choix du nombre de véhicule possédé. Le niveau inférieur est le choix du type de véhicule posséder compte tenu du nombre de véhicules. Implicitement dans ce niveau inférieur de hiérarchisation, il y a un troisième niveau qui est le choix des marques / modèles à acquérir étant donné le nombre et le type de véhicules. Ce troisième niveau n'est pas estimé à l'aide d'un modèle, car cela n'ajouterait, le cas échéant, qu'une faible amélioration de l'outil de simulation des politiques, en contrepartie d'une plus grande complexité.

Le premier sous-modèle envisagé est donc le choix du nombre de véhicules à acquérir. Ici, le choix de posséder zéro, un ou deux véhicules dépend des caractéristiques des ménages telles que le nombre de personnes, leur âge, leur cadre socio professionnel et leurs revenus, la qualité et le prix des transports publics et les infrastructures routières dans la région du ménage et, surtout, les caractéristiques des options de choix de types de véhicules, y compris les coûts fixes et variables. Ce dernier facteur constitue le lien entre le nombre de véhicules à acquérir et les sous-modèles du choix de type de véhicule. Il faut noter que l'option de disposer de trois véhicules ou plus est exclue de ce sous-modèle afin de simplifier considérablement l'étude. Nous appliquons une correction lors de l'agrégation de la demande afin de veiller à ce que cette exclusion ne conduise pas à biaiser nos estimations de la demande globale de véhicules.

Le deuxième sous-modèle que nous considérerons est celui du choix du type de véhicule (catégorie/millésime/type de carburant/émissions de CO₂) par les ménages, étant donné qu'ils ont choisi de posséder un seul véhicule. Le choix de ce sous-modèle dépend des caractéristiques moyennes des véhicules de chaque type, et comment celles-ci interagissent avec les caractéristiques du ménage. Par exemple, les ménages avec enfants attacheront une plus grande valeur à la taille du véhicule que les ménages composés d'une seule personne ou en couple. Ce sous-modèle dépend également des caractéristiques des options de choix disponibles selon la marque / le modèle présent dans chaque type de véhicule. Le choix de la marque / modèle selon le type de véhicule sera donc implicite plutôt que d'être explicitement formulé dans un sous-modèle distinct.

Le troisième et dernier sous-modèle que nous considérons est le choix de la paire de véhicules que le ménage décide d'acquérir, étant donné qu'il a choisi de posséder deux véhicules. Les options de choix disponibles dans ce cas sont toutes les paires d'options possibles de catégorie/millésime/émissions de CO₂, où les options de catégorie / millésime et émissions de CO₂ sont définies comme dans le sous-modèle de choix d'un seul véhicule. Ce modèle est donc plus large que celui décrit précédemment, et nécessite donc une plus grande puissance informatique pour l'estimer.

4.2. Agrégation de la Demande de Véhicules, Prévision et Simulation

Le modèle peut être utilisé dans un outil de simulation pour prévoir la demande globale des véhicules par catégorie/millésime/type de carburant/émissions de CO₂, dans un cas de base et selon des simulations de politiques. Les données nécessaires pour cet outil sont les prévisions des caractéristiques de chaque véhicule, y compris les coûts fixes et variables des véhicules pour chaque année de simulation, et les prévisions des variables socio-économiques telles que la distribution de la population, des revenus et de la taille des ménages. De plus, le champ d'application et les montants des niveaux du "bonus écologique" seront nécessaires pour créer un outil spécifique pour les simulations pertinentes des politiques souhaitées par le MEDDTL.

Compte tenu de ces données, les résultats du modèle sont les probabilités estimées pour chaque ménage :

- d'acquérir zéro, un ou deux véhicules ;

- d'acquérir chaque type de véhicule, étant donné qu'ils ont choisi d'acquérir un seul véhicule ; et
- d'acquérir chaque paire de véhicules possible, étant donné qu'ils ont choisi de posséder deux véhicules.

Le modèle lie la taille du marché des véhicules neufs à la typologie des nouveaux millésimes. Par exemple, la demande estimée de véhicules neufs en 2009 correspond à la demande prévue pour l'acquisition d'un véhicule du millésime 2009 en 2009. Le modèle peut être calibré pour le marché des véhicules neufs, le cas échéant, en utilisant les données sur la demande de l'année la plus récente ayant des données disponibles.

Les estimations de la demande globale de véhicules par catégorie/millésime/type de carburant/émissions de CO2 sont obtenues en appliquant la formule suivante:

$$(1) \quad N_i = \sum_n \left[\left(P_1^n \times P_{i|1}^n \right) + \left(k P_2^n \times \sum_j P_{ij|2}^n \right) + \left(k P_2^n \times P_{ii|2}^n \right) \right],$$

où:

N_i est l'estimation de la demande globale des véhicules de catégorie/ millésime/ émissions de CO2 i ,

P_1^n et P_2^n sont les probabilités estimées que le ménage n acquière respectivement 1 et 2 véhicules,

k est le nombre moyen de véhicules détenus par les ménages qui détiennent au moins 2 véhicules, divisé par deux

$P_{i|1}^n$ est la probabilité estimée qu'un ménage n acquière un véhicule de catégorie/millésime/émissions de CO2 i étant donné qu'il a choisi d'acquérir 1 véhicule, et

$P_{ij|2}^n$ est la probabilité estimée qu'un ménage n acquière une paire véhicule de catégorie/millésime/émissions de CO2 i et j étant donné qu'il a choisi d'acquérir 2 véhicules.

De manière littéraire, cette formule indique que l'estimation de la demande globale pour les véhicules de catégorie/millésime/type de carburant/émissions de CO2 de type i est égale à la somme de trois termes sur chaque ménage. Le premier terme est la probabilité estimée que le ménage acquière un seul véhicule multiplié par la probabilité que ce véhicule soit de type i . Le deuxième terme est égal à une constante k , multipliée par la probabilité d'être propriétaire de deux véhicules, multipliée par la probabilité que l'un de ces deux véhicules soit de type i . La constante k , ici, est le nombre moyen de véhicules appartenant à des ménages qui possèdent au moins deux véhicules, divisé par deux. Elle est incluse pour corriger le fait que les ménages ayant trois véhicules ou plus sont exclus de la spécification du modèle. Ce terme n'est pas prédit par le modèle, et doit être retenu pour les années à venir, ce qui est une faiblesse de l'approche. Le troisième et dernier terme est k multiplié par la probabilité estimée que le ménage détienne deux véhicules, multiplié par la probabilité que ses deux véhicules soient de type i . Ce terme permet de prendre en compte le fait que certains ménages peuvent posséder deux véhicules de même type.

5. Données

La méthodologie décrite dans la section précédente a été développée de façon à utiliser au mieux la série de données de l'INSEE "Budget des Familles 2006" (BdF 2006), qui nous a été fournie pour mener à bien cette étude. Cette série de données contient des informations relatives à 10 240 ménages en 2006, avec de nombreuses variables dont une permettant l'identification des véhicules détenus, une liée au moyen de financement de l'achat de ces véhicules, et d'autres liées aux caractéristiques socioéconomiques. En plus de cette série de données, la méthodologie nécessite des données sur les caractéristiques des véhicules, que nous avons apportées, et des données sur les prix des carburants.

Cette section est divisée en trois parties. La première décrit les sources de données de façon plus détaillée et évalue les données disponibles selon les exigences de l'étude. La seconde partie décrit ensuite le traitement des données pour créer une série de données adaptée à l'analyse. La dernière partie présente des résumés statistiques sur les séries de données finaux.

5.1. Exigences Relatives aux Données et Disponibilité

La méthodologie nécessite des données relatives aux ménages, aux véhicules et au marché. Cette sous-section décrit et évalue les sources de données nous fournissant les informations nécessaires pour chaque catégorie.

5.1.1. Données Relatives aux Ménages

La série de données de l'INSEE, fournie par le MEDDTL, contient les réponses de ménages à un sondage réalisé en 2006. La liste complète des variables pertinentes est fournie en 0, l'0 fournissant un résumé statistique de ces variables. Dans le cadre de notre étude, les variables clés incluses dans le modèle sont les suivantes :

- Les variables de choix du véhicule – marque, modèle, l'année d'achat, l'année du modèle
- Variables de coût – prix d'achat, coûts fixes, coûts de fonctionnement
- Variables d'utilité du véhicule – utilisation pour les déplacements et/ou le travail et/ou le loisir
- Variables socioéconomiques – âge, éducation, revenus, pour le ménage et pour la personne désignée comme utilisateur principal du véhicule

Afin d'évaluer la portée de la série de données, le Tableau 5.1 montre les tailles des échantillons et les cadres de quelques études similaires réalisées récemment, ainsi que les caractéristiques de l'échantillon disponible dans la série de données de l'INSEE.

Tableau 5.1
Taille des Echantillons et Cadres d'Études Précédemment Réalisées

Variables	CE/BMRB1	T+W2	Eftec3	Train4	Disponible par l'INSEE
Nombre de ménages	900	458	N/A	1,095	10,240
Seulement les nouveaux achats?	√	√	√	X	X
Tous les véhicules détenus?	X	X	X	√	√
Non achetés modélisés?	X	X	√	√	√
Nombre d'année	1	1	5	1/Toutes	1/Toutes
Nombre de types de véhicules	1,700	200	2,190	120	150

(1) *Cambridge Econometrics/BMRB (2008)*; (2) *Train et Winston (2007)*; (3) *Eftec (2008)*; (4) *Train (1986)*

Comme ce tableau le montre, le nombre de ménages dans la série de données de l'INSEE est beaucoup plus élevé que pour les autres études. De plus, cette série de données ne se limite pas seulement les neufs car il contient des informations sur les véhicules neufs et d'occasions, ce qui enrichit cette analyse. Le set comprend également des données sur les non-acheteurs, plutôt que sur la population ayant acheté un véhicule une certaine année. Ceci nous permet de modéliser la demande agrégée ainsi que les parts de marché.

Par rapport aux données disponibles, le Tableau 5.2 compare les variables relatives aux ménages incorporées dans les récentes analyses comparables et montre les variables disponibles dans le set de l'INSEE.

Tableau 5.2
Variables Relatives aux Ménages Utilisées dans les Etudes Précédentes et Disponibles dans le set de l'INSEE

Variables	CE/BMRB1	T+W2	Eftec3	Train4	Disponible par l'INSEE
Véhicule acheté / détenu	✓	✓	✓	✓	✓
Date d'achat	✓	✓	✓	✓	✓
Véhicules considérés	✓	✓	X	X	X
Véhicules détenus précédemment	X	✓	X	X	X
Nombre de véhicules détenus	✓	X	X	✓	✓
Kilomètres parcourus dans chaque véhicule	X	X	X	✓	X
Utilisation du véhicule	X	X	X	✓	✓
Age du véhicule	✓	X	✓	✓	✓
Localisation	X	✓	✓	✓	✓
Age	✓	✓	X	✓	✓
Sexe	✓	✓	X	✓	✓
Revenu	✓	✓	X	✓	✓
Taille du ménage	✓	✓	X	✓	✓

(1) *Cambridge Econometrics/BMRB (2008)*; (2) *Train et Winston (2007)*; (3) *Eftec (2008)*; (4) *Train (1986)*

Comme indiqué dans le Tableau 5.2 le montre, la série de données de l'INSEE contient presque toutes les variables relatives aux ménages utilisées dans les principales études sélectionnées. Les exceptions (les variables absentes du set de l'INSEE) sont les véhicules pour lesquels les ménages sont intéressés mais qui ne sont finalement pas achetés, les véhicules précédemment détenus par les ménages, et les distances parcourues par les véhicules.

Savoir quels véhicules sont considérés par les ménages, mais non achetés, peut permettre d'utiliser des modèles à paramètre aléatoire afin d'examiner la variance de préférence entre les ménages attribuable à des facteurs non observables. L'information a été utilisée de la sorte dans *Cambridge Econometrics/BMRB (2008)* et *Train et Winston (2007)* par exemple. Sans cette information, les modèles à paramètre aléatoire sont probablement impossibles à estimer de façon pertinente. A notre connaissance, aucune étude n'a pour le moment utilisé les méthodes de paramètre aléatoire pour estimer la détention d'un type de véhicule, et nous n'aurions en tout état de cause pas choisi d'utiliser de telles méthodes afin de ne pas ajouter de la complexité au modèle. Nous ne considérons donc pas l'absence de cette information comme étant important pour notre étude.

Par rapport à la seconde variable manquante (les véhicules précédemment détenus par les ménages) l'information a été utilisée par *Train et Winston* pour estimer l'importance de l'attachement à une marque. Ce n'est pas un élément crucial pour notre étude, et la structure du modèle qui regroupe les marques/modèles par catégories, n'aurait pas permis pas une telle analyse, même si les données avaient été disponibles. L'absence de variable relative à l'historique de la détention de véhicules n'est donc pas un sujet pour cette étude.

La troisième variable utilisée dans les études précédentes, mais absente de la série de données de l'INSEE, est le kilométrage parcouru par chaque véhicule. Cette variable est utilisée dans la modélisation de *Train (1986)* de manière importante. Plutôt que d'être utilisée pour expliquer le choix des types de véhicules ou du nombre à acquérir, cette variable est utilisée comme la variable dépendante de deux sous-modèles additionnels. Dans ces deux sous-modèles, le choix du nombre de kilomètres parcouru, lié au fait de détenir un ou deux véhicules, est expliqué par les coûts de fonctionnement des véhicules détenus et par un certain nombre de caractéristiques des ménages et de la zone géographique. Nous ne sommes pas capables de répliquer ce modèle pour cette étude car nous ne disposons pas de données suffisantes sur les distances parcourues des véhicules. Nous sommes cependant capables d'estimer un modèle similaire qui utiliserait des données sur les dépenses en carburant des ménages pour leurs véhicules. Nous pensons qu'il serait utile pour le MEDDTL d'avoir la possibilité de prédire la consommation de carburant dans un cadre intégré à celui du choix de véhicule, mais ceci va au-delà de la présente étude.

5.1.2. Données relatives au véhicule

Les données sur les caractéristiques du véhicule, qui sont nécessaires pour estimer l'importance de celles-ci sur les choix de véhicules des ménages, proviennent de JATO. L'O fournit la liste complète des variables pertinentes fournies par cette organisation.

Pour évaluer la source de données JATO, nous avons également comparé les données disponibles aux données utilisées dans les principales études sélectionnées. Le Tableau 5.3 présente cette comparaison.

Tableau 5.3
Caractéristiques des Véhicules Utilisés dans les Etudes Précédentes et celles Disponibles dans la série de données JATO

Variable	CE/BMRB ¹	T+W ²	Eftec ³	Train ⁴	Disponible par JATO
Prix d'achat	✓	✓	✓	✓	✓
Valeur retenue	X	✓	✓	X	✓
Droits d'accises du véhicule	✓	X	✓	X	✓
Coûts d'assurances	X	X	✓	X	✓
Consommation de carburant	✓	✓	✓	✓	✓
Type de carburant	✓	X	✓	X	✓
Segment de marché	✓	✓	✓	X	✓
Constructeur	✓	✓	✓	X	✓
Taille du moteur	X	X	✓	X	✓
Emissions de CO2	✓	X	✓	X	✓
Nombre d'airbags	✓	X	✓	X	✓
Type de transmission	✓	✓	✓	X	✓
Nombre de vitesses	X	X	✓	X	✓
Taille du véhicule	✓	✓	✓	X	✓
Nombre de portes	X	X	✓	X	✓
Puissance du moteur	X	✓	✓	✓	✓
Accélération	X	X	✓	X	✓
Poids	X	✓	X	X	✓
Fiabilité	X	✓	X	X	X
Air conditionné	X	X	✓	X	✓
Jantes en alliage	X	X	✓	X	✓
Anti-lock Braking System (ABS)	X	X	✓	X	✓

(1) Cambridge Econometrics/BMRB (2008); (2) Train et Winston (2007); (3) Eftec (2008); (4) Train (1986)

Comme le montre le Tableau 5.3, la série de données de JATO est une source d'informations extrêmement riche sur les caractéristiques des véhicules, couvrant toutes les variables utilisées dans les études précédentes, exceptée la fiabilité, une mesure utilisée dans Train et Winston (2007). La même source de données a été utilisée par Cambridge Econometrics/BMRB (2008) et Eftec (2008). Les mêmes variables que pour ces deux précédentes études sont donc disponibles pour cette étude. Concernant la variable fiabilité, l'information aurait pu être utile, si elle avait été disponible, mais son absence n'aura que peu d'impact sur l'étude.

5.1.2.1. Prix des véhicules d'occasion

Les prix des véhicules neufs en 2006 sont disponibles à partir des données JATO qui ne prend pas en compte les véhicules d'occasion. La source des données des prix des véhicules d'occasion est l'Argus. Nous utilisons des données du premier trimestre 2006. Celles-ci fournissent les prix pour chaque millésime de véhicules entre 1997 et 2005, ainsi que pour les véhicules ayant entre 10 et 15 ans. Les prix sont donnés selon des combinaisons de versions de marques et de modèle pour des véhicules en moyen état.

5.1.2.2. Données sur les ventes de véhicules

Lorsque les données de l'INSEE donnent lieu à un choix ambigu de véhicule, nous utilisons les données de AAA sur les ventes de véhicules afin de pondérer l'observation possible. Ces données sont disponibles pour la période 2003-2009.

5.1.3. Données relatives au marché

Le dernier type de données nécessaires pour l'étude est l'information relative au marché, concernant les droits d'accises et sur les prix des carburants. Ces informations se combinent avec les caractéristiques des véhicules afin de déterminer leurs coûts de fonctionnement. Les taux de droits d'accises proviennent du magazine Autoplus, No. 908. Les prix des carburants proviennent eux du Ministère de l'Économie, des Finances et de l'Industrie, DGEMP (http://www.industrie.gouv.fr/energie/petrole/se_cons_fr.htm).

Tableau 5.4

Variables Relatives au Marché Utilisées dans des Etudes Précédentes et Disponibles dans l'Etude Présente

Variable	CE/BMRB ¹	T+W ²	Eftec ³	Train ⁴	Disponible pour cette étude
Droit d'accise du véhicule	✓	X	✓	X	✓
Prix des carburants	✓	X	✓	✓	✓
Champ du réseau de distribution du fabricant	X	✓	X	X	X

(1) *Cambridge Econometrics/BMRB (2008)*; (2) *Train et Winston (2007)*; (3) *Eftec (2008)*; (4) *Train (1986)*

La seule taxe appliquée au début de l'année 2006 était la taxe d'immatriculation (sur la carte grise). Cette taxe, appliquée sur les véhicules neufs et d'occasion au moment de l'immatriculation, varie selon la région et la puissance en chevaux fiscaux du véhicule. Nous avons agrégé ces taxes pour les 22 régions françaises en huit zones d'études et d'aménagement du territoire puis pondéré ces taxes en utilisant les données sur la population en 2006, afin de correspondre aux définitions géographiques du set de l'INSEE.

5.2. Traitement des données

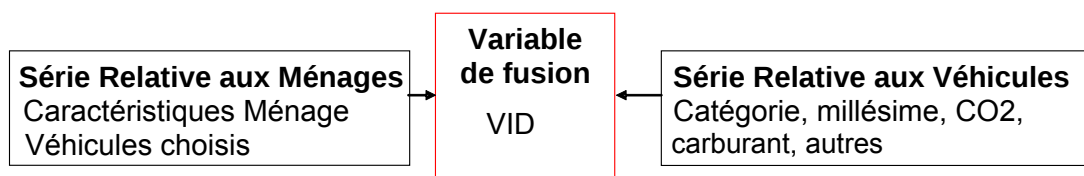
Les données utilisées pour l'analyse proviennent de six sources principales, décrites ci-dessous :

- INSEE
- JATO
- L'Argus Auto
- AAA
- Les prix des carburants proviennent du MEDDTL

Données sur la taxe d'immatriculation provenant d'Auto Plus

Deux séries de données ont été créées à partir de ces sources – une série comprenant des données relatives aux ménages, et une série sur les données relatives au type de véhicule - avec une correspondance directe entre eux. La Graphique 5.1 illustre la relation entre ces deux séries de données.

Graphique 5.1
Lien entre les Séries de données Relatifs au Ménage et aux Types de Véhicules



La série de données relatif aux ménages provient de l'INSEE, auquel nous avons ajouté une variable permettant d'identifier le type de véhicule (VID) de chaque véhicule détenu. Il existe plusieurs occurrences du même type de véhicule dans la série de données relatives aux ménages. La série de données relatives au type de véhicules, au contraire, contient une observation pour chaque type de véhicule sur le marché en 2006. Chaque observation dans cette série de données contient le nombre de véhicules de même type sur le marché en janvier 2006, ainsi que la moyenne et la variance de toutes les caractéristiques importantes présentes dans les séries de données de l'ARGUS et de JATO (par exemple le prix, la puissance du moteur, la taille du véhicule...).

Les types de véhicules sont définis comme combinaison unique de catégorie, millésime, type de carburant et fourchette d'émissions de CO₂. Le Tableau 5.5 montre les niveaux potentiels de chacune de ces variables qui sont choisies pour définir les types de véhicules dans notre analyse.

Tableau 5.5
Variables Définissant les "Types de Véhicules"

Catégorie	Millésime	Type de carburant	Emission de CO ₂ (g/km)
A – Voitures Utilitaires / de ville	Nouveau (moins d'1 an)	Essence	Moins de 100
B - Petite	Age moyen (entre 1 et 5 ans)	Diesel	101-120
C – Moyenne-basse	Agé (Plus de 5 ans)	Autre	121-130
D – Moyenne-haute			131-160
E – Grande, Exécutive et de luxe			161-165
Mini MPV			166-200
MPV			201-250
Sports			Plus de 250
SUV			

Source: NERA

Des exemples de véhicules appartenant à chaque catégorie sont données dans le Tableau 5.6 ci-dessous.

Tableau 5.6
Exemple de Véhicules par Catégorie

Catégorie	Exemple de véhicules
A	Renault Twingo, Ford Ka
B	Renault Clio, Ford Fiesta, Citroen Saxo/AX, Peugeot 106/206
C	Renault Megane, Ford Focus, Citroen Xsara/ZX, Peugeot 306/307/309
D	Renault Laguna, Ford Mondeo, Citroen Xantia, Peugeot 406
E	Renault Safrane, Citroen XM, Peugeot 605, BMW 5 Series
Mini MPV	Renault Express/Kangoo/Scenic, Citroen Berlingo/C15/Picasso, Peugeot Partner
MPV	Renault Espace, Citroen Evasion, Peugeot 806
Sports	Porsche 911
SUV	BMW X5

Créer la correspondance entre chaque véhicule considéré par l'INSEE et son type correspondant n'est pas simple. La catégorie et l'intervalle d'émissions de CO₂ ne sont pas pris en compte par l'INSEE, et doivent donc être définis. De plus, les données qui sont enregistrées par l'INSEE – marque, modèle, millésime et type de carburant – sont parfois absentes, mal retranscrites ou erronées.

Le Tableau 5.7 ci-dessous montre la qualité de la série de données de l'INSEE, indiquant le nombre de véhicules ayant des données concernant la marque, le modèle, le millésime et le type de carburant. Sur les 13 414 véhicules enregistrés dans la série de données complet de l'INSEE, 10 199 ont des valeurs pour toutes ces données. Les autres observations ont chacune au moins une donnée manquante. Par exemple, plusieurs personnes interrogées ont donné des descriptions approximatives de leur véhicule, telle que « Audi » ou « Renault ». Il était difficile de lier une telle description à un modèle exact. Ces véhicules sont présentés dans Tableau 5.7 comme n'ayant pas de données sur leur modèle.

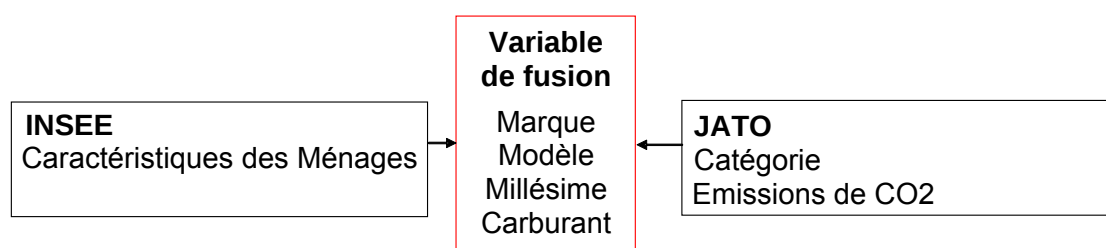
Tableau 5.7
Données de l'INSEE Relatives à la Qualité

Cas	Marque	Modèle	Millésime	Type de Carburant	Observations
1	✓	✓	✓	✓	10,199
2	✓	✓		✓	230
3	✓		✓	✓	1,355
4	✓		✓		1
5	✓			✓	38
6	✓				1
7			✓	✓	1,559
8				✓	30
9					1
					13,414

Source: Série de données de l'INSEE et analyse NERA

Pour les véhicules dans la série de données de l'INSEE disposant d'informations complètes sur la marque, le modèle, le millésime et le type de carburant, les données ont été fusionnées avec celles de JATO à l'aide de ces variables. Ceci permet d'identifier la catégorie et la fourchette d'émissions de CO₂ de chaque véhicule de la série de données de l'INSEE. Le processus de fusion est illustré dans la Graphique 5.2.

Graphique 5.2
Groupeement en Catégorie et Fourchette d'Emissions de CO₂ des Données de l'INSEE
Cas 1: Information Complète de l'INSEE sur la Marque, le Modèle, le Millésime et le Type de Carburant



Dans les cas où il manque des données dans la série de l'INSEE, nous avons élargi les données initiales afin de créer une nouvelle série de données pour chaque combinaison possible de catégorie, millésime, type de carburant et d'émissions de CO₂ étant donné les informations provenant des variables non-manquantes. En effet, nous avons assigné un poids à chaque observation simulée selon la distribution empirique d'observations similaires dans la série de données non-manquantes.

Pour ceci, nous avons adopté une approche qui implique de répliquer le ménage pour chacun des véhicules possibles lui correspondant. Nous avons ensuite donné à chacune de ces possibilités une pondération relative à sa probabilité empirique relatives aux ventes des véhicules de telle sorte que la somme des pondérations soit

égale à un pour chaque ménage.⁵¹ Cette approche nous permet d'assurer que le ménage soit inclus dans l'analyse, avec un poids de un, et que toutes les correspondances possibles (en prenant en compte les descriptions disponibles dans le sondage) soient considérées.

Certains véhicules de l'INSEE ne trouvent pas de liens pour d'autres raisons, par exemple parce que les combinaisons marque/modèle ne sont pas dans la base JATO, ou encore en cas d'informations manquantes. Concernant ces véhicules, certains se sont vu assignés une catégorie manuellement, d'autres ont été liés à une catégorie unique, et les observations restantes ont été élargies à tous les véhicules possibles et pondérées.⁵²

Afin de créer la série de données sur les véhicules, nous avons dans un premier temps fusionné les séries de données de l'Argus et de JATO par marque, modèle, millésime et type de carburant. La série de données JATO contenait presque tous les enregistrements présents dans celui de l'Argus, mais aussi un nombre important de données supplémentaires. Nous avons retenu seulement les enregistrements qui étaient dans les deux séries de données. Ceci assure la consistance entre les prix et les autres caractéristiques du véhicule, et limite les véhicules considérés à ceux qui ont été vendus en quantité suffisante pour mériter une cotation dans l'Argus. Nous considérons la présence dans l'Argus comme étant le critère le plus approprié pour définir l'étendue réelle du marché des voitures d'occasion en France au moment de l'étude.

Les moyennes et variances des caractéristiques des véhicules et le nombre de marque/modèle ont été calculées pour chaque combinaison de catégorie, millésime, type de carburant et de fourchette d'émissions de CO₂. Les caractéristiques moyennes ont été calculées tout d'abord pour chaque combinaison (marque/modèle/millésime/carburant) en considérant toutes les versions de véhicules. Ces caractéristiques moyennes ont ensuite été utilisées pour créer de nouveau une moyenne des caractéristiques par combinaisons de marque et de modèles en utilisant la catégorie de chaque véhicule. Il résulte de ceci l'identification de caractéristiques moyennes pour 211 types de véhicules (combinaisons de catégorie, millésime, type de carburant et émissions de CO₂).

Une fois que la série de données relatif au ménage a été créé, 36 véhicules appartenant à 16 ménages ont été retirés en raison de leur nature : ces 36 véhicules étaient des vans. Or la base de données JATO ne considère pas ce type de véhicule. 642 ménages ont également été enlevés car ils avaient plus de deux véhicules. 40 autres n'ont pas été considérés pour d'autres raisons. Ceci inclut le cas où les données n'étaient pas disponibles pour un certain type de véhicule (catégorie, millésime, type de carburant et émissions de CO₂) aussi bien dans les données de l'Argus que de JATO.

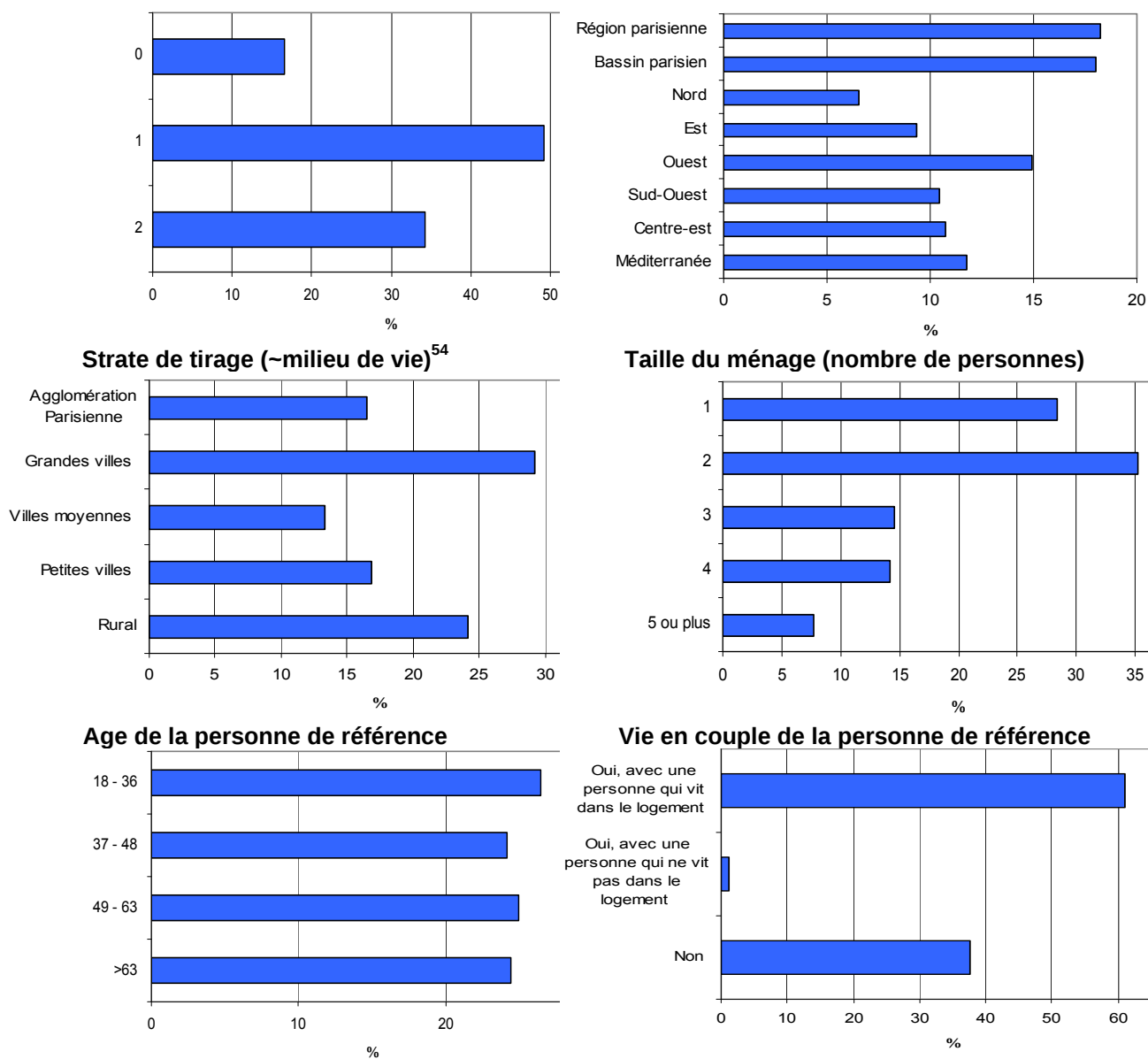
Propriétés des Séries de données Relatifs aux Ménages et aux Véhicules

La série de données finale relative aux ménages contient 9 542 ménages sur les 10 240 ménages d'origine. Les chiffres ci-dessous nous montrent la distribution des ménages dans l'échantillon selon certaines caractéristiques, incluant le nombre de véhicules détenus, la taille du ménage, leurs revenus et le niveau d'éducation.

⁵¹ Les données agrégées de vente sur le marché de AAA ont été utilisées, pour 2003 à 2006. Pour les millésimes plus récents, où les données de AAA n'étaient pas disponible, les données de l'INSEE ont été utilisées pour déterminer les poids.

⁵² Les informations sur les émissions de CO₂ de certains véhicules sont manquantes dans la base de données JATO. Dans ce cas, nous utilisons les données sur la consommation de carburant (si disponible) afin d'estimer les émissions de CO₂. Lorsque les données sur la consommation de carburant ne sont pas disponibles, nous avons utilisé les données, lorsque possible, du site: <http://www.vcacarfueldata.org.uk/search/usedcar/make-model.asp>

Graphique 5.3
Caractéristiques de l'Echantillon sur les Ménages
Nombre de véhicules possédés **Région de résidence du ménage**⁵³

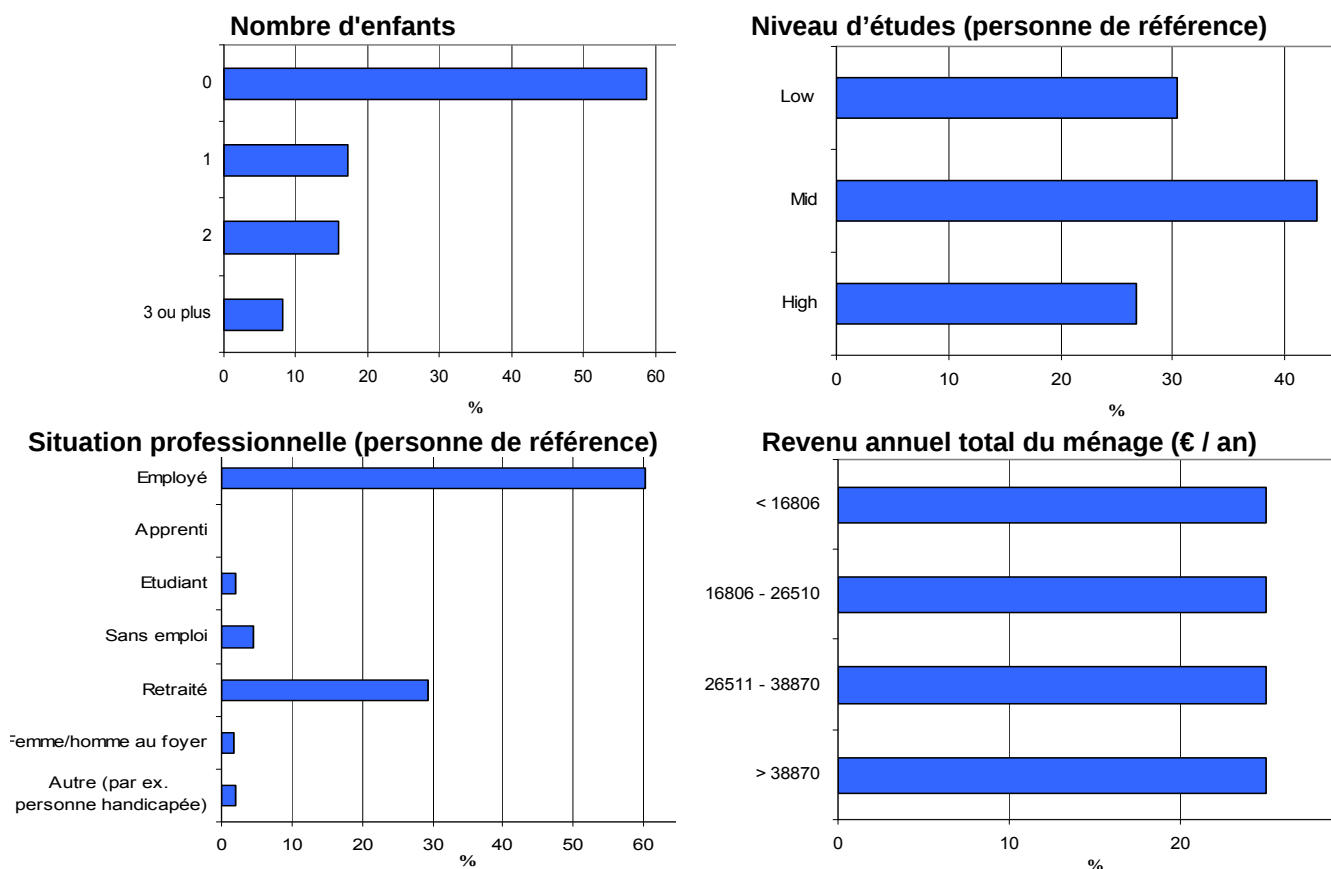


Source : Série de données de l'INSEE et analyses NERA (suite page suivante)

⁵³ Ces régions sont définies par le premier niveau de la Nomenclature of Territorial Units for Statistics (NUTS), développée par l'Union Européenne afin de référencer les divisions administratives des pays pour des raisons statistiques.

⁵⁴ Agglomération Parisienne réfère à l'unité urbaine de Paris, une zone définie par l'INSEE, composée d'environ 400 communes d'Île de France. Grandes villes sont les "communes appartenant à des unités urbaines de plus de 100 000 habitants". Villes moyennes étant les "communes appartenant à des unités urbaines de 20 000 à 100 000 habitants". Petites villes sont les "communes appartenant à des unités urbaines de moins de 20 000 habitants". Rural sont les communes rurales.

Graphique 5.4 (suite)
Caractéristiques de l'Echantillon sur les Ménages



Source : Série de données de l'INSEE et analyses NERA

L'échantillon original relatif aux ménages a été obtenu par l'INSEE de manière à être représentatif de la population française en 2005-2006. L'échantillon final utilisé pour notre analyse est également représentatif à l'exception du fait qu'il exclut les ménages possédant plus de deux véhicules et ceux qui sont propriétaires de vans ou d'autres types de véhicules assez rares.

Le Tableau 5.8 montre les caractéristiques moyennes des types de véhicules inclus dans la série de données, par catégorie, millésime, émissions de CO₂ et type de carburant. Le tableau nous montre également le prix moyen, la consommation de carburant et la capacité du coffre ainsi que le nombre de modèles pour chaque groupe de véhicules.

Tableau 5.8
Caractéristiques des Véhicules Disponibles en 2006

	Prix (000 €)	Consommation de carburant (litres par 100 km)	Capacité du Coffre (m³)	Nombre de modèles de véhicules disponibles
Catégorie				
A	4.3	5.6	0.2	143
B	5.5	5.8	0.3	361
C	6.7	6.7	0.4	495
D	9.5	7.4	0.5	517
E	15.0	9.2	0.5	326
MPV	10.3	8.5	0.6	309
MiniMPV	8.8	6.8	0.5	298
SUV	14.8	10.0	0.5	402
Sports	12.8	8.7	0.3	191
Millésime				
Moins de 1 an	18.6	7.1	0.4	297
1-5 ans	12.4	7.7	0.4	1444
Plus de 5 ans	5.0	7.8	0.4	1301
Emissions de CO2 (g/km)				
Moins de 100	5.2	3.4	0.2	7
101-120	6.7	4.4	0.2	94
121-130	6.2	4.9	0.3	79
131-160	7.1	5.8	0.3	717
161-165	7.0	6.5	0.4	126
166-200	8.9	7.3	0.4	903
201-250	11.2	9.0	0.5	749
Plus de 250	17.4	11.5	0.5	367
Type de carburant				
Essence	9.0	8.5	0.4	1666
Diesel	11.1	6.6	0.5	1335
Autre	5.4	9.3	0.4	41
TOTAL	9.8	7.7	0.4	3042

Source : Analyses NERA

5.2.1. Appréciation Globale de la Qualité des Données

La série de données construit pour l'analyse contient un échantillon large et représentatif des ménages français. Il y a cependant un certain nombre de questions relatives à ces données, signifiant qu'il faut interpréter les résultats avec précaution. Les problèmes les plus importants sont les suivants :

- Les données de l'INSEE ne donnent pas exactement le type de véhicule détenu.
- Au mieux, ces données permettent de connaître la marque, le modèle, le millésime, et le type de carburant du véhicule. Cependant nous ne connaissons jamais la version du modèle, et donc les émissions en

CO2 des véhicules détenus. Cela implique que nous sommes incapables de modéliser, par exemple, les changements de véhicule d'une version à une autre qui peut être due à une variation de prix.

– Au pire, les données sur les modèles des véhicules détenus sont manquantes. Par exemple, certains véhicules sont décrits comme étant simplement des "Audi" ou "Renault". La série de données est donc inévitablement moins robuste du fait de ce manque d'information.

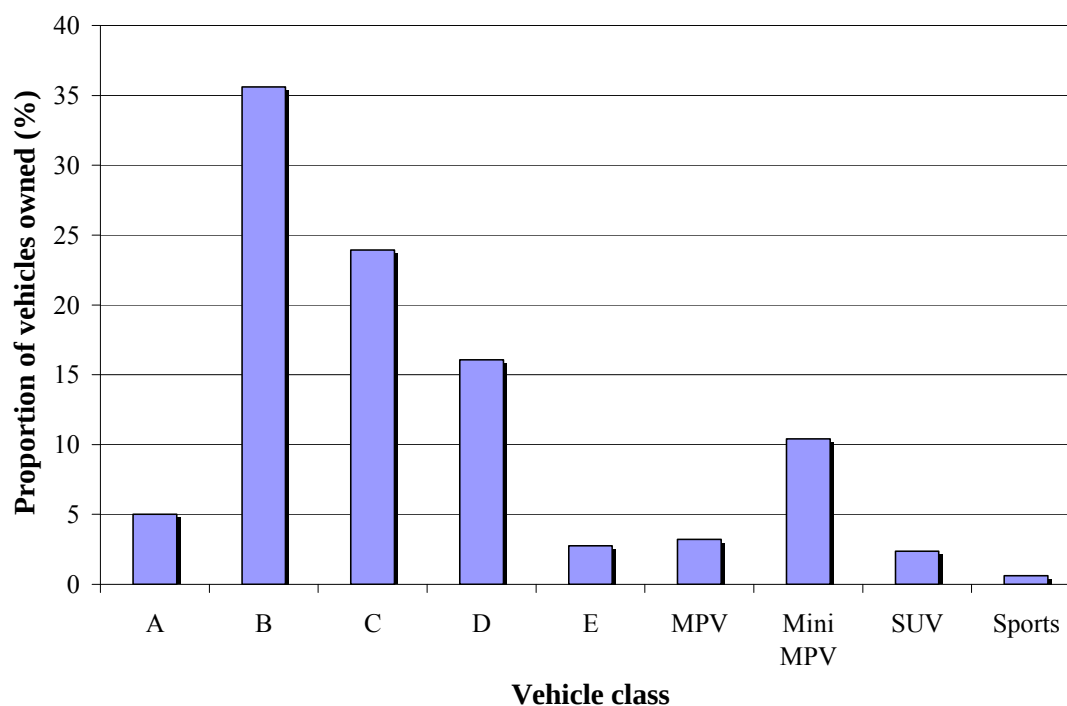
- Les données JATO et L'Argus sont incomplètes ;
 - Il n'y a pas de données JATO pour les véhicules de plus de 10 ans ;
 - Les données relatives émissions de CO2 sont manquantes pour les véhicules relativement anciens (5-10 ans) ;
 - JATO ne contient pas de données sur les vans;
 - Les véhicules peu vendus n'étaient pas bien représentés dans l'Argus (par exemple, les véhicules à carburant alternatif);
- Nous n'avons pas de données sur les véhicules précédemment détenus ne permettant pas à notre modèle de prendre en compte cette dynamique.

6. Résultats Descriptifs

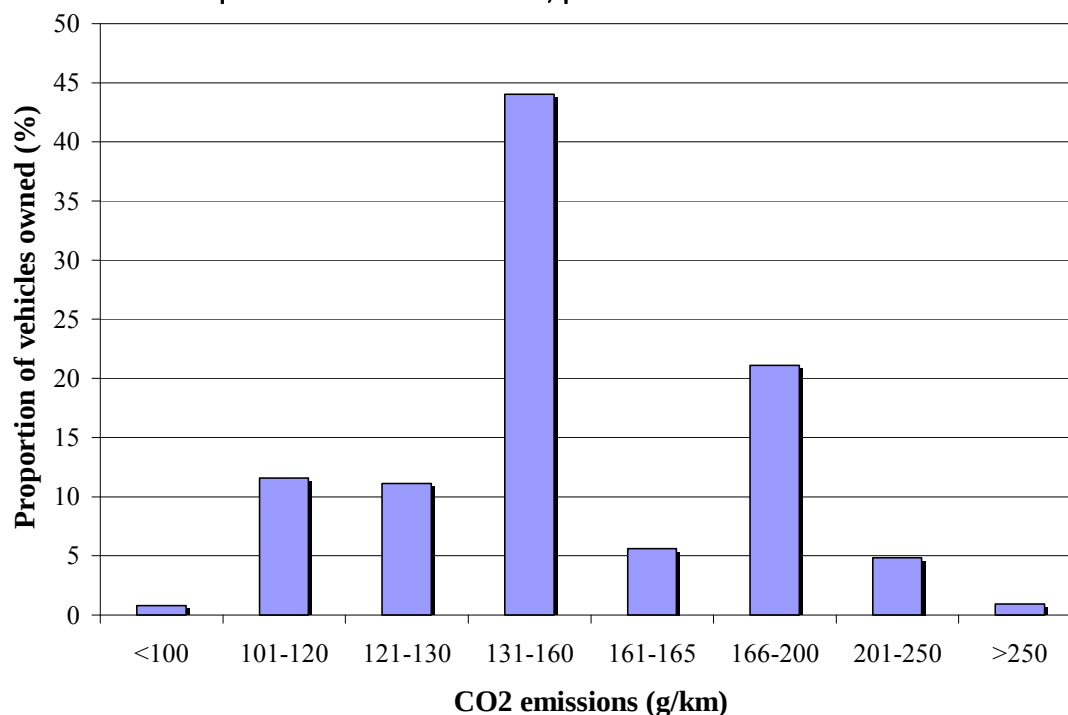
Dans cette section, nous exposons plusieurs représentations graphiques mettant en évidence la distribution des véhicules détenus selon un certain nombre de caractéristiques relatives aux véhicules ainsi qu'aux ménages.

Les trois représentations graphiques ci-dessous montrent la proportion de véhicules de chaque type qui sont détenus par les ménages français selon la catégorie du véhicule, ses émissions de CO₂, et son millésime. Elles montrent que les véhicules de catégories B et C sont les plus communes, que la plupart des véhicules émettent entre 131 et 160 g/km de CO₂, et que la majorité des véhicules détenus ont plus de cinq ans.

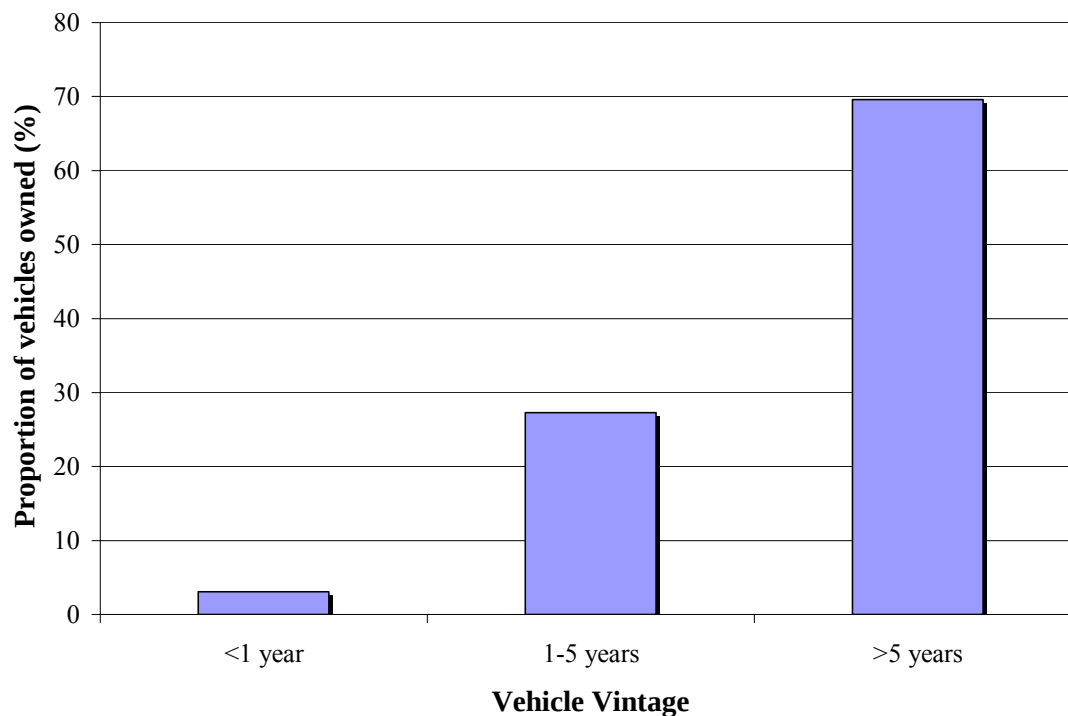
Graphique 6.1
Proportion de Véhicules Détenus, par Catégorie de Véhicule



Graphique 6.2
Proportion de Véhicules Détenus, par Classe d'Emissions de CO2



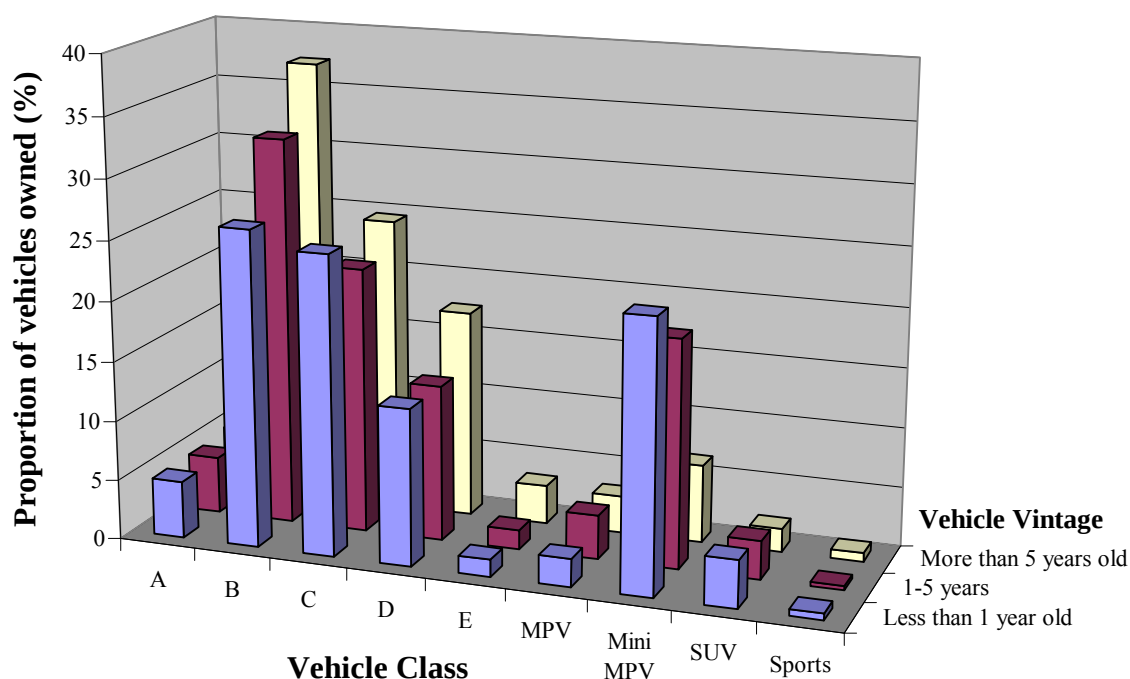
Graphique 6.3
Proportion de Véhicules Détenus, par Millésime de Véhicule



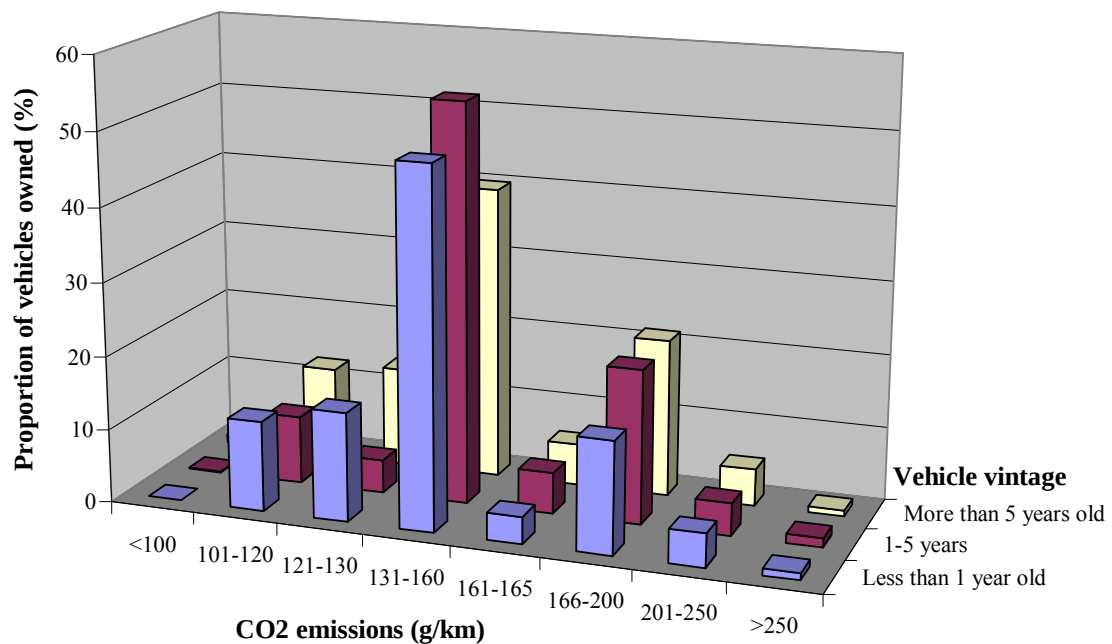
Les trois représentations graphiques ci-dessous permettent de décomposer la distribution de détention de véhicule selon différentes paires de caractéristiques des véhicules. Dans chaque représentation, la somme des proportions est égale à 100 % pour chaque ligne.

- Tout d'abord, le Graphique 6.4 montre qu'il y a une proportion plus grande de Mini MPVs dans les nouveaux véhicules achetés que dans les anciens. Il y a également relativement moins de véhicules de catégories B, D et E dans les nouveaux véhicules que dans les anciens.
- Le Graphique 6.5 montre que les nouveaux véhicules sont moins présents dans les classes hautes d'émissions de CO₂ et sont plus représentés dans les classes faibles que les anciens véhicules. Cependant ces différences sont assez faibles. Globalement, la distribution des véhicules selon la classe d'émissions de CO₂ est assez stable entre les trois millésimes.
- Le Graphique 6.6 met en avant la distribution de véhicules détenus par classe d'émissions de CO₂ et par catégorie. Cette représentation vérifie que les véhicules les plus imposants et les plus puissants sont ceux qui sont le plus représentés dans les classes hautes d'émissions de CO₂

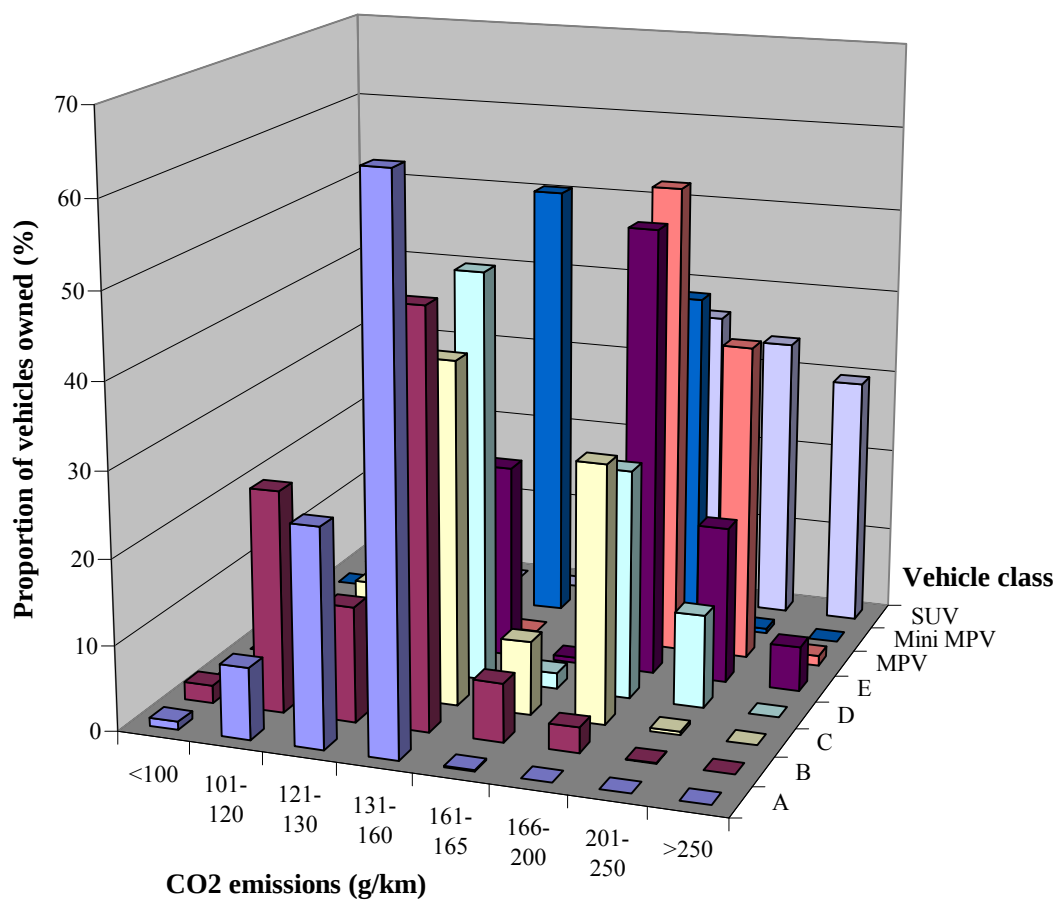
Graphique 6.4
Proportion de Véhicules Détenus par Catégorie et par Millésime



Graphique 6.5
Proportion de Véhicules Détenus par Classe d'Emission de CO2 et par Millésime

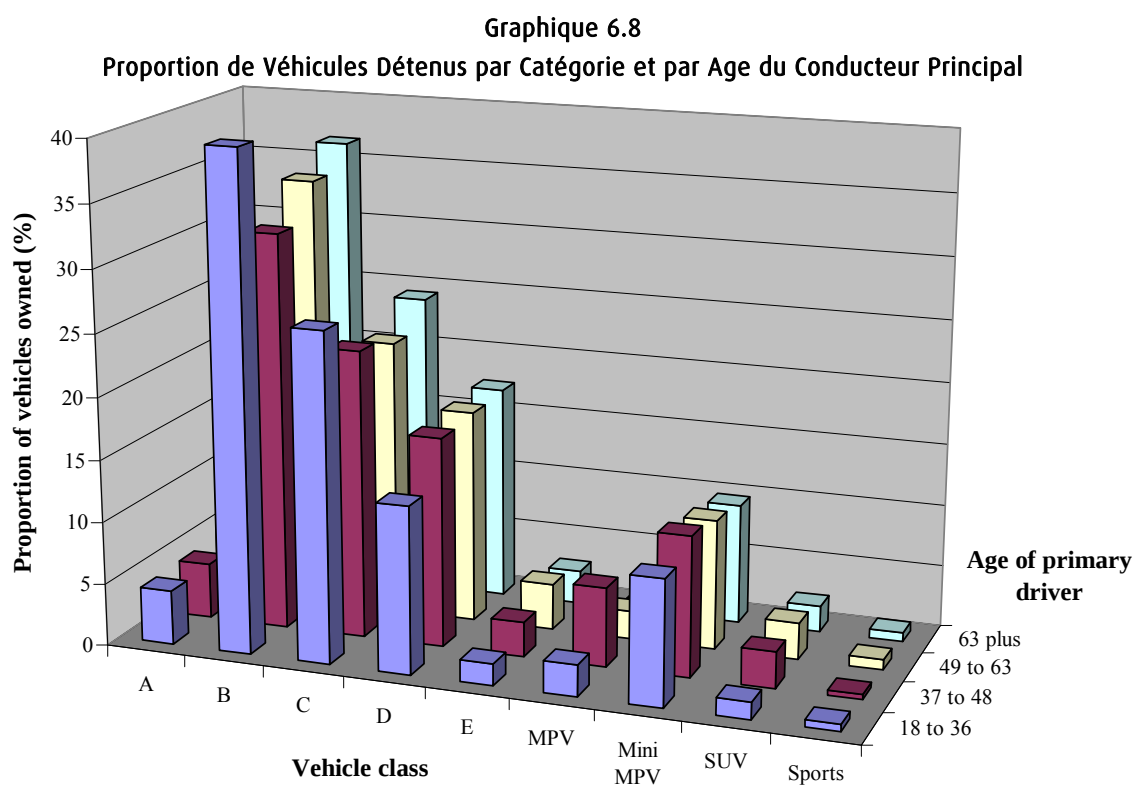
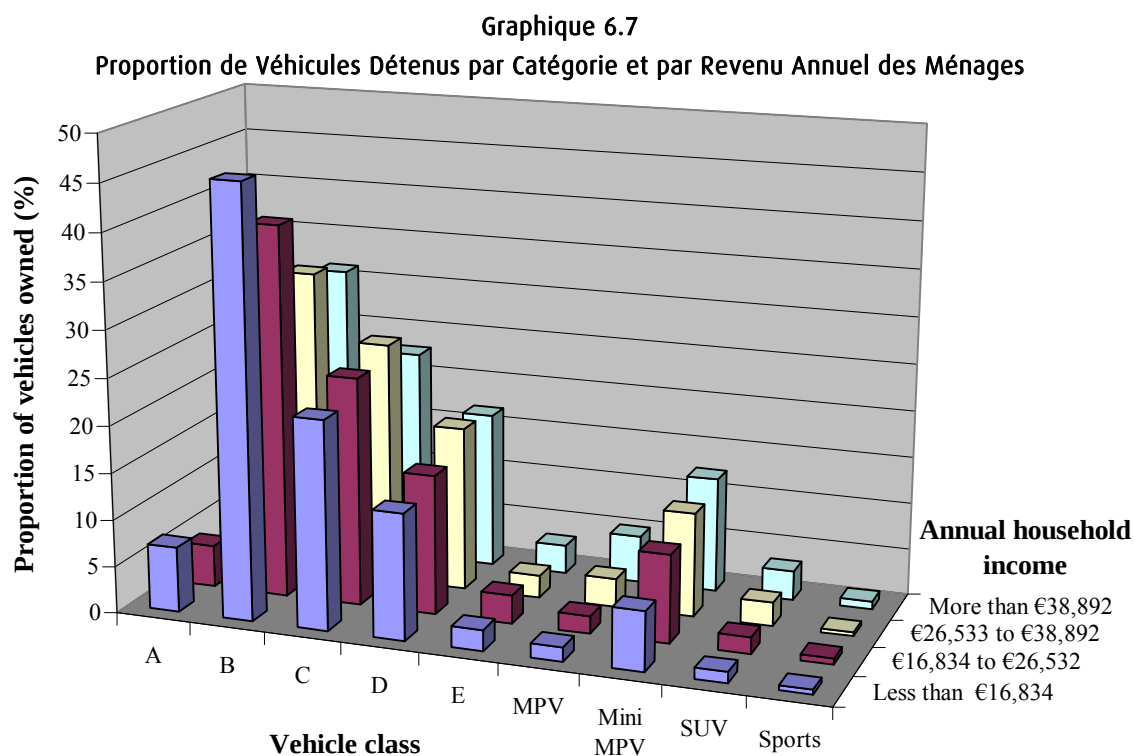


Graphique 6.6
Proportion de Véhicules Détenus par Classe d'Emission de CO2 et par Catégorie

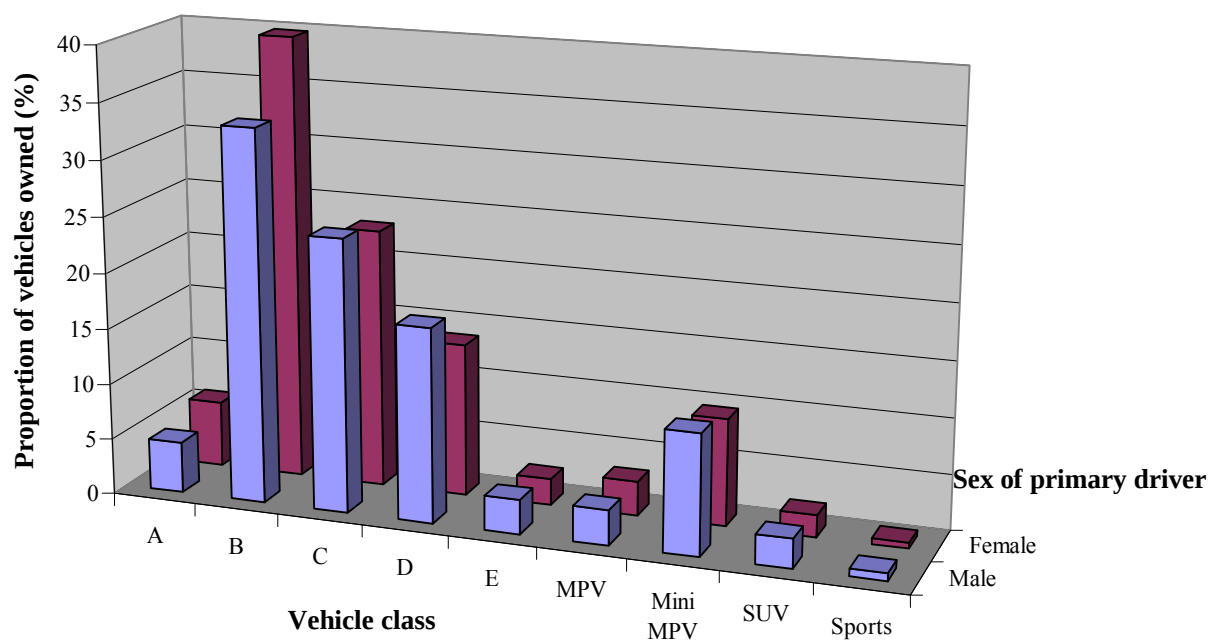


Les représentations graphiques ci-dessous mettent en évidence la proportion de véhicules détenue pour chaque catégorie de véhicules selon un éventail de caractéristiques des ménages. Les chiffres sont établis de sorte que la somme des proportions soit égale à 100 % à travers chaque catégorie de véhicules, permettant ainsi de gommer les différences dans la fréquence relative de chaque type de ménage dans l'échantillon. Ainsi, sans surprise, les principales caractéristiques peuvent être observées :

- Le Graphique 6.7 met en avant le fait que les ménages ayant les revenus les plus élevés ont tendance à acquérir des véhicules de catégorie E, MPV, Mini MPV et Sports davantage que les autres ménages. Les ménages ayant les revenus les plus faibles ont tendance à détenir une part relative plus importante de véhicules de catégorie A et B. Les ménages ayant des revenus moyens ont tendance à acquérir des véhicules de catégories moyennes.
- Le Graphique 6.8 met en avant que les conducteurs les plus jeunes (entre 18 et 36 ans) ont tendance à détenir relativement plus de véhicules de catégories A, B et C et moins de catégorie D, E et SUV. Les ménages compris dans le groupe des 37-48 ans ont tendance à préférer les MPV aux autres groupes d'âge, au détriment des catégories B et C.
- Le
-
- Graphique 6.9 montre que les femmes ont une préférence relativement plus forte pour les véhicules de catégories A et B que les hommes. Cependant, la distribution des catégories de véhicules est sensiblement la même entre les hommes et les femmes.
- Le Graphique 6.10 montre que les MPVs, et dans une moindre mesure, les mini MPVs, sont préférés relativement plus souvent par les ménages ayant le plus d'enfants. Cependant, les ménages sans enfants ont une préférence relative pour les véhicules de catégories A et B que les ménages avec enfants.
- Enfin, le Graphique 6.11 met en avant que les ménages localisés dans les plus grandes villes, à l'exception de l'agglomération Parisienne, ont tendance à détenir relativement plus de véhicules de catégorie B et C, et relativement moins de catégorie D, MPV et mini MPV. Paris semble être une exception à ces caractéristiques, apparaissant plus en ligne avec les zones rurales et les petites villes que les grandes villes dont elle fait partie.

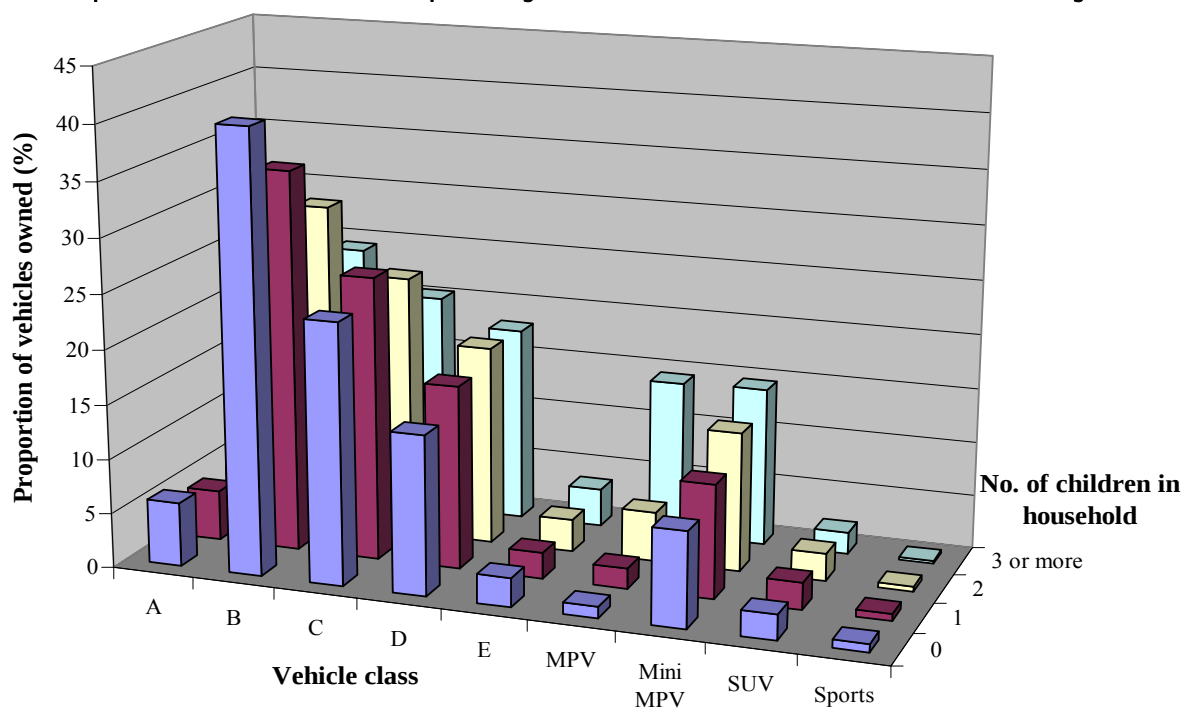


Graphique 6.9
Proportion de Véhicules Détenus par Catégorie et selon le Sexe du Conducteur Principal



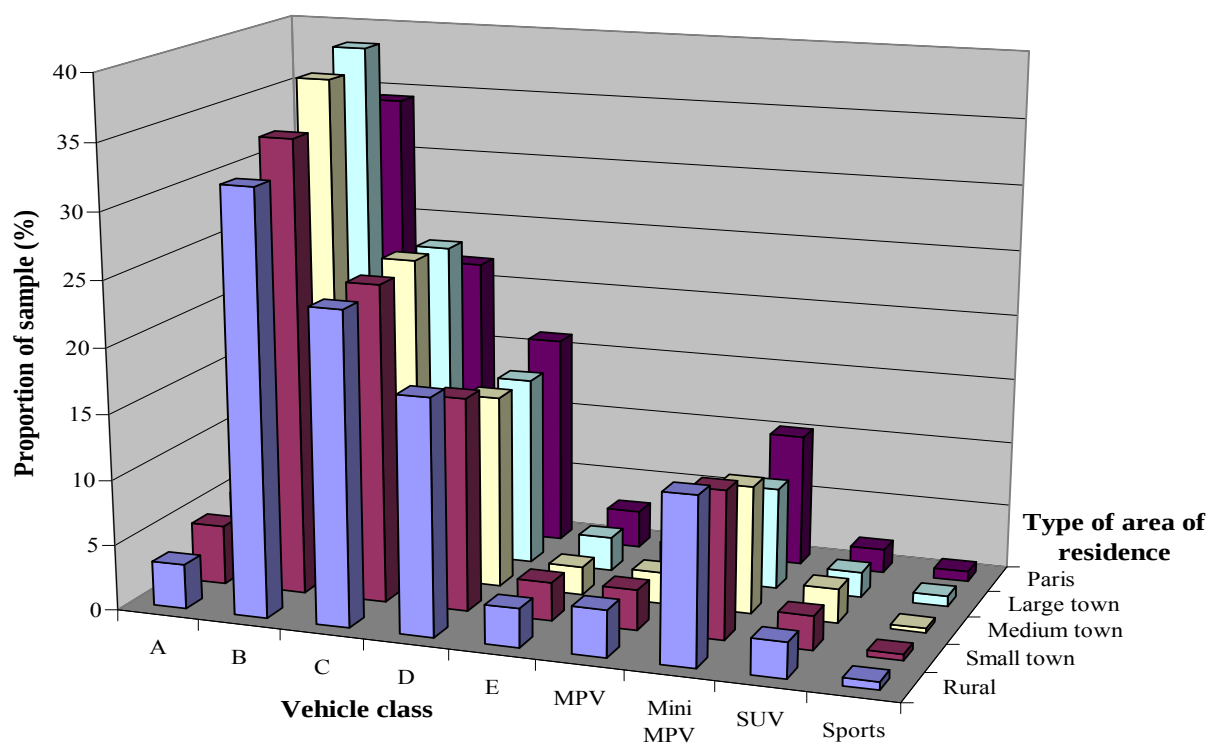
Graphique 6.10

Proportion de Véhicules Détenus par Catégorie et selon le Nombre d'Enfants dans le Ménage



Graphique 6.11

Proportion de Véhicules Détenus par Catégorie et par Type de Régions de Résidence



7. Résultats Econométriques

Le modèle complet d'acquisition de véhicules tel qu'il a été décrit en Section 4 comprend trois sous-modèles de choix :

- le nombre de véhicules choisis par le ménage soit zéro, un ou deux ;
- le choix du type de véhicule à acquérir, selon si l'option de détenir un véhicule est choisie ; et
- le choix de la paire de type de véhicule à acquérir, selon si l'option de détenir deux véhicules est choisie.

Cette section décrit les résultats de l'estimation de ces sous-modèles. Nous commençons par les sous modèles de choix du type de véhicules selon si le ménage désire un ou deux véhicules, avant d'expliquer le modèle de choix du nombre de véhicule à acquérir. Cet ordre de présentation reflète la séquence requise pour l'estimation.

7.1. Sous-Modèle de Choix de Véhicules pour les Ménages à Un Véhicule

Le sous-modèle de choix du type de véhicule permet d'estimer la probabilité pour chaque ménage de choisir un véhicule de chaque type, étant donné qu'ils ont choisi d'en posséder un seul. Il est estimé seulement sur l'échantillon des ménages qui possèdent un seul véhicule.

Le nombre d'alternatives qui s'offrent à chaque ménage dans le modèle est de 211, une par type de véhicule. Afin d'accélérer le calcul nécessaire pour l'estimation, le nombre d'alternatives réellement inclus dans l'estimation était de 50. Cela comprenait la solution choisie, et 49 autres choisies au hasard, sans remise, dans les 210 autres types de véhicules. Cette procédure d'estimation sur un sous-ensemble aléatoire se traduit par des résultats cohérents (Train, 1986).

Les variables présentes dans le modèle de choix d'un seul véhicule, ainsi que leurs coefficients associés, standard errors et t-statistics, sont exposés dans le Tableau 7.1. La somme des coefficients multipliés par les valeurs des variables leur correspondant donne une représentation de l'utilité pour la $n^{\text{ième}}$ alternative, V_n . Étant donné les utilités représentant les N alternatives, les probabilités sont alors égales à :

$$P_n = \frac{\exp(V_n)}{\sum_{i=1}^N \exp(V_i)}$$

Le modèle s'adapte raisonnablement bien aux données. Toutes les variables ressortent avec le signe attendu et sont statistiquement significatives.

Les quatre premières variables dans le modèle reflètent l'effet du prix d'achat sur l'utilité d'une alternative pour les ménages, ceci dans les quatre intervalles de revenu de l'échantillon. Comme prévu, les quatre coefficients du prix d'achat sont négatifs, ce qui indique que les gens préfèrent les véhicules moins chers à des véhicules plus coûteux, *ceteris paribus*. En outre, plus le revenu des ménages est élevé, moins les coefficients sont négatifs. Cela indique que le coût du véhicule est une préoccupation plus importante pour les ménages les moins aisés financièrement que pour les ménages les plus riches. Autrement dit, les véhicules à coût élevé sont relativement moins attrayants pour les ménages les moins aisés que pour les ménages les plus riches, comme attendu.

Les variables 5, 6 et 7 permettent de saisir l'impact des coûts variables de l'automobile sur l'utilité des alternatives entre type de véhicule. Ces variables ont également été séparées en différents groupes de revenus, mais dans ce cas, les deux intervalles de revenu les plus bas ont été combinés (il n'y avait pas de différence significative entre les coefficients des variables de coûts pour les deux intervalles inférieurs de revenu.). Les coefficients de ces variables sont tous négatifs, ce qui indique que les ménages préfèrent les véhicules qui sont moins chers à l'entretien que les véhicules plus coûteux, *ceteris paribus*. Comme précédemment, les coefficients sont plus négatifs pour les groupes à faible revenu que pour les groupes à revenu élevé, ce qui indique que les coûts variables des différents types de véhicules sont plus préoccupants pour les ménages les moins aisés financièrement que pour les ménages les plus riches, comme on pouvait s'y attendre.

Tableau 7.1
Sous-Modèle de Choix de Véhicules pour les Ménages Ayant un Véhicule

Variables Explicatives		Coefficients Estimés	Standard Error	t-statistic
1	Prix d'achat (€'000) pour les ménages ayant des revenus inférieurs à €16 834	-0.4048	0.0165	-24.55
2	Prix d'achat (€'000) pour les ménages ayant des revenus compris entre €16 834 et €26 532	-0.2777	0.0132	-21.04
3	Prix d'achat (€'000) pour les ménages ayant des revenus compris entre €16 833 et €38 892	-0.1975	0.0124	-15.95
4	Prix d'achat (€'000) pour les ménages ayant des revenus supérieurs à €38 892	-0.1195	0.0116	-10.34
5	Coûts variables (€/100km) pour les ménages ayant des revenus inférieurs ou égaux à €26 532	-0.2839	0.0130	-21.76
6	Coûts variables (€/100km) pour les ménages ayant des revenus compris entre €26 533 et €38 892	-0.2559	0.0177	-14.43
7	Coûts variables (€/100km) pour les ménages ayant des revenus supérieurs à €38 892	-0.1793	0.0186	-9.65
8	Contenance du coffre (m³) pour les ménages composés d'au moins quatre membres	4.0370	0.3857	10.47
9	Log du nombre de marques et de modèles de type de véhicule (en proportion de toutes les marques et modèles)	1.0459	0.0168	62.18
10	Variance du prix autour de sa moyenne pour le type de véhicule	0.0131	0.0020	6.69
11	Dummy nouveau millésime	0.3991	0.1233	3.24
12	Dummy millésime âgé de 1 à 5 ans	0.1768	0.0567	3.12
13	Dummy MPV	-1.8591	0.1773	-10.49
14	Dummy MPV pour les ménages composés d'au moins quatre membres	0.9590	0.2514	3.82
15	Dummy voiture de sport	-2.4931	0.3003	-8.30
16	Dummy SUV	-0.8590	0.1207	-7.11

Source: NERA

Modèle estimé par logit conditionnel sous Stata. Nombre d'observations : 505,900; Log Vraisemblance à convergence : 13,148.83.

Le ratio entre les coefficients des coûts variables et du prix d'achat a une signification économique du fait qu'il mesure le montant qu'un ménage est prêt à payer en plus du prix d'achat, à la réduction d'une unité de coût de l'automobile. Cette relation peut être utilisée comme un test de la validité du modèle, en regardant si l'ampleur de la propension implicite à payer lors d'une baisse de prix est raisonnable, compte tenu de certaines hypothèses habituelles sur la durée de vie du véhicule, le nombre de kilomètre parcourus et le taux d'actualisation.

Le Tableau 7.2 présente les résultats d'une telle analyse. Il montre le seuil de rentabilité des distances qui auraient besoin d'être parcourues par le véhicule pour que la propension implicite à payer soit rationnelle. Il montre les distances à parcourir pour atteindre le seuil de rentabilité pour les quatre intervalles de revenu dans l'échantillon selon deux hypothèses de taux d'actualisation, 3 % et 5 %, et deux hypothèses de durée de vie du véhicule, 8 ans et 10 ans.

Les résultats semblent raisonnables. Les ménages à faible revenu doivent conduire entre 8 200 km et 10 800 km par an pour justifier leur propension implicite à payer pour des coûts variables réduits. Les ménages à revenu

élevé devraient conduire entre 17 600 km et 23 200 km pour justifier leurs dépenses sur des véhicules plus efficaces énergétiquement.

Tableau 7.2
Seuil de Rentabilité du Véhicule selon les Distances Parcourues Annuellement, Induit par les Coefficients
Estimés du Prix et du Coût Variable

Taux d'actualisation	Vie du véhicule (années)	Seuil de rentabilité induit du véhicule selon la distance parcourue ('000 km / an)			
		Revenus inférieurs à €16,834	Revenus entre €16,834 et €26,532	Revenus entre €16,833 et €38,892	Revenus supérieurs à €38,892
3 %	8	10.0	14.6	18.5	21.4
3 %	10	8.2	12.0	15.2	17.6
5 %	8	10.8	15.8	20.1	23.2
5 %	10	9.1	13.2	16.8	19.4

Source: Analyses NERA

La huitième variable du Tableau 7.1 reflète l'importance de la capacité du coffre pour les ménages composés d'au moins quatre membres. Les plus petits ménages se sont révélés n'avoir pas de propension à payer significative pour des capacités de coffre élevées.

Les neuvième et dixième variables sont incluses afin d'approximer l'utilité disponible lors du choix de la marque/modèle, une fois qu'un type de véhicule a été choisi. Cette approximation permet d'éviter la nécessité d'estimer un sous-modèle de choix de marque/modèle de façon explicite.

La neuvième variable est le logarithme du nombre de marques et de modèles par type de véhicule, en tant que proportion de toutes les marques et modèles. Cette variable ressort avec un coefficient positif, ce qui indique que les ménages préfèrent les types de véhicules où il y a davantage de marques et de modèles dans lesquels ils peuvent choisir. Idéalement, le coefficient devrait prendre une valeur comprise entre zéro et un (McFadden, 1978). Le fait qu'il soit de 1 046, soit supérieur à un, signifie que le modèle n'est pas totalement compatible pour maximiser l'utilité à travers une gamme complète de choix.

La dixième variable mesure la variance du prix des véhicules autour de leur moyenne par type de véhicule. Elle ressort également avec un coefficient positif, tel que requis par la théorie, indiquant que les ménages préfèrent les types de véhicules où il existe un plus grand éventail de prix entre ses marques/modèles.

Les variables 11 à 16 dans le Tableau 7.1 sont toutes des dummies (variables indicatrices), capturant l'utilité résiduelle moyenne d'une série de millésimes et de catégorie, non pris en compte par les variables des caractéristiques du véhicule. Ainsi, les variables 11 et 12 montrent que les véhicules neufs sont plus appréciés que ceux ayant entre 1 et 5 ans, qui sont eux-mêmes plus appréciés que les véhicules plus âgés (omis), *ceteris paribus*.

Les variables 13 et 14 montrent que la dummy MPV (monospace) a un coefficient négatif, indiquant que les monospaces sont moins attrayants que les autres véhicules *ceteris paribus*. Le coefficient de la dummy MPV pour les ménages composés d'au moins quatre membres est positif, indiquant que les monospaces sont relativement plus attractifs pour ce type de ménages que pour les autres. Toutefois, le coefficient de cette variable est plus faible en valeur absolue que la dummy MPV pour tous les ménages, indiquant que les ménages composés d'au moins quatre membres préfèrent généralement les autres voitures aux MPV, *ceteris paribus*.

Enfin, les coefficients des dummies pour les voitures de sport et les SUV (véhicule utilitaire) sont tous deux négatifs, indiquant que ces types de véhicules sont moins attrayants que les autres types de véhicules, *ceteris paribus*.

Dans le cadre de la spécification du modèle à un véhicule, quelques variables ont été testées et enlevées⁵⁵. Dans ces variables figurent la puissance du moteur, l'accélération, la vitesse maximale et la capacité du véhicule en terme de places, l'âge, l'éducation et la situation d'emploi du conducteur primaire, et s'il vit en couple, le nombre d'enfants et nombre de personnes employées composant son ménage ainsi que la région du ménage et le type de zone de résidence.

7.2. Sous Modèle de Choix de Véhicules pour les Ménages à Deux Véhicules

Le choix d'une paire de types de véhicules est modélisé à l'aide d'une approche similaire à celle pour les ménages dans l'alternative de posséder un véhicule. Le sous-modèle estime la probabilité de choisir une paire de types de véhicules parmi toutes les paires possibles. Il existe 211 types de véhicules, et donc 22,366 paires possibles y compris les paires composées de type de véhicule identique. Tout comme pour le sous-modèle du choix d'un véhicule, le nombre d'alternatives réellement inclus dans l'estimation est de 50. Cela comprend la solution choisie, et 49 autres choisies au hasard, sans remise.

Le calcul des probabilités à partir du sous-modèle de choix de deux véhicules est le même que pour le sous-modèle de choix d'un véhicule. L'équation estimée génère l'utilité V_n , pour chaque paire de types de véhicules. La probabilité de choisir la $n^{\text{ième}}$ paire de type de véhicule est alors égale à :

$$P_n = \frac{\exp(V_n)}{\sum_{i=1}^N \exp(V_i)}.$$

Le Tableau 7.3 présente l'estimation du sous-modèle pour les ménages ayant deux véhicules, montrant les variables qui contribuent à l'utilité, ainsi que leurs coefficients, leurs standard errors et t-statistics.

L'ensemble des variables du modèle est similaire à celui du modèle de choix d'un véhicule. Dans l'ensemble, les variables du modèle correspondent à des sommes des deux types de véhicules dans la paire. Par exemple, la première variable est la somme des prix d'achat des deux types de véhicules (en milliers d'euros) pour les ménages dont le revenu est inférieur à 16 834€. Cette spécification suppose que les ménages accordent la même importance à chaque véhicule de la paire.

Les coefficients des variables composant le modèle ont tous les signes attendus, ont des importances relatives plausibles, et sont statistiquement significatifs. En outre, les résultats de ce modèle sont très similaires à ceux qui découlent du sous-modèle de choix d'un véhicule. Les ménages sont sensibles au prix d'achat et aux coûts variables de l'automobile, et sont d'autant plus sensibles que leurs revenus sont faibles. Ils ont une capacité de coffre plus grande, ce qui est plus important pour les ménages composés de quatre membres ou plus. Ils sont plus attirés par les millésimes plus récents et moins par les MPV, les SUV et les voitures de sport, *ceteris paribus*. Les MPV sont à nouveau relativement plus attrayants pour les ménages de quatre membres ou plus.

⁵⁵ Notre décision de savoir si, et comment, inclure une variable candidate dans le modèle a été basée non seulement sur la signification statistique de la variable, bien que cela ait joué un rôle important, mais plutôt sur une appréciation équilibrée de l'opportunité d'inclure la variable afin d'améliorer la performance de l'échantillon, et le cas échéant, dans quelle mesure l'amélioration justifierait la complexité supplémentaire causée par l'inclusion de la variable.

Tableau 7.3

Sous-Modèle du Choix du Type de Véhicules pour les Ménages à Deux Véhicules

Variables Explicatives		Coefficients Estimés	Standard Error	t-statistic
1	Somme des prix d'achat des deux véhicules (€'000) pour les ménages ayant des revenus inférieurs à €16 834	-0.3057	0.0241	-12.69
2	Sommes des prix d'achat des deux véhicules (€'000) pour les ménages ayant des revenus compris entre €16 834 et € 26 532	-0.2642	0.0129	-20.50
3	Sommes des prix d'achat des deux véhicules (€'000) pour les ménages ayant des revenus compris entre €16 833 et € 38 892	-0.2021	0.0077	-26.31
4	Sommes des prix d'achat des deux véhicules (€'000) pour les ménages ayant des revenus supérieurs à € 38 892	-0.1107	0.0054	-20.66
5	Somme des coûts variables des deux véhicules (€/100km)	-0.3162	0.0093	-33.99
6	Somme des contenances du coffre des deux véhicules (m³) pour les ménages composés d'au moins quatre membres	1.3787	0.2186	6.31
7	Somme des contenances du coffre des deux véhicules (m³) pour les ménages composés de trois membres ou moins	1.0502	0.1708	6.15
8	Différence absolue estimée dans les contenances des coffres des deux véhicules (m³) pour les ménages composés d'au moins quatre membres	1.4931	0.3003	4.97
9	Différence absolue estimée dans les contenances des coffres des deux véhicules (m³) pour les ménages composés de trois membres ou moins	0.6637	0.2649	2.51
10	Log du nombre de paire de marques et de modèles dans la paire du type de véhicule (en proportion de toutes paires possibles de marques et de modèles)	0.9749	0.0151	64.57
11	Nombre de véhicules du nouveau millésime dans la paire	0.2525	0.1705	1.48
13	Nombre de MPV dans la paire	-1.6320	0.1484	-11.00
14	Nombre de MPV dans la paire pour les ménages composés d'au moins quatre membres	1.3466	0.1898	7.10
15	Nombre de SUV dans la paire	-0.2943	0.0825	-3.57
16	Nombre de voitures de sport dans la paire	-1.3067	0.1533	-8.52

Source : NERA

Modèle estimé par logit conditionnel sous Stata. Nombre d'observations : 1,176,850 ; Log Vraisemblance à convergence: 6,976.28.

Comme dans le cas du sous-modèle de choix d'un véhicule, un certain nombre de variables ont été testées lors du développement de la spécification du sous-modèle de choix de deux véhicules. Les suivantes ont été jugées insignifiantes dans le choix de la paire de véhicules dans le modèle de choix de deux véhicules, par rapport aux variables finalement incluses dans le modèle: Puissance du moteur, accélération, vitesse maximale et capacité des véhicules en terme du nombre de places, l'âge, l'éducation et situation professionnelle de la personne de référence du ménage, si cette personne vit dans en couple, le nombre d'enfants et nombre d'employés composant le ménage ainsi que la région du ménage et le type de zone de résidence.

7.3. Modèle de Choix du Nombre de Véhicule

Le troisième et dernier sous-modèle comprenant le modèle complet de détention de véhicule correspond à la quantité de véhicules demandés. Ce sous-modèle estime la probabilité de posséder zéro, un et deux véhicules en utilisant l'échantillon complet des ménages disponibles. Le sous-modèle considère le fait de détenir zéro véhicule comme base de référence, en attribuant une valeur d'utilité de 1 à cette alternative. Les utilités de posséder un et deux véhicules par rapport à cela, V_1 et V_2 , sont alors estimées comme une fonction linéaire de variables explicatives. La probabilité de choisir un véhicule est calculée comme suit :

$$P_1 = \frac{\exp(V_1)}{1 + \exp(V_1) + \exp(V_2)}$$

Les facteurs déterminants du choix de la quantité de véhicules comprennent la zone géographique et les caractéristiques des ménages, et surtout une paire de variables capturant l'utilité moyenne obtenue par les ménages qui choisissent un véhicule, et deux véhicules, respectivement. Ces variables permettent le lien vers les sous-modèles de choix du type d'un véhicule et de deux véhicules. Les termes d'utilité sont calculés comme étant le log des dénominateurs des équations de probabilité provenant des sous-modèles de choix d'un et de deux véhicules, évalué pour chaque ménage de l'échantillon complet.

Le Tableau 7.4 présente les variables qui entrent en compte dans le sous-modèle de choix de la quantité de véhicules à détenir ainsi que leurs coefficients associés, leurs standard errors et t-statistics.

Tableau 7.4
Sous Modèle du Choix du Nombre de Véhicules

Variables Explicatives		Coefficients Estimés	Standard Error	t-statistic
1	Log du revenu des ménages (€), étant dans l'alternative à un véhicule	0.8120	0.0744	10.91
2	Log du revenu des ménages (€), étant dans l'alternative à deux véhicules	1.5916	0.1303	12.21
3	Nombre de travailleurs dans le ménage, étant dans l'alternative à un véhicule	0.3670	0.0485	7.57
4	Nombre de travailleurs dans le ménage, étant dans l'alternative à deux véhicules	0.7804	0.0567	13.76
5	Log du nombre de personnes composant le ménage, étant dans l'alternative à deux véhicules	1.2658	0.0796	15.91
6	Dummy pour petite ville, étant dans l'alternative à un véhicule	-0.4837	0.1205	-4.01
7	Dummy pour petite ville, étant dans l'alternative à deux véhicules	-0.8631	0.1355	-6.37
8	Dummy pour ville moyenne, étant dans l'alternative à un véhicule	-0.6289	0.1208	-5.21
9	Dummy pour ville moyenne, étant dans l'alternative à deux véhicules	-1.3451	0.1412	-9.53
10	Dummy pour grande ville, étant dans l'alternative à un véhicule	-0.9337	0.1006	-9.28
11	Dummy pour grande ville, étant dans l'alternative à deux véhicules	-1.7478	0.1164	-15.02
12	Dummy pour Paris, étant dans l'alternative à un véhicule	-2.0676	0.1128	-18.34
13	Dummy pour Paris, étant dans l'alternative à deux véhicules	-3.8671	0.1398	-27.67
14	Utilité moyenne dans le choix du type de véhicule pour l'alternative à un véhicule	0.5352	0.0542	9.87
15	Utilité moyenne dans le choix du type de véhicule pour l'alternative à deux véhicules	0.9440	0.1011	9.33
16	Constante pour l'alternative à un véhicule	-4.1170	0.8767	-4.70
17	Constante pour l'alternative à deux véhicules	-9.0976	1.9056	-4.77

Source : NERA - Modèle estimé par logit conditionnel sous Stata. Nombre d'observations : 105,663 ; Log Vraisemblance à convergence : 7,245.01.

Le sous-modèle inclut des termes distincts pour chacune des alternatives de posséder un et deux véhicules, mais avec les mêmes variables. Ceci s'explique par le fait que les mêmes facteurs qui déterminent l'utilité d'un ménage de posséder un véhicule sont susceptibles de déterminer aussi leur utilité de posséder deux véhicules, mais avec un poids différent.

Les deux premières lignes du Tableau 7.4 mettent en évidence l'importance du revenu dans la détermination du nombre de véhicules à acquérir. Les variables sont déterminées comme étant le logarithme des revenus des ménages, en euros par an, intégrées séparément pour les alternatives de détenir un ou deux véhicules. Les coefficients sont positifs et significatifs, comme prévu, et le coefficient pour l'alternative à deux véhicules est plus élevé que le coefficient pour l'alternative à un véhicule. Cela montre, comme attendu, que l'augmentation

du revenu augmente la probabilité qu'un ménage choisisse deux véhicules plutôt qu'un, ainsi que la probabilité qu'ils choisissent un véhicule plutôt qu'aucun.

Les deux lignes suivantes comprennent le nombre de travailleurs dans le ménage, entré séparément selon les deux alternatives. Les coefficients sont positifs et significatifs, comme prévu, et le coefficient pour l'alternative de posséder deux véhicules est plus élevé que le coefficient pour l'alternative à un véhicule. Cela montre qu'un travailleur supplémentaire dans le ménage augmente la probabilité qu'un ménage choisisse deux véhicules plutôt qu'un ainsi que la probabilité qu'ils choisissent un véhicule plutôt qu'aucun.

La cinquième variable du Tableau 7.4 est le log du nombre de membres composant le ménage. Cette variable entre en compte dans les deux alternatives avec le même coefficient. Lors des tests de spécification, il a été constaté que cette variable a des coefficients statistiquement similaires dans l'utilité d'avoir un ou deux véhicules, et a donc été combinée dans une seule variable. Elle ressort du modèle avec un coefficient positif et significatif, en indiquant comme prévu, que les ménages de grande taille sont davantage susceptibles de choisir de posséder une ou deux voitures plutôt qu'aucune.

Les lignes 6 à 13 du Tableau 7.4 contiennent des variables indicatrices pour le type de zone de résidence du ménage. Chacune de ces variables entre dans les alternatives à un et à deux véhicules séparément. Les caractéristiques des coefficients de ces variables indiquent ce qui suit.

- En comparaison avec les zones rurales (la catégorie omise) les ménages dans les moyennes et grandes villes sont moins susceptibles de posséder un véhicule (tous les coefficients sont négatifs et significatifs).
- La différence entre les zones rurales et les autres régions est plus grande pour l'alternative à deux véhicules que pour l'alternative à un véhicule (les coefficients pour l'alternative à deux véhicules sont tous plus importants, en valeur absolue, que les coefficients de l'alternative à un véhicule). Cela signifie que le passage d'une zone rurale à une ville diminuera l'opportunité de posséder deux véhicules plutôt qu'un, ainsi que l'opportunité de posséder un véhicule par rapport à aucun.
- Plus la zone urbaine est importante, moins l'utilité de posséder un véhicule est importante (les coefficients sont de plus en plus négatifs que l'on passe d'une petite ville à une ville moyenne, à une grande ville, et à Paris).

Les 14^{ème} et 15^{ème} variables du Tableau 7.4 mettent en évidence l'utilité moyenne obtenue par les choix de l'alternative de posséder un véhicule et deux véhicules. Comme décrit ci-dessus, ces variables sont calculées comme étant le log des dénominateurs des équations de probabilité des sous-modèles de choix d'un et de deux véhicules, évaluées pour chaque ménage de l'échantillon complet, et permet le lien vers les sous-modèles de choix de type d'un et de deux véhicules. Les coefficients de ces termes ressortent tous les deux avec des coefficients compris entre zéro et un, en vérifiant qu'ils sont compatibles avec la théorie de maximisation de l'utilité. En outre, le coefficient de l'alternative à deux véhicules est supérieur à celui de l'alternative à un véhicule, ce qui indique qu'une plus grande variété d'options disponibles dans une paire de type de véhicule est préférable pour les ménages par rapport à un seul type de véhicule.

Les deux dernières variables dans le sous-modèle de quantité de véhicules sont des constantes pour les alternatives à un et deux véhicules. Ces constantes sont estimées pour s'assurer que les parts estimées dans chaque alternative du sous-modèle sont égales à la part réelle de l'échantillon d'estimation.

Comme dans le cas des sous-modèles de choix d'un véhicule et de deux véhicules, un certain nombre de variables a été testé lors du développement de la spécification du sous modèle de choix de quantité de véhicules. Les variables suivantes ont été jugées insignifiantes, par rapport aux variables incluses dans le modèle choisi: l'âge, l'éducation et situation professionnelle de la personne de référence du ménage, le fait que cette personne vit en couple, le nombre d'enfants composant le ménage, et la région de résidence du ménage.

7.4. Élasticités

Les résultats décrits ci-dessus peuvent être utilisés pour calculer un intervalle d'estimations d'élasticités. Dans cette section, nous traitons l'élasticité-prix propre et l'élasticité-prix croisée de substitution par catégorie de véhicule, par millésime et par émissions de CO₂. De plus, nous considérons l'élasticité-revenu globale de la demande de véhicules. Nous comparons ces élasticités avec des résultats d'études similaires.

Le Tableau 7.5 montre les élasticités de la demande au prix d'achat, par catégorie de véhicule, provenant de nos résultats. La demande est élastique au prix pour toutes les catégories de véhicule, avec des élasticités allant de -1.15 pour les véhicules de catégorie A à -7.67 pour les voitures de sport. Les élasticités-prix augmentent pour les catégories de A à E, avec une élasticité particulièrement élevée de -4.07 pour les véhicules de catégorie E. Les voitures de sports ont également une élasticité de la demande élevée de -7.67.

Tableau 7.5
Elasticités de la Demande au Prix d'Achat, par Catégorie de Véhicule

	A	B	C	D	E	MPV	MiniMPV	SUV	Sports	TOTAL
A	-1.15	0.15	0.14	0.13	0.12	0.08	0.12	0.11	0.11	-0.03
B	0.41	-1.23	0.39	0.37	0.37	0.27	0.35	0.33	0.32	-0.06
C	0.31	0.31	-1.40	0.32	0.34	0.29	0.31	0.30	0.28	-0.04
D	0.26	0.27	0.28	-1.62	0.34	0.32	0.30	0.30	0.26	-0.03
E	0.11	0.12	0.13	0.15	-4.07	0.20	0.16	0.19	0.17	0.01
MPV	0.03	0.03	0.04	0.05	0.06	-1.85	0.05	0.06	0.04	0.00
MiniMPV	0.30	0.32	0.33	0.35	0.40	0.40	-2.01	0.36	0.30	-0.03
SUV	0.07	0.07	0.08	0.09	0.12	0.13	0.10	-3.64	0.10	0.00
Sports	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.03	-7.67	0.00

Source: NERA. Les données du tableau montrent la variation en pourcentage du nombre total de véhicules détenus par catégorie en colonne, étant donné une simulation d'augmentation de prix de 1 % des véhicules de catégories définies en ligne.

Ces résultats sont relativement en ligne avec ceux d'études précédentes, résumées dans le Tableau 7.8. Eftec (2008), CE et BMRB (2008), Train et Winston (2007) et Berry et al (1995), lesquels sont tous arrivés au résultat que la demande de véhicule est élastique au prix. Cependant, il faut noter que ces études ne sont pas toutes directement comparables aux résultats exposés précédemment car les méthodologies utilisées ne sont pas les mêmes (voir Tableau 7.8).

Le résultat montrant que l'élasticité-prix est plus forte pour les plus grosses voitures et luxueuses (*i.e.* augmentation de l'élasticité entre les catégories A à E) est en ligne avec les résultats de CE and BMRB (2008). Cependant, d'autres études n'ont pas démontré une élasticité-prix de la demande aussi forte pour les voitures de sport.

Nous trouvons des élasticités-prix croisées assez faibles (inférieure à un) pour toutes les catégories de véhicules. Ces résultats sont consistants avec les études de CE et BMRB (2008) et Eftec (2008).

Le Tableau 7.6 met en avant les élasticités-prix d'achat de la demande par groupe de millésime, provenant des résultats obtenus par notre étude. Les élasticités-prix propres de la demande diminuent avec l'âge du véhicule; la demande de nouveaux véhicules étant fortement élastique au prix alors que celle des véhicules de plus de 5 ans est assez inélastique. Ces résultats sont en partie expliqués par le fait que la catégorie des véhicules « neufs » contient beaucoup moins de véhicules que la catégorie des véhicules de « Plus de 5 ans ».

Tableau 7.6
Elasticités de la Demande au Prix d'Achat, par Groupe de Véhicules de Différents Millésimes

Groupe du Millésime	Neufs (Moins d'un an)	1-5 ans	Plus de 5 ans	TOTAL
Nouveaux (Moins d'un an)	-5.32	0.42	0.30	-0.01
1-5 ans	0.89	-2.56	0.74	-0.04
Plus de 5 ans	0.44	0.55	-0.42	-0.13

Source : NERA. Les données du tableau montrent la variation en pourcentage du nombre total de véhicules détenus par catégorie en colonne, étant donnée une simulation d'augmentation de prix de 1 % des véhicules de catégories définies en ligne.

Des caractéristiques similaires des élasticités-prix diminuant lorsque l'âge du véhicule augmente ont été mises en avant par d'autres études. Par exemple, Bento et al (2009) ont établi que les voitures ont une élasticité

d'acquisition de -2.0 à l'égard des prix de location, mais que l'élasticité diminue à -1.0 pour les véhicules âgés de 1 à 2 ans et est inférieure à 1 pour les voitures de plus de trois ans.

Le Tableau 7.7 met en évidence les élasticités de la demande au prix d'achat de nouveaux véhicules par groupe de niveau d'émissions de CO₂. Les élasticités-prix propres ont tendance à augmenter avec la quantité de CO₂ émise, avec des élasticités très élevées pour les véhicules émettant 201 g/km de CO₂ ou plus. Toutes les élasticités-prix propres sont élevées dans le tableau ci-dessous, qui offre un résultat consistant avec le résultat montrant qu'il existe une substitution forte entre les nouveaux véhicules et les plus anciens compte tenu de l'augmentation des prix des nouveaux véhicules.

Tableau 7.7
Elasticités de la Demande pour les Nouveaux Véhicules, par Groupe de Véhicules Selon leurs Emissions de CO₂ (g/km)

Groupe d'Emissions de CO ₂	Moins de 100	101-120	121-130	131-160	161-165	166-200	201-250	Plus de 250	TOTAL
Moins de 100	-4.70	0.01	0.01	0.01	0.01	0.01	0.01	0.01	-0.12
101-120	0.12	-4.16	0.13	0.13	0.14	0.14	0.17	0.22	-1.93
121-130	0.03	0.03	-4.79	0.04	0.04	0.04	0.06	0.07	-0.48
131-160	0.11	0.12	0.13	-6.03	0.17	0.19	0.25	0.33	-1.53
161-165	0.01	0.02	0.02	0.02	-7.74	0.03	0.04	0.05	-0.20
166-200	0.05	0.06	0.07	0.08	0.09	-9.83	0.15	0.21	-0.65
201-250	0.02	0.02	0.02	0.03	0.03	0.04	-23.75	0.09	-0.19
Plus de 250	0.02	0.02	0.02	0.03	0.04	0.04	0.06	-60.97	-0.19

Source : NERA. Les données du tableau montrent la variation en pourcentage du nombre total de véhicules détenus par catégorie en colonne, étant donnée une simulation d'augmentation de prix de 1 % des véhicules de catégories définies en ligne.

CE et BMRB (2008) et Efttec (2008) ont également étudié le degré de substitution entre les différents groupes d'émissions de CO₂ suite à un changement de prix. Cependant ces résultats ne sont pas comparables à ceux trouvés dans notre étude car ils ne prennent pas en considération la possibilité de substitution entre nouveaux et anciens véhicules qui ne sont pas sujets aux mêmes changements de prix. Pour cette raison, les résultats de ces études montrent des élasticités de substitutions plus faibles. CE et BMRB (2008) trouvent des élasticités-prix propres comprises entre -4 et -2, par exemple, selon la fourchette d'émissions de CO₂. De plus, CE et BMRB (2008) ne trouvent pas, à l'inverse de nous, que les élasticités-prix propres augmentent avec le niveau d'émissions de CO₂.

Le Tableau 7.8 ci-dessous résume les résultats sur les élasticités-prix d'autres études clés.

L'élasticité-revenu de la demande est calculée comme étant le pourcentage d'augmentation du nombre de véhicules détenus par les ménages étant donné une augmentation de 1 % du revenu de tous les ménages dans l'échantillon. Les estimations résultant de ce calcul est de 0.33. Nous n'avons pas trouvé de résultats comparables dans la littérature sur l'élasticité-revenu de la demande pour la détention de véhicules.

Tableau 7.8
Résumé des Résultats sur les Elasticités des Principales Etudes

Etude	Elasticité-Prix	Commentaires
Bento et al (2009)	L'étude considère l'élasticité du choix de détention de véhicules selon le prix de location de voitures. Elle met en évidence des élasticités entre -0.62 et -1.25 pour toutes les voitures, et entre -1.32 and -3.14 pour les voitures neuves.	Le Modèle a pour sujet les décisions d'acquisition de voitures et de kilomètres à parcourir. Pas d'interaction entre les véhicules déjà détenus par les ménages.
Eftec (2008)	A la médiane, la demande est élastique. L'élasticité-prix de la demande est plus élevée pour les petites voitures que pour les 4x4. Les élasticités-prix sont comprises entre -1.5 pour les mini MPVs (eg Renault Scenic) à -5.2 pour les voitures de catégorie C (eg Honda Civic). Les élasticités-prix croisées sont plus élevées pour les petites voitures. Par exemple, les élasticités-prix croisées dans le groupe de voitures de catégorie A sont de +0.105.	L'étude modélise les décisions d'achat de nouveaux véhicules.
CE et BMRB (2008)	Les élasticités-prix directes et indirectes sont généralement supérieures à -1 (en valeur absolue) dans une même catégorie. L'élasticité-prix de la demande est comprise entre -0.96 pour les 4x4 à -3.5 pour les voitures de luxe. Les élasticités-prix croisées sont plus élevées pour les catégories de véhicules similaires. Les élasticités-prix croisées de la demande pour d'autres catégories de véhicules en réponse à un changement de prix des voitures de luxe sont inférieures à 0.055.	L'étude a pour sujet les achats de véhicules neufs et quasi-neufs.
Train et Winston (2007)	L'élasticité-prix moyenne est de -2.32 pour toutes marques et modèles.	L'étude estime le choix conditionnel d'achat d'un véhicule neuf.
Prieto (2006)	Les élasticités-prix croisées sont plus élevées pour les véhicules de types similaires.	
Greene et al (2005)	La probabilité de choisir un nouveau véhicule est une fonction négative du prix	
Mohammadian et Miller (2003)	Les auteurs prennent des hypothèses sur la valeur des élasticités dans le modèle	
Cecchini (2001)	Elasticités-prix non présentées	
McCarthy et Tay (1998)	Le prix influe sur la décision d'achat. Les élasticités sont présentées pour 12 combinaisons de marques et de modèles. L'augmentation du prix d'un véhicule fait baisser sa demande. Les élasticités-prix propres sont dans l'intervalle (-1.1,-1.0) pour les véhicules peu efficaces en consommation de carburant et (-0.9,-0.8) pour les plus efficaces. La demande globale de véhicule a une élasticité-prix propre juste en dessous de 1. Les élasticités-prix croisées à l'intérieur d'un groupe de type de consommation de carburant et entre groupes sont également considérées.	Etude de marché des petites voitures. Modèles de demande de nouveaux véhicules.
Berry et al (1995)	La demande est élastique au prix pour tous les modèles sur un échantillon. Les élasticités-prix calculées sont comprises entre -3.1 pour un LS400 à -6.8 pour un Buick Century. Les semi-élasticités prix croisés sont plus élevées entre les véhicules ayant des caractéristiques comparables.	
Berry et al (2004)	L'élasticité implicite du marché se situe entre -0.2 and -1, selon les hypothèses. Les semi-élasticités prix diminuent avec le prix, et à un niveau constant de prix, sont plus faibles pour les vans, pickups, SUVs et (à degré moindre) les voitures de sport.	Examine les achats de véhicules.
Berkovec et Rust (1984)	Le prix a un effet négatif sur la demande. Les élasticités-prix ne sont pas calculées.	Modèle de choix de véhicule pour les ménages ayant un seul véhicule.

Source: NERA

8. Résultats des Simulations

Le modèle de choix de véhicules développé dans cette analyse peut être utilisé pour simuler l'impact sur les choix des ménages, des politiques publiques (dont le Bonus Ecologique) qui affectent le prix d'achat des véhicules ou leurs coûts d'entretien. Le modèle est capable de prévoir la demande de véhicules par fourchette d'émissions de CO₂, par millésime, et par catégorie de véhicule compte tenu des changements qui ont eu lieu dans la structure du marché automobile (nature des véhicules proposés), de la structure des ménages dans la population française et des changements des politiques publiques. Il n'est toutefois pas à même de prédire les évolutions de l'offre de véhicules sur le marché.

Dans cette section, nous utilisons le modèle afin de simuler l'impact que l'introduction du bonus Ecologique en 2008 pourrait avoir sur les tendances d'achat de véhicules compte tenu des changements intervenus dans la structure du marché automobile français entre 05/06, date du sondage INSEE, et janvier 2009, date à laquelle la nature des véhicules disponibles sur le marché et leurs prix ont été mis à jour.

Nous comparons les scénarii suivants :

- Scénario de base : Données véhicules et ménages de 2006, pas de taxe à l'achat
- Scénario 1 : Données véhicules et ménages de 2006, avec taxe à l'achat
- Scénario 2 : Données véhicules et ménages de 2009, pas de taxe à l'achat
- Scénario 3 : Données véhicules et ménages de 2006, avec taxe à l'achat

Les résultats de ces différents scénarii sont présentés ci-après. Pour les nouveaux véhicules, nous comparons également les résultats obtenus par le biais de la simulation avec les données de ventes réelles telles qu'obtenues dans la base de données AAA.

8.1. Données à renseigner

Les montants retenus pour les taxes utilisées dans les scénarii 1 et 3 sont présentés ci-dessous :

Tableau 8.1
Bonus écologique

Fourchette d'émission de CO ₂ (g/km)	Taxe (en Euros)	
	Montant Nominal	Prix de 2006
Less than 100	-1000	-949.26
101-120	-700	-664.48
121-130	-200	-189.85
131-160	0	0
161-165	200	189.85
166-200	750	711.94
201-250	1600	1518.81
More than 250	2600	2468.06

Source : http://www.effet-de-serre.gouv.fr/bonus_ecologique_voiture; les montants nominaux ont été convertis en valeurs de 2006 en utilisant HICP (source : Eurostat).

Pour les scénarii de 2009, les montants des taxes ainsi que les prix des véhicules ont été convertis en valeurs de 2006 en utilisant HICP (Source : Eurostat). Les prix des carburants et les données démographiques relatives aux ménages présentés ci-après sont également renseignés dans le modèle pour les besoins de la simulation. Les données 2006 et 2009 sont les mêmes dans certains cas compte tenu de l'indisponibilité de données mises à jour.

Tableau 8.2
Éléments renseignés pour les besoins de la Simulation

	Données 2006	Données 2009
Coût des Carburants (Euro/litre, prix 2006)⁽¹⁾		
Essence	1.1975	1.2578
Diesel	1.0325	1.1367
Autres	0.6575	0.7024
No. Ménages ('000)⁽²⁾	26,069	26,069
No. moyen de véhicules par ménage possédant au moins deux véhicules⁽³⁾	2.2	2.2
Distribution des Revenus Annuels des Ménages (prix 2006)⁽⁴⁾		
Moins de €16 806	25.0 %	24.0 %
Entre €16 806 et €26 510	25.0 %	24.5 %
Entre €26 511 et €38 870	25.0 %	25.3 %
Plus de €38 870	25.0 %	26.3 %
Distribution par nombre de travailleurs dans le ménage⁽³⁾		
0	31.6 %	31.6 %
1	32.1 %	32.1 %
2+	36.3 %	36.3 %
Distribution par type de zones de résidence^{(3)(a)}		
Zone rurale	24.1 %	24.1 %
Petite ville	16.8 %	16.8 %
Ville moyenne	13.2 %	13.2 %
Grande ville	29.3 %	29.3 %
Paris	16.6 %	16.6 %
Distribution par nombre de personnes dans le ménage⁽³⁾		
1	28.7 %	28.7 %
2	35.2 %	35.2 %
3	14.3 %	14.3 %
4+	21.8 %	21.8 %

*Sources : (1) MEDDTL, prix 2006 = 0.75 * prix 2005 + 0.25 * prix 2006 ; prix 2009 = 0.75 * prix 2008 + 0.25 * prix 2009 ; prix nominal converti en prix 2006 en utilisant HICP (source: Eurostat) ; (2) INSEE, "Population census 2006" ; (3) INSEE "Budget des Familles 2006" ; (4) Source données 2006 = INSEE "Budget des Familles 2006", Données 2009 obtenues en augmentant tous les revenus de l'échantillon INSEE 2006 de 2.34 %, ce qui équivaut à la croissance réelle du PIB entre le Q1 de 2006 et le Q1 de 2009 (Source pour le PIB : Eurostat). Notes : (a) Agglomération Parisienne fait référence à l'unité urbaine de Paris, une zone définie par l'INSEE et composée de presque 400 communes en Île de France. Les Grandes villes sont des « communes appartenant à des unités urbaines de plus de 100 000 habitants ». Les Villes moyennes sont des « communes appartenant à des unités urbaines de 20 000 à 100 000 habitants ». Les Petites villes sont des « communes appartenant à des unités urbaines de moins de 20 000 habitants ». Rural correspond à des communes rurales.*

Les dernières informations à entrer pour les besoins de la simulation sont les données relatives au type de véhicule. Le Tableau 8.3 donne les prix moyens, la consommation moyenne de carburant, le volume moyen du

coffre, et le nombre de modèles disponibles par classe, millésime, fourchette d'émissions de CO₂ et type de carburant. Ces données sont fournies pour 2006 et 2009.

La comparaison entre 2006 et 2009 appelle les remarques suivantes :

- Les prix ont augmenté de façon substantielle. Parmi toutes les catégories, millésimes, fourchettes d'émissions de CO₂ et types de carburants, seuls les véhicules anciens ont vu leurs prix baisser. Le prix des nouveaux véhicules en particulier a connu une très forte hausse entre 2006 et 2009.
- L'efficacité des véhicules en termes de consommation s'est améliorée pour les catégories A, B et C mais a empiré pour les catégories D, E, MPV et Sports. En moyenne, l'efficacité en termes de consommation est plus faible pour les nouveaux véhicules de 2009 que pour les nouveaux véhicules de 2006.
- Le volume du coffre est resté le même dans toutes les catégories
- Le nombre de nouveaux véhicules proposés a été multiplié par 3 entre 2006 et 2009. Il y a un plus ample choix de véhicules au sein de chaque catégorie, chaque fourchette d'émissions de CO₂ et chaque type de carburant. Le nombre de véhicules entrant dans la catégorie « Autres » pour les types de carburants a été quasiment multiplié par 5.

Tableau 8.3
Caractéristiques des Véhicules Disponibles en 2006 et 2009

	Prx (000 €)		Consommation de carburant (litres par 100km)		Capacité du coffre (m³)		Nombre de modèles de véhicules disponibles	
	2006	2009	2006	2009	2006	2009	2006	2009
Catégorie								
A	4.3	4.9	5.6	5.2	0.2	0.2	143	180
B	5.5	6.7	5.8	5.6	0.3	0.3	361	511
C	6.7	9.0	6.7	6.5	0.4	0.4	495	529
D	9.5	12.2	7.4	7.5	0.5	0.5	517	592
E	15.0	24.6	9.2	9.4	0.5	0.5	326	369
MPV	10.3	13.2	8.5	8.6	0.6	0.6	309	328
MiniMPV	8.8	9.9	6.8	6.8	0.5	0.5	298	532
SUV	14.8	19.8	10.0	9.9	0.5	0.5	402	570
Sports	12.8	25.0	8.7	9.0	0.3	0.3	191	227
Millésime								
Moins d'1 an	18.6	26.5	7.1	7.5	0.4	0.4	297	918
1-5 ans	12.4	13.9	7.7	7.5	0.4	0.4	1444	1534
More than 5 years	5.0	4.6	7.8	7.9	0.4	0.4	1301	1386
Emissions de CO2 (g/km)								
Moins de 100	5.2	6.1	3.4	3.5	0.2	0.2	7	15
101-120	6.7	8.7	4.4	4.6	0.2	0.2	94	225
121-130	6.2	9.5	4.9	5.0	0.3	0.3	79	138
131-160	7.1	8.7	5.8	5.9	0.3	0.4	717	955
161-165	7.0	9.9	6.5	6.5	0.4	0.4	126	140
166-200	8.9	12.2	7.3	7.3	0.4	0.5	903	1100
201-250	11.2	15.8	9.0	9.1	0.5	0.5	749	772
Plus de 250	17.4	27.0	11.5	11.9	0.5	0.5	367	493
Type de carburant								
Essence	9.0	14.0	8.5	8.6	0.4	0.4	1666	1922
Diesel	11.1	13.4	6.6	6.5	0.5	0.4	1335	1720
Autre	5.4	9.8	9.3	7.9	0.4	0.4	41	196
TOTAL	9.8	13.5	7.7	7.6	0.4	0.4	3042	3838

Source: NERA

8.2. Résultats des Scenarii

Les tableaux et représentations graphiques ci-dessous présentent les résultats de simulation de scenarii décrit précédemment. Ces simulations ont été réalisées en utilisant le NERA Vehicle Choice Simulation Model v1.2 accompagnant ce rapport.

Tableau 8.4
Résultats des Simulations des Scenarii de Politiques (Proportions)

	Base	Parts estimées			Parts observées (2006)
		Scenario 1	Scenario 2	Scenario 3	
Par Catégorie					
A	11 %	11 %	13 %	13 %	5 %
B	24 %	24 %	27 %	27 %	36 %
C	26 %	26 %	20 %	20 %	24 %
D	18 %	18 %	16 %	16 %	16 %
E	5 %	5 %	3 %	3 %	3 %
MPV	3 %	3 %	2 %	2 %	3 %
Mini MPV	10 %	10 %	16 %	16 %	10 %
SUV	2 %	2 %	2 %	2 %	2 %
Sports	1 %	1 %	0 %	0 %	1 %
Par Millésime					
Neuf (Moins d'1 an)	4 %	4 %	6 %	7 %	3 %
Moyennement-Agé (1-5 ans)	27 %	27 %	25 %	25 %	28 %
Vieux (Plus de 5 ans)	69 %	69 %	69 %	68 %	69 %
Par Carburant					
Essence	42 %	42 %	32 %	32 %	50 %
Diesel	54 %	54 %	56 %	56 %	49 %
Autre	4 %	4 %	12 %	12 %	0 %
Par émissions de CO₂ (g/km)					
bonus faible	1 %	1 %	2 %	2 %	1 %
bonus moyen	8 %	8 %	12 %	12 %	10 %
bonus élevé	8 %	8 %	5 %	5 %	11 %
Neutre	42 %	42 %	44 %	44 %	45 %
malus faible	6 %	6 %	5 %	5 %	6 %
malus moyen	27 %	27 %	24 %	24 %	21 %
malus élevé	8 %	8 %	6 %	6 %	5 %
malus très élevé	1 %	1 %	1 %	1 %	1 %
Par émissions de CO₂, Nouveau millésime seulement					
bonus faible	2 %	2 %	2 %	3 %	0 %
bonus moyen	19 %	21 %	43 %	47 %	12 %
bonus élevé	11 %	11 %	11 %	11 %	15 %
Neutre	51 %	50 %	29 %	28 %	48 %
malus faible	3 %	3 %	3 %	3 %	4 %
malus moyen	12 %	11 %	9 %	8 %	15 %
malus élevé	2 %	2 %	1 %	1 %	5 %
malus très élevé	0 %	0 %	0 %	0 %	1 %

Source : NERA

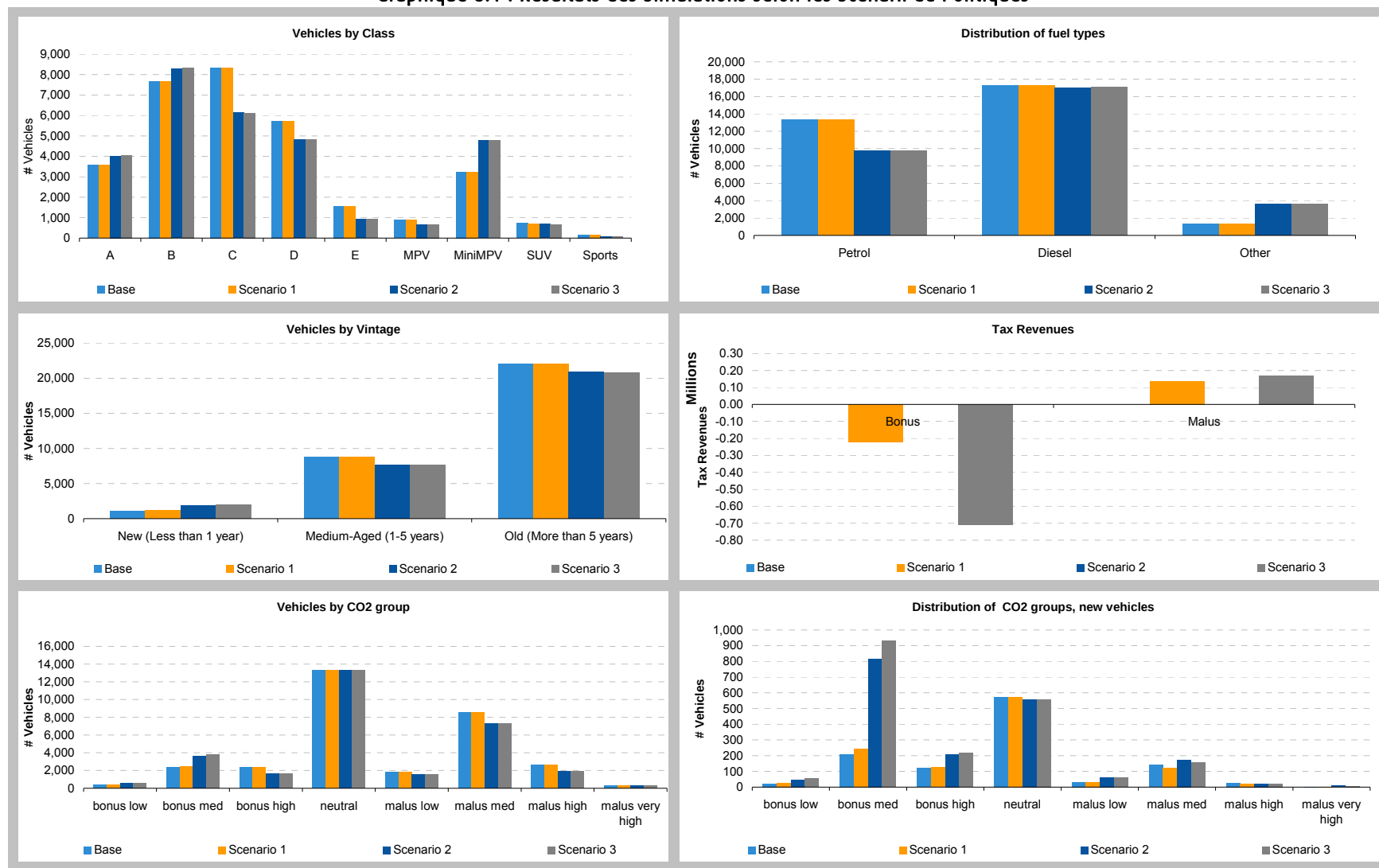
Tableau 8.5
Résultats des Simulations des Scenarii de Politiques (Nombre)

	Total Vehicles ('000)			
	Base	Scenario 1	Scenario 2	Scenario 3
Par Catégorie				
A	3,565	3,581	4,039	4082
B	7,671	7,688	8,365	8,405
C	8,339	8,332	6,205	6,193
D	5,742	5,736	4,888	4,873
E	1,562	1,561	941	938
MPV	895	894	696	694
Mini MPV	3,248	3,240	4,870	4,858
SUV	723	720	704	699
Sports	166	166	83	82
Par Millésime				
Neuf (Moins d'1 an)	1,127	1,146	1,962	2,065
Moyennement-Agé (1-5 ans)	8,750	8,748	7,811	7,793
Vieux (Plus de 5 ans)	22,033	22,024	21,016	20,966
Par Carburant				
Essence	13,352	13,338	9,901	9,895
Diesel	17,276	17,298	17,207	17,247
Autre	1,282	1,281	3,682	3,681
Par émissions de CO2 (g/km)				
bonus faible	426	431	593	601
bonus moyen	2,405	2,437	3,718	3,825
bonus élevé	2,397	2,401	1,692	1,695
Neutre	13,292	13,286	13,483	13,449
malus faible	1,843	1,842	1,583	1,577
malus moyen	8,559	8,540	7,412	7,378
malus élevé	2,646	2,640	2,004	1,996
malus très élevé	341	341	305	302
Total	31,910	31,917	30,790	30,823
Par émissions de CO2, Nouveau millésime seulement				
bonus faible	21	26	47	57
bonus moyen	210	243	841	956
bonus élevé	120	125	215	223
Neutre	574	574	578	577
malus faible	32	31	64	62
malus moyen	140	124	181	163
malus élevé	27	21	25	20
malus très élevé	2	2	10	7
Total	1,127	1,146	1,962	2,065
Revenu de la taxe sur le Co2 ('000)				
Bonus	€ 0	(€ 220,819)	€ 0	(€ 731,331)
Malus	€ 0	€ 137,435	€ 0	€ 176,425
Revenu Net	€ 0	(€ 83,384)	€ 0	(€ 554,906)

Source : NERA

Base : Données 2006 relatives aux véhicules et aux ménages, sans taxe à l'achat ; Scenario 1 : Données 2006 relatives aux véhicules et aux ménages, avec une taxe à l'achat ; Scenario 2 : Données 2009 relatives aux véhicules et aux ménages, sans taxe à l'achat ; Scenario 3 : Données 2009 relatives aux véhicules et aux ménages, avec une taxe à l'achat.

Graphique 8.1 : Résultats des Simulations selon les Scenarii de Politiques



En comparant le scénario 3 (données 2009 sur les véhicules avec une taxe) au scénario de base (données 2006 sur les véhicules sans taxe), nous trouvons une diminution du nombre de voitures détenues d'environ 1,1 million, équivalant à une baisse de 3,4 % du nombre total de véhicules détenus. Il y a plus de véhicules de catégories A, B et Mini MPV, mais moins d'autres catégories. Il y a davantage de nouveaux véhicules, et moins d'âge intermédiaire et de vieux véhicules. Il y a davantage de véhicules à carburants alternatifs et moins de véhicules à moteurs diesel et essence. Il y a également plus de véhicule à faibles émissions de CO₂ et moins de véhicules à fortes émissions de CO₂ et ceci est d'autant plus vrai pour les véhicules neufs que pour tous les autres millésimes. Le revenu estimé du Bonus écologique est de 554 900 €.

Il faut noter que l'effet de mettre en place la taxe « Bonus écologique » sur les véhicules détenus, par catégorie, millésime, type de carburant et émissions de CO₂, est substantiellement moins prononcé que de changer de véhicule de 2006 pour un véhicule de 2009. Cela se traduit par le fait que les résultats des scénarii 2 et 3 sont très similaires, tout comme le scénario de base et le scénario 1 ; en revanche il existe de grandes différences entre les scénarii de base et le 2 et le 1 et le 3. Ce résultat suggère que la plus grande part de la diminution des émissions de CO₂ des véhicules des ménages sera due à des changements au niveau de l'offre des véhicules et non au niveau de la demande. Train, Davis et Levine (1997) sont arrivés à des résultats similaires pour le marché des États-Unis.

Les résultats pour les nouveaux véhicules peuvent être comparés aux données des ventes sur le marché français. Une telle comparaison est présentée dans le Tableau 8.6, qui met en évidence les proportions estimées des véhicules détenus par fourchette d'émissions de CO₂ à partir du modèle de simulation NERA, pour 2006 et 2009 (scénario de base et scénario 3) et à partir de la série de données de AAA sur les ventes de véhicules en France.

Tableau 8.6
Comparaison des Résultats des Simulations et des Données AAA sur les Ventes

Emissions de CO ₂ (g/km)	2006		2009	
	Simulation du Scénario de Base ⁽¹⁾	Données AAA sur les Ventes ⁽²⁾	Simulation Scénario 3 ⁽¹⁾	Données AAA sur les Ventes ⁽²⁾
<100	2 %	0 %	3 %	0 %
101-120	18 %	13 %	46 %	46 %
121-130	11 %	9 %	11 %	8 %
131-160	51 %	43 %	28 %	36 %
161-165	3 %	6 %	3 %	1 %
166-200	13 %	20 %	8 %	6 %
201-250	2 %	6 %	1 %	2 %
>250	0 %	2 %	0 %	0 %

Sources: (1) NERA Vehicle Choice Simulation v1.1 ; (2) AAA. Notes: Les données NERA issues des simulations montrent les proportions estimées des véhicules détenus qui ont moins de un an. Les données AAA pour 2009 sont basées sur les ventes effectuées de janvier à Septembre 2009.

Les résultats du modèle ne sont pas directement comparables avec les ventes car ils reflètent le nombre de véhicules détenus ayant moins d'un an, plutôt que le nombre de ventes effectuées dans l'année. Cependant, cette comparaison est encourageante car elle met en évidence des caractéristiques similaires dans la détention de véhicules neufs en 2006 et 2009. Les mêmes tendances sont également observées entre les deux années.

9. Conclusions

Dans ce rapport, nous avons présenté les conclusions de trois sous-modèles composant notre modèle d'acquisition de véhicule. Les sous-modèles sont relatifs au choix du type de véhicule pour les ménages ayant choisi d'avoir un seul véhicule, au choix de la paire véhicules pour les ménages ayant choisi d'en avoir deux, et au choix du nombre de véhicule pour tous les ménages. Les estimations de ces modèles englobent le cœur de l'outil de simulation qui accompagne ce rapport. De plus, ce rapport a présenté une revue de la littérature, une description des séries de données utilisées, une description de la méthodologie choisie pour les constituer et quelques résultats de simulations de politiques identifiant l'effet de l'introduction du Bonus écologique en France en 2007-2008.

En guise de conclusions, cette partie résume les principales conclusions de cette étude, critique leurs robustesses, et met en avant deux extensions possibles que nous considérons comme pertinentes pour de futures études, mais qui dépassent le périmètre de cette étude. Ces extensions comprennent : (i) inclure des sous-modèles pour expliquer la demande des ménages en carburant en établissant un lien avec les sous-modèles de choix de quantité et de types de véhicules ; et (ii) inclure l'option de choisir trois véhicules dans les sous-modèles de quantité au lieu de restreindre le choix à zéro, un ou deux véhicules.

9.1. Résumé des Principales Conclusions

Le modèle de choix d'acquisition de véhicule s'appuie sur le modèle économique standard de choix - le modèle d'utilité aléatoire - et utilise la formulation d'un logit hiérarchisé pour assurer des liens théoriques cohérents entre les choix de quantité et de type de véhicule(s).

Ces trois sous-modèles apparaissent bien fondés. Les variables supposées influencer chaque décision apparaissent, dans les sous-modèles, avec des coefficients de signes anticipés, statistiquement significatifs, et d'un ordre de grandeur plausible.

Pour les ménages qui choisissent d'acquérir un seul véhicule, nous trouvons que :

- Les personnes préfèrent les véhicules les moins chers aux véhicules les plus chers, *ceteris paribus*, mais le prix d'achat et les coûts de fonctionnement des véhicules sont plus importants pour les ménages les moins aisés que pour les plus aisés.
- La propension implicite à payer pour bénéficier d'une réduction des coûts variables relatifs au véhicule, en terme de prix additionnel à payer, semble être raisonnable. Les ménages à faibles revenus auraient besoin de parcourir entre 8 200 km et 10 800 km par an pour justifier leur propension implicite à payer pour des coûts variables réduits. Les ménages à revenus élevés auraient besoin de parcourir entre 17 600 et 23 200 km par an pour justifier leurs dépenses en véhicules plus efficaces en consommation d'essence.
- La taille du coffre est importante pour les ménages composés d'au moins quatre membres, mais n'est pas significative pour les autres.
- Les personnes préfèrent des types de véhicules où ils ont le plus grand choix de types de marques et de modèles et où il y a un grand intervalle de prix.
- Les véhicules neufs ont une plus grande valeur que les véhicules ayant entre 1 et 5 ans, qui eux-mêmes ont une plus grande valeur que les véhicules les plus vieux, *ceteris paribus*.
- Les MPVs sont moins attractifs que les autres véhicules *ceteris paribus*, mais sont relativement plus attractifs pour les ménages composés d'au moins quatre membres. Les ménages composés d'au moins quatre membres préfèrent cependant les autres types de voitures aux MPVs, *ceteris paribus*. Pour les ménages, les voitures de sports et les SUVs sont moins attractives que les autres types de voitures, *ceteris paribus*.

Pour les ménages détenant deux véhicules nous trouvons des résultats très similaires à ceux pour le sous-modèle des ménages détenant un seul véhicule, à savoir :

- Les ménages sont sensibles aux prix d'achat et aux coûts variables des véhicules, et sont plus sensibles lorsque leurs revenus sont faibles.

- Ils désirent une grande capacité de coffre, et c'est d'autant plus important pour les ménages composés d'au moins quatre membres.
- Ils sont plus intéressés par les nouveaux millésimes, et moins par les MPVs, SUVs et voitures de sports, *ceteris paribus*.
- Les MPVs sont aussi relativement plus attractifs pour les ménages composés de quatre membres ou plus.
- Les personnes préfèrent les types de véhicules où ils ont le plus grand choix de marques et de modèles. Concernant le sous-modèles de quantité, nous avons trouvé que :
- Une hausse des revenus augmente la probabilité qu'un ménage choisisse deux véhicules plutôt qu'un, et un véhicule plutôt qu'aucun.
- Un travailleur supplémentaire dans le ménage augmente la probabilité que le ménage choisisse deux véhicules plutôt qu'un et un véhicule plutôt qu'aucun. Un membre supplémentaire dans le ménage aura des effets similaires.
- Déménager d'une zone rurale à une ville diminue la probabilité de détenir deux véhicules plutôt qu'un, et de détenir un véhicule plutôt qu'aucun. De plus, plus la zone urbaine est importante, plus la probabilité de détenir un véhicule est faible.

9.2. Discussion

La série de données pour notre analyse contient un échantillon très important et représentatif des ménages français, lié à un autre échantillon exhaustif des véhicules existants. La taille de l'échantillon est plus de dix fois supérieure, par exemple, à la récente étude de Cambridge Econometrics/BMRB pour le DfT, et la quantité de variables est au moins aussi riche, si ce n'est plus. En utilisant cette série de données, nous avons été capables d'implémenter avec succès une méthodologie de modélisation qui est consistante avec la théorie économique et qui aboutie à des résultats plausibles. Les résultats de cette modélisation sont le cœur de l'outil de simulation qui accompagne ce rapport.

Cependant, il y a certaines caractéristiques de la série de données qui doivent être considérées afin que les résultats de ce rapport, et de l'outil de simulation, soient interprétés avec une certaine prudence. Les principales raisons sont :

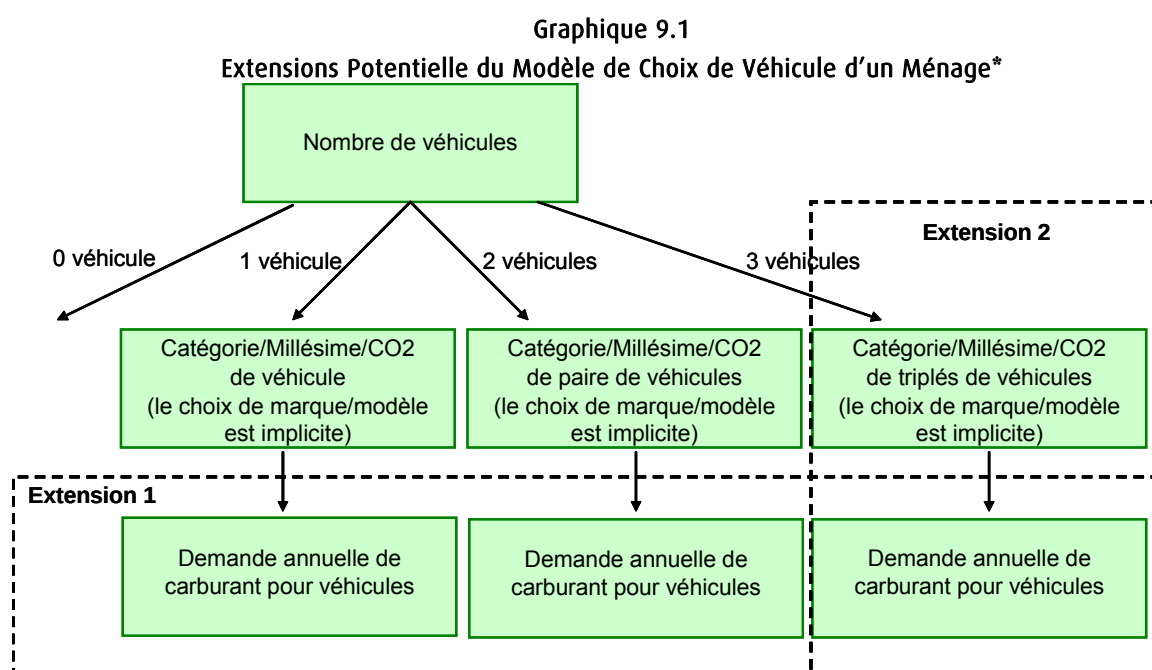
- Les données de l'INSEE ne donnent pas exactement le type de véhicule détenu.
 - Au mieux, ces données permettent de connaître la marque, le modèle, le millésime, et le type de carburant du véhicule. Cependant nous ne connaissons jamais la version du modèle, et donc les émissions en CO₂ des véhicules détenus. Cela implique que nous sommes incapables de modéliser, par exemple, les changements de véhicule d'une version à une autre qui peut être due à une variation de prix.
 - Au pire, les données sur les modèles des véhicules détenus sont manquantes. Par exemple, certains véhicules sont décrits comme étant simplement des "Audi" ou "Renault". La série de données est donc inévitablement moins robuste du fait de ce manque d'information.
- Les données JATO et L'Argus sont incomplètes ;
 - Il n'y a pas de données JATO pour les véhicules de plus de 10 ans ;
 - Les données relatives émissions de CO₂ sont manquantes pour les véhicules relativement anciens (5-10 ans) ;
 - JATO ne contient pas de données sur les vans ;
 - Les véhicules peu vendus n'étaient pas bien représentés dans l'Argus (par exemple, les véhicules à carburant alternatif) ;
- Nous n'avons pas de données sur les véhicules précédemment détenus ne permettant pas à notre modèle de prendre en compte cette dynamique.

9.3. Recommandations pour des recherches Futures

La méthodologie de modélisation implémentée dans cette étude remplit les objectifs dans le sens où elle fournit une bonne estimation des paramètres influant sur le choix de véhicule qui sont utilisés dans l'outil de simulation permettant de prévoir la demande agrégée de véhicule, par type de véhicule, selon des scénarii de politiques prédéfinis. Il y a un nombre important de sujets qu'il semblerait être intéressant de traiter dans le futur, si plus de résultats ou un plus grand degré de robustesse sont désirés. Les extensions potentielles sont :

- Ajouter la demande des véhicules en carburant en tant que sous-modèle ; et
- Ajouter le fait de détenir trois véhicules comme une option de choix.

La Graphique 9.1 met en avant la façon dont ces ajouts seraient annexés à la structure du modèle principal décrit précédemment. La suite de cette section décrit l'intérêt de chaque ajout et discute des problèmes qui pourraient être rencontrés dans leur réalisation.



- Chaque rectangle représente un sous-modèle spécifique

9.4. Ajout de la demande de Carburant des Véhicules comme Sous-Modèle

La série de données de l'INSEE fournit une variable potentiellement importante mesurant les dépenses annuelles des ménages en carburant pour les véhicules. En principe, nous pouvons convertir ces données en quantité de carburant, en déflétant par le prix de marché de chaque type de carburant, et estimer un sous-modèle, pour exprimer cette quantité comme une fonction des coûts de fonctionnement de tous les véhicules détenus, qui incorpore la consommation et le prix des carburants, les caractéristiques du ménage (telles que la situation professionnelle et le revenu), et les caractéristiques de la zone géographique (par exemple, la qualité des transports publics). L'avantage d'estimer un tel modèle est qu'il peut prédire une demande agrégée annuelle de carburant, par type, pour des scénarii de base, et par simulations de politiques.

L'approche globale adoptée pour formuler et évaluer ce modèle étendu, est similaire à celle de Train (1986). Cette méthode inclut des sous-modèles expliquant le nombre de kilomètres parcourus par chaque véhicule. Malheureusement, notre modèle serait moins riche que celui de Train (1986), du au fait que les dépenses en carburant sont observées au niveau des ménages et non au niveau des véhicules. Cependant, il serait possible d'estimer une relation entre la demande de carburant et les coûts de fonctionnement de chaque véhicule détenu.

Un des problèmes à prendre ces aspects en compte dans l'estimation du modèle de demande de carburant, et le point de connexion avec les niveaux supérieurs du modèle dans la Graphique 9.1, est le fait que les coûts de fonctionnement sont endogènes étant déterminés en temps que la demande en carburant. Les ménages utilisant potentiellement leur véhicule de manière importante pourraient être tentés d'acquiescer un véhicule consommant moins de carburant, en comparaison à des ménages utilisant peu leur véhicule. Si cette endogénéité n'est pas prise en compte, les coefficients estimés pour les coûts de fonctionnement, qui seront certainement négatifs, seront biaisés à la baisse du fait que les coûts de fonctionnement sont engagés par les ménages qui auraient choisi de conduire moins de kilomètres de toute manière, et également du fait que des coûts de fonctionnement élevés désincitent à consommer du carburant. Heureusement, l'endogénéité des coûts de fonctionnement peut être prise en compte dans la modélisation économétrique en utilisant la méthode des variables instrumentales. Cette approche utilise les variables qui sont corrélées à la variable endogène (les coûts de fonctionnement) mais qui ne sont pas déterminées en même temps que la variable expliquée (la demande de carburant). Pour le cas présent, les instruments appropriés sont les caractéristiques du ménage et de la zone géographique. Avec ces instruments, il serait possible d'estimer une équation robuste expliquant la demande des ménages en carburant de sorte qu'elle soit directement liée à leur choix de véhicules à détenir.

Nous pensons que l'extension décrite ici serait utile à , et nous restons à votre entière disposition pour discuter plus en détail de cette option.

9.5. Ajout de la Détention de Trois Véhicules Comme une Option de Choix

Le modèle principal considère comme choix possible de détenir zéro, un ou deux véhicules dans le sous-modèle de choix de quantité, sans considérer, au moment de l'estimation, les ménages choisissant de détenir trois véhicules ou plus. Un terme est inclus dans la formule pour agréger la demande à la population pour corriger le biais potentiel qui peut-être causé par l'omission des ménages détenant trois véhicules ou plus. Cependant, cette correction est un ajustement plutôt que l'approche idéale car le terme de correction n'est pas prédit par le modèle, et doit donc être supposé pour les années futures. Ceci est une faiblesse de l'approche que la seconde extension cherche à corriger.

Le principal avantage à étendre le modèle pour inclure les ménages ayant trois véhicules est que l'ajustement nécessaire pour la correction du fait de ne pas considérer tous les ménages est moins important. Mais il n'élimine pas la nécessité d'un ajustement car certains ménages détiennent quatre véhicules ou plus. Il y a cependant peu de ménages dans ce cas : alors que 5,2 % de l'échantillon des ménages de l'INSEE possèdent trois véhicules, seulement 1 % des ménages détiennent quatre véhicules ou plus. L'erreur venant d'une hypothèse erronée sur la croissance du nombre moyen de véhicules détenus par les ménages possédant au moins trois véhicules pourraient certainement être considérée comme insignifiante. On ne peut pas forcément en dire autant pour les ménages possédant deux véhicules ou plus.

Les avantages de l'extension du champ d'application est la puissance de modélisation offerte par cette approche, mais en ajoutant de la complexité, et donc nécessitant plus de temps, que la spécification du modèle utilisé dans cette analyse. La raison principale à cet ajout de complexité est le besoin d'étendre le nombre d'options de choix dans les sous-modèles de types de véhicules pour inclure tous les triplés de types de véhicules. Un travail supplémentaire important serait nécessaire afin d'explorer de manière adéquate les différentes options de modélisation disponibles avec un ensemble aussi large de choix, et chaque exécution du modèle prendrait beaucoup plus longtemps que lors de la modélisation des paires de types de véhicules. Nous pensons que ce travail serait intéressant pour le MEDDTL, dans une étude future, pour améliorer la robustesse du modèle.

Références

- Bento A.M., L. H. Goulder, M. R. Jacobsen et R. H. von Haefen, "Distributional and Efficiency Impacts of Increased U.S. Gasoline Taxes," Draft Report, Juillet 2006
- Berkovec, J. et J. Rust, "A nested logit model of automobile holdings for one-vehicle households," *Transportation Research B*, 19 (4), 1984, pp 275-286
- Berry, S., Levinsohn, J., Pakes, A., *Automobile Prices in Market Equilibrium* *Econometrica*, Vol. 63, No. 4, pp. 841-890, Juillet 1995
- Berry, S., Levinsohn, J., Pakes, A., *Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market*, *Journal of Political Economy*, vol 112, no.1, Summer 2004
- Bhat, C.R. et S. Sen, "Household vehicle type holdings and usage: an application of the multiple discrete-continuous extreme value (MDCEV) model," *Transportation Research B*, 40, 2006, pp 35-53
- Bunch, D.S., B. Chen, "Automobile demand and type choice," D.A. Hensher and K. Button (eds), *Handbook of Transport Modelling*, 2ème édition, Elsevier Science, Oxford, 2007
- Bunch, D.S., M. Bradley, T.F. Golob, R. Kitamura et G.P. Occhiuzzo, "Demand for clean fuel vehicles in California: a discrete-choice stated-preference pilot project," *Transportation Research A*, 27A (3), 1993, pp 237-253
- Cambridge Econometrics (CE) et BMRB, "Demand for cars and their attributes," Final Report, Department for Transport, 23 Janvier 2008
- Cecchini, L., "A logit analysis of the car market," *Applied Economics Letters* 8, 2001, pp 199-201
- De Jong, G.C., J. Fox, A. Daly, M. Pieters et R. Smit, "Comparison of car ownership models," *Transport Reviews* 24 (4), 2004, pp 379-408
- De Jong, G.C., "A disaggregated model system of vehicle holding duration, type choice and use," *Transportation Research B*, 30(4), 1996, pp 263-276
- De Jong, G.C., "An indirect utility model of car ownership and private use," *European Economic Review* 34, 1990, pp 971-985
- Eftic, "Demand for cars and their attributes," Final Report, Department for Transport, Janvier 2008
- Greene, D. L., P. D. Patterson, M. Singh, et J. Li, "Feebates, rebates and gas-guzzler taxes: a study of incentives for increased fuel economy," *Energy Policy* 33, 2005, pp. 757-775
- Hensher, D.A., Rose, J.M., Black, I., *Interactive Agency Choice in Automobile Purchase Decisions: The Role of Negotiation in Determining Equilibrium Choice Outcomes*, *Journal of Transport Economics and Policy*, Vol 42, Part 2, May 2008
- Hensher, D.A. et W.G. Greene, "Choosing between conventional, electric and lpg/cng vehicles in single-vehicle households," Hensher, D.A. (eds) *The Leading Edge of Travel Behaviour Research*, Pergamon, Oxford, 2001
- Hensher, D. A. et V. le Plastrier, "Toward a dynamic discrete-choice model of household automobile fleet size and composition," *Transportation Research*, 19B, 1985, pp 481-495
- Lave, C. A., et K.E. Train, "A disaggregate model of auto-type choice," *Transportation Research* 13A, 1979, pp 1-9
- Mannering, F., H. Mahmassani, "Consumer valuation of foreign and domestic vehicle attributes: econometric analysis and implications for auto demand," *Transportation Research A*, 19A (3), 1985, pp 243-251
- Manski, C. F., L. Sherman, "An empirical analysis of household choice among motor vehicles," *Transportation Research*, 14A, 1980, pp 349-366
- McCarthy, P.S. et R.S. Tay, "New vehicle consumption and fuel efficiency: a nested logit approach," *Transportation Research E*, 34(1), 1998, pp 39-51
- Mohammadian, A. et E. J. Miller, "An empirical investigation of household vehicle type choice decisions," *Journal of the Transportation Research Record* 1854, 2003, pp 99-106
- Prieto, M., "Équipement des ménages et comportements de demande de produits différenciés: application des méthodes d'analyse des données et de choix discrets au cas de l'automobile en France," Thèse de Doctorat en Sciences Economiques, Université de Rouen, Faculté de Droit, Sciences Economiques et Gestion, 8 décembre 2006
- Train, K.E. et C. Winston, "Vehicle choice behaviour and the declining market share of U.S. automakers," *International Economic Review*, Vol. 48, No. 4, pp. 1469-1496, 2007
- Train, K.E., "Qualitative Choice Analysis (Theory, Econometrics and an Application to Automobile Demand)," MIT Press, Cambridge MA, 1986
- Verboven, F., Goldberg, P.K., "The Evolution of Price Dispersion in the European Car Market," *Review of Economic Studies*, Vol 68(4), 2004, pp 811-848

ANNEXES

Annexe A Spécification du Modèle Technique

Dans cette annexe, nous exposons les équations constituant le modèle du choix de véhicule décrit de façon non technique en partie 4.1. La présentation ci-dessous provient principalement de Train (1986).

Nous supposons que chaque ménage choisit le nombre de véhicule à détenir (n), leurs type (c_n), et leurs marques/modèles (m_{c_n}) de telle sorte qu'il maximise son utilité, spécifiée:

$$(1) \quad V_{nc_n m_{c_n}} = f(Y, P_{nc_n m_{c_n}}, X_{nc_n m_{c_n}})$$

Où, $V_{nc_n m_{c_n}}$ est l'utilité indirecte du ménage, exprimée en fonction de son revenu (Y), des coûts variables du véhicule ($P_{nc_n m_{c_n}}$) pour chaque véhicule détenu, et une série d'autres variables explicatives ($X_{nc_n m_{c_n}}$), incluant les caractéristiques observés et non-observés des véhicules, qui affectent le niveau d'utilité que le ménage obtiendra à détenir n véhicules de types c_n et de marques/modèles m_{c_n} .

La probabilité jointe qu'un ménage choisisse n^* , c_n^* et $m_{c_n}^*$ est alors:

$$(2) \quad P_{n^* c_n^* m_{c_n}^*} = \Pr(V_{n^* c_n^* m_{c_n}^*} > V_{nc_n m_{c_n}} \text{ for all } n, c_n, m_{c_n} \text{ other than } n^*, c_n^*, m_{c_n}^*)$$

Dans un premier temps, afin d'estimer cette probabilité, nous décomposons l'utilité par l'utilité moyenne des ménages, et un terme d'erreur, comme suit:

$$(3) \quad V_{nc_n m_{c_n}} = \bar{V}_{nc_n m_{c_n}} + e_{nc_n m_{c_n}},$$

Ensuite, nous décomposons l'utilité moyenne, $\bar{V}_{nc_n m_{c_n}}$, en trois composantes:

$$(4) \quad \bar{V}_{nc_n m_{c_n}} = \bar{V}_n + \lambda \bar{V}_{c_n|n} + \theta \lambda \bar{V}_{m_{c_n}|nc_n}$$

où:

$\bar{V}_n$ est la moyenne de l'utilité de détenir n véhicules, (i.e. la moyenne de $\bar{V}_{nc_n m_{c_n}}$ sur tous les types de véhicules et de marques/modèles),

$\bar{V}_{c_n|n}$ est la moyenne de l'utilité, nette de $\bar{V}_n$, de détenir des véhicules de types c_n étant donné que n véhicules sont détenus, (égal à la différence entre $\bar{V}_{nc_n m_{c_n}}$ et $\bar{V}_n$ pour toutes les marques/modèles de véhicules de types c_n),

$\bar{V}_{m_{c_n}|nc_n}$ est l'utilité de détenir une marque/modèle m_{c_n} étant donné que n véhicules de types c_n sont détenus, nette de la moyenne de nombre et de types de véhicules détenus,

, sont les paramètres de la fonction de la Valeur Extrême Généralisée (GEV) supposée caractériser la distribution du terme d'erreur $e_{nc_n m_{c_n}}$:

$$(5) \quad \text{CDF}(e_{nc_n m_{c_n}}, \text{ for all } n, c_n \text{ and } m_{c_n}) = \exp \left\{ - \sum_n \left[\sum_{c_n} \left[\sum_{m_{c_n}} \left[\exp(-e_{nc_n m_{c_n}}) \right]^{1/\theta} \right]^{\theta/\lambda} \right]^\lambda \right\}$$

Le premier des composants de cette équation, $\bar{V}_n$, dépend seulement des facteurs variant selon n .

$\bar{V}_{c_n|n}$ dépend des facteurs qui varient selon n et c_n , mais pas selon m_{c_n} . $\bar{V}_{m_{c_n}|nc_n}$ dépend des facteurs variant selon n , c_n et m_{c_n} .

La probabilité jointe qu'un ménage choisisse n^* , c_n^* et $m_{c_n}^*$ peut-être décomposée comme le produit des probabilités marginales et conditionnelles:

$$(6) \quad P_{n^* c_n^* m_{c_n}^*} = P_{n^*} \cdot P_{c_n^*|n^*} \cdot P_{m_{c_n}^*|n^* c_n^*}$$

où:

P_{n^*} est la probabilité marginale de détenir n^* véhicules de tous types, marques et modèles,

$P_{c_n^*|n^*}$ est la probabilité marginale de détenir des véhicules de type c_n^* de toutes marques et modèles, conditionnellement à détenir n^* véhicules, et

$P_{m_{c_n}^*|n^* c_n^*}$ est la probabilité de détenir $m_{c_n}^*$ marques/modèles conditionnellement au fait de détenir n^* véhicules de types c_n^*

Etant donné la distribution GEV ci dessus de $e_{nc_n m_{c_n}}$, ces probabilités deviennent:

$$(7) \quad P_{n^*} = \frac{\exp(\bar{V}_{n^*} + \lambda I_{n^*})}{\sum_n \exp(\bar{V}_n + \lambda I_n)}$$

$$(8) \quad P_{c_n^*|n^*} = \frac{\exp(\bar{V}_{c_n^*|n^*} + \theta J_{c_n^*})}{\sum_{c_n} \exp(\bar{V}_{c_n|n^*} + \theta J_{c_n})}$$

$$(9) \quad P_{m_{c_n}^*|n^* c_n^*} = \frac{\exp(\bar{V}_{m_{c_n}^*|n^* c_n^*})}{\sum_{m_{c_n}} \exp(\bar{V}_{m_{c_n}|n^* c_n^*})}$$

où:

$$(10) \quad J_{c_n^*} = \ln \sum_{m_{c_n}} \exp(\bar{V}_{m_{c_n}|n^* c_n^*}), \text{ et}$$

$$(11) \quad I_{n^*} = \ln \sum_{c_n} \exp(\bar{V}_{c_n|n^*} + \theta J_{c_n})$$

Les équations (7) et (8) représentent les sous-modèles caractérisant respectivement la quantité de véhicules et les types de véhicules. L'équation (7) montre que la probabilité de choisir n^* véhicules dépend des éléments de $\bar{V}_n$ plus un terme, I_{n^*} , captant l'utilité que le ménage gagneraient par son choix de type de véhicule. De même, l'équation (8) montre que la probabilité de choisir le type de véhicule c_n^* , étant donné qu'il a choisi de détenir n^* véhicules, dépend des éléments de $\bar{V}_{c_n|n^*}$ plus un terme, $J_{c_n^*}$, qui capte l'utilité que le ménage gagneraient par son choix de marque / modèle.

L'équation (9) n'est pas estimée dans notre approche, en raison du niveau de complexité et la difficulté qu'elle introduirait dans l'analyse, pour une faible amélioration, le cas échéant, de l'outil de simulation des politiques. L'impact qu'elle a sur les sous-modèles de choix de type de véhicule et du nombre à acquérir est capturé par le

terme $J_{c_n}^*$, qui doit être inclus dans (8) pour éviter qu'elle soit biaisée. McFadden (1978) a démontré que, sous des hypothèses assez générales, et en particulier lorsque le nombre de marques / modèles augmente, $J_{c_n}^*$ peut être approximé par:

$$J_{c_n}^* \rightarrow \ln(r_{c_n}) + \frac{1}{2} W_{c_n}^2 ,$$

où r_{c_n} est le nombre de marques/modèles pour les véhicules de types c_n , exprimé comme une proportion de toutes les marques et modèles, et $W_{c_n}^2$ est la variance de $\bar{V}_{m_{c_n}|nc_n}$ autour de $\bar{V}_{c_n|n}$. Nous avons adopté cette approximation dans le sous-modèle de choix du type de véhicule.

Annexe B Variables Pertinentes du Série de données de l'INSEE

Noms des Variables	Description
Variables de Choix du Véhicule	
numvehic	Véhicule ID
acqvoi	Neuf / Occasion
anvoi	Année du modèle
recvoi	Année d'achat
codemar	Modèle du véhicule
carbu	Type de carburant du véhicule
Variables Socio-Economiques	
<i>Variables des Ménages</i>	
ident_men	Ménage ID
agpr	Age des personnes de référence au 31 Décembre de l'année du sondage
sexep	Sexe de la personne de référence
couplepr	La personne de référence vit en couple
situapr	Situation professionnel de la personne de référence
codcspr	Groupe socioprofessionnel de la personne de référence
revact	Revenu total du travail (euros)
revtot	Revenu total sans les revenus exceptionnels
rev200_d	Montant des salaires (euros)
npers	Nombre de personnes composant le ménage
nactifs	Nombre de salariés composant le ménage
nenfants	Nombre d'enfants composant le ménage
nbvehic	Nombre de véhicule détenu par le ménage
zeat	Région de résidence
dip14pr	Niveau de qualification le plus élevé de la personne de référence du ménage
<i>Variables Individu</i>	
ident_men	Ménage ID
ident_ind	Individu ID
ag	Age au 31 Décembre de l'année du sondage
sexe	Sexe
couple	Vit en couple
natio7	Nationalité
zeat	Région de résidence
dip14	Niveau de qualification le plus élevé

Variables d'Utilisation du Véhicule

persvoi	Le véhicule a un utilisateur principal
quivoi	ID de l'utilisateur principal
expvoi1	Véhicule utilisé pour se rendre au travail
protrav	Proportion des journées d'utilisation pour les déplacements (%)
expvoi2	Véhicule utilisé pour les voyages d'affaires
provoi	Proportion des journées d'utilisation pour les voyages d'affaires (%)
expvoi3	Véhicule utilisé pour d'autres raisons

Variable Name**Label****Variables de Financement**

repvoi_d	Montant provenant de la vente de l'ancien véhicule (ajusté)
apvoi	Mode d'achat (Location-vente/gratuit/acheté par une personne du ménage)
apervoi_d	Montant de la contribution personnelle dans l'achat du véhicule (ajusté)
crevoi	Utilisation d'un crédit pour l'achat du véhicule
empvoi_d	Montant emprunté pour l'achat (ajusté)
mcrevoi_d	Montant annuel de remboursement du crédit (ajusté)
pervoi	Périodicité du remboursement du crédit
assvoi	Si l'assurance a été payée pour les 12 derniers mois
tassvoi	Type d'assurance
massvoi_d	Montant des dépenses annuelles dans l'assurance du véhicule (ajusté)
passvoi	Périodicité des paiements de l'assurance en mois

Variables de Dépense en Carburant

C072211	Super, ordinaire, super sans plomb, mélange 2T
C072212	Gaz oil, diesel
C072213	GPL
C072214	Electricité

Annexe C Caractéristiques de l'échantillon de l'INSEE

	Fréquence	Pourcentage
Nombre de Véhicules Détenus		
0	1 567	15.30
1	4 700	45.90
2	3 331	32.53
3	536	5.23
4	93	0.91
5	9	0.09
6	3	0.03
9	1	0.01
Taille de l'échantillon (nombre de ménages)	10 240	100
Utilisation du Véhicule		
Véhicule utilisé pour les déplacements	6 941	51.7
Véhicule utilisé pour les affaires	1 321	9.85
Véhicule utilisé pour d'autres trajets	10 553	78.7
Taille de l'échantillon (nombre de véhicules)	13 414	100
Age du Véhicule		
Pre-1986	533	4.06
1986-1990	1 557	11.87
1991-1995	3 041	23.19
1996	751	5.73
1997	672	5.12
1998	810	6.18
1999	817	6.23
2000	1 094	8.34
2001	914	6.97
2002	874	6.66
2003	858	6.54
2004	794	6.05
2005	394	3.00
2006	5	0.04
Taille de l'échantillon (nombre de véhicules)	13 114	100
Localisation des Ménages		
Région parisienne	1 803	17.61
Bassin parisien	1 852	18.09
Nord	665	6.49
Est	971	9.48
Ouest	1 544	15.08
Sud-ouest	1 098	10.72
Centre-Est	1 108	10.82

	Fréquence	Pourcentage
Méditerranée	1 199	11.71
Taille de l'échantillon	10 240	100.00
Taille du Ménage (Nombre de Personnes)		
1	2 717	26.53
2	3 499	34.17
3	1 607	15.69
4	1 552	15.16
5	654	6.39
6	165	1.61
7	26	0.25
8	10	0.10
9	6	0.06
10	1	0.01
11	2	0.02
12	1	0.01
Taille de l'échantillon	10 240	100.00
Age de la personne de référence		
<21	129	1.26
21-30	1 253	12.24
31-40	2 094	20.45
41-50	2 083	20.34
51-60	1 932	18.87
61-70	1 257	12.28
71-80	1 075	10.50
>80	417	4.07
Taille de l'échantillon	10 240	100.00
Niveau d'Éducation de la Personne de Référence		
Diplôme de 3e cycle universitaire, doctorat	16	0.16
Diplôme d'ingénieur, d'une grande école	97	0.95
Diplôme de 2e cycle universitaire	1 582	15.45
Diplôme de 1er cycle universitaire	102	1.00
BTS, DUR or équivalent	665	6.49
Diplôme des professions sociales et de la santé de niveau BAC+2	126	1.23
Baccalauréat général, brevet supérieur, capacité en droit, DAEU	697	6.81
Baccalauréat technologique	340	3.32
Baccalauréat professionnel	220	2.15
Brevet professionnel ou de technicien	127	1.24
CAP, BEP ou diplôme du même niveau	2 763	26.98
Brevet des collèges, BEPC	683	6.67

	Fréquence	Pourcentage
Certificat d'Etude Primaire	1 258	12.29
Sans diplôme	1 564	15.27
Taille de l'échantillon	10 240	100.00
Statut Professionnel de la Personne de Référence		
Salarié	6 304	61.56
Apprenti	22	0.21
Etudiant	187	1.83
Sans emploi	460	4.49
Retraité	2 901	28.33
Femme/home au foyer	168	1.64
Autre (e.g. personne handicapée)	198	1.93
Taille de l'échantillon	10 240	100.00
Groupe Socioprofessionnel de la Personne de Référence		
Agriculteurs	189	1.85
Artisans	202	1.97
Petits détaillant	153	1.49
Chefs d'entreprises de 10 salariés ou plus	32	0.31
Professions libérale (professionnels indépendants)	88	0.86
Agents de la fonction publique, Professeurs du secondaire/université, media et divertissement	477	4.66
Gestionnaires	782	7.64
Professeurs des écoles, travailleurs médicaux et sociaux	599	5.85
Cadres commerciaux et d'administration	490	4.79
Techniciens	401	3.92
Contremaîtres et superviseurs	225	2.20
Ouvriers du service civil, policiers et militaires	715	6.98
Employés de bureau	294	2.87
Assistants de boutique	182	1.78
Aide à domicile	288	2.81
Ouvriers qualifiés	1,258	12.29
Ouvriers non qualifiés	386	3.77
Fermiers	59	0.58
Anciens fermiers	228	2.23
Anciens artisans, petits commerçants et chefs d'entreprises	268	2.62
Anciens cadres et cadres supérieurs	1 033	10.09
Anciens travailleurs et superviseurs	1 426	13.93
Sans emploi (jamais travaillé)	16	0.16
Autres inactifs (autre que retraité)	445	4.35
Non spécifié	4	0.04
Taille de l'échantillon	10 240	100.00

	Fréquence	Pourcentage
Revenu Annuel Total du Ménage (€/an)		
<10k	717	7.00
10k-20k	2 539	24.79
20k-30k	2 416	23.59
30k-40k	1 951	19.05
40k-50k	1 105	10.79
50k-70k	974	9.51
70k-100k	358	3.50
>100k	180	1.76
Taille de l'échantillon	10 240	100.00

Annexe D Variables du série de données JATO

Date d'Enregistrement:

Date

Détails des Voitures:

Année du modèle

Marque

Modèle

Version

Etat de la voiture

Type de carrosserie

Nombre de portes

Type de carburant

Segment du Marché

Détails des Moteurs

Taille du Moteur (l)

Puissance (hp/PS)

0-100km/h (secs)

Roues Motrices

Type de Transmission

Nombre de vitesses

ABS

Réservoir

Capacité (l)

Dimensions:

Poids à vide (kg)

Longueur (m)

Largeur (m)

Hauteur (m)

Détails Efficacité/Pollution:

Contrôle du Niveau d'Emission

Norme respectée

CO₂g/km combiné

Consommation d'Essence

Standard

Combiné (mpg)

Combiné (l/100km)

Détails de l'Etat de la voiture:

Air conditionné

Vitres électriques

Type de jantes

Airbag frontal

Airbag latéral

Airbag supérieurs

Détails sur les coûts:

Prix

Assurance

Description

Liste des Tableaux

Tableau 3.1 Résumé des Principales Études.....	19
Tableau 5.1 Taille des Echantillons et Cadres d'Etudes Précédemment Réalisées	25
Tableau 5.2 Variables Relatives aux Ménages Utilisées dans les Etudes Précédentes et Disponibles dans le set de l'INSEE	26
Tableau 5.3 Caractéristiques des Véhicules Utilisés dans les Etudes Précédentes et celles Disponibles dans la série de données JATO.....	27
Tableau 5.4 Variables Relatives au Marché Utilisées dans des Etudes Précédentes et Disponibles dans l'Etude Présente	28
Tableau 5.5 Variables Définissant les "Types de Véhicules"	29
Tableau 5.6 Exemple de Véhicules par Catégorie.....	30
Tableau 5.7 Données de l'INSEE Relatives à la Qualité	31
Tableau 5.8 Caractéristiques des Véhicules Disponibles en 2006.....	35
Tableau 7.1 Sous-Modèle de Choix de Véhicules pour les Ménages Ayant un Véhicule	46
Tableau 7.2 Seuil de Rentabilité du Véhicule selon les Distances Parcourues Annuellement, Induit par les Coefficients Estimés du Prix et du Coût Variable.....	47
Tableau 7.3 Sous-Modèle du Choix du Type de Véhicules pour les Ménages à Deux Véhicules	49
Tableau 7.4 Sous Modèle du Choix du Nombre de Véhicules.....	51
Tableau 7.5 Elasticités de la Demande au Prix d'Achat, par Catégorie de Véhicule.....	53
Tableau 7.6 Elasticités de la Demande au Prix d'Achat, par Groupe de Véhicules de Différents Millésimes	53
Tableau 7.7 Elasticités de la Demande pour les Nouveaux Véhicules, par Groupe de Véhicules Selon leurs Emissions de CO2 (g/km).....	54
Tableau 7.8 Résumé des Résultats sur les Elasticités des Principales Etudes	55
Tableau 8.1 Bonus écologique.....	56
Tableau 8.2 Eléments renseignés pour les besoins de la Simulation.....	57
Tableau 8.3 Caractéristiques des Véhicules Disponibles en 2006 et 2009	59
Tableau 8.4 Résultats des Simulations des Scenarii de Politiques (Proportions)	60
Tableau 8.5 Résultats des Simulations des Scenarii de Politiques (Nombre).....	61
Tableau 8.6.....	63

Liste des graphiques

Graphique 4.1 Modèle de Choix de Véhicule des Ménages*	22
Graphique 5.1 Lien entre les Séries de données Relatifs au Ménage et aux Types de Véhicules.....	29
Graphique 5.2 Groupement en Catégorie et Fourchette d'Emissions de CO2 des Données de l'INSEE Cas 1: Information Complète de l'INSEE sur la Marque, le Modèle, le Millésime et le Type de Carburant.....	31
Graphique 5.3 Caractéristiques de l'Echantillon sur les Ménages	33
Graphique 5.4 (cont.) Caractéristiques de l'Echantillon sur les Ménages.....	34
Graphique 6.1 Proportion de Véhicules Détenus, par Catégorie de Véhicule	37
Graphique 6.2 Proportion de Véhicules Détenus, par Classe d'Emissions de CO2	38
Graphique 6.3 Proportion de Véhicules Détenus, par Millésime de Véhicule.....	38
Graphique 6.4 Proportion de Véhicules Détenus par Catégorie et par Millésime.....	39
Graphique 6.5 Proportion de Véhicules Détenus par Classe d'Emission de CO2 et par Millésime.....	40
Graphique 6.6 Proportion de Véhicules Détenus par Classe d'Emission de CO2 et par Catégorie	40
Graphique 6.7 Proportion de Véhicules Détenus par Catégorie et par Revenu Annuel des Ménages.....	42
Graphique 6.8 Proportion de Véhicules Détenus par Catégorie et par Age du Conducteur Principal	42
Graphique 6.9 Proportion de Véhicules Détenus par Catégorie et selon le Sexe du Conducteur Principal	42
Graphique 6.10 Proportion de Véhicules Détenus par Catégorie et selon le Nombre d'Enfants dans le Ménage.....	44
Graphique 6.11 Proportion de Véhicules Détenus par Catégorie et par Type de Régions de Résidence.....	44
Graphique 8.1 : Résultats des Simulations selon les Scenarii de Politiques.....	62
Graphique 9.1 Extensions Potentielle du Modèle de Choix de Véhicule d'un Ménage*	66

Commissariat général au développement durable

Service de l'économie, de l'évaluation et de l'intégration du développement durable

Tour Voltaire

92055 La Défense cedex

Tél : 01.40.81.21.22

Retrouver cette publication sur le site :

<http://www.developpement-durable.gouv.fr/developpement-durable/>, rubrique « Publications »

Résumé

Cette étude réalisée par la société NERA pour le compte du CGDD/MEDDTL vise à développer un modèle permettant de simuler les effets de différentes politiques telles que le « bonus écologique » octroyé aux véhicules à faibles émissions de CO₂, sur le choix de véhicule des ménages et sur la demande globale de véhicules.

Le modèle développé est un modèle de choix discret de type « logit hiérarchisé » qui permet de représenter les différentes phases de choix de possession (et pas seulement d'achat) de véhicules par les ménages : (i) le nombre de véhicules à posséder, et (ii) la catégorie, le millésime, le type de carburant et les émissions de CO₂ de chaque véhicule possédé. Le modèle a été calibré principalement sur la base de données de l'INSEE « Budget des Familles 2006 » (BdF 2006), ainsi que sur la base JATO pour les caractéristiques techniques des véhicules et l'Argus pour les prix des véhicules d'occasion.

Les variables supposées influencer chaque décision apparaissent avec des coefficients de signes anticipés, statistiquement significatifs, et d'un ordre de grandeur plausible. Ainsi, il ressort par exemple que les ménages sont sensibles aux prix d'achat et aux coûts variables des véhicules, et d'autant plus sensibles que leurs revenus sont faibles ; ils sont également d'autant plus sensibles à la taille du véhicule que la taille du ménage est grande.

Les élasticités au prix (directes et croisées) ont été évaluées par catégorie de véhicule, millésime et tranches d'émissions de CO₂. Ainsi, les élasticités-prix augmentent, en valeur absolue, pour les catégories de A (économique, élasticité de -1,15) à E (luxe, -4,07), avec une élasticité particulièrement élevée pour les voitures de sport (-7,67). Les élasticités-prix propres de la demande diminuent avec l'âge du véhicule ; la demande de nouveaux véhicules étant fortement élastique au prix (de l'ordre de -5,32) alors que celle des véhicules de plus de 5 ans est assez inélastique (-0,42). Enfin, les élasticités-prix propres ont tendance à augmenter avec la quantité de CO₂ émise, de l'ordre de -4 pour les véhicules faiblement émetteurs, elles dépassent -20 pour les véhicules émettant 201 g/km de CO₂ ou plus.

La principale limite du modèle est qu'il n'intègre pas l'évolution des performances des véhicules proposés aux ménages, autrement dit l'offre des constructeurs automobile dont la sensibilité à une politique de type bonus écologique est réelle.



Dépôt légal : Janvier 2011
ISSN : 2102 - 4723