

Direction de la Diffusion et de l'Action régionale

H2012/03

**Estimations régionales dans l'enquête nationale
Transport et Déplacements 2007-2008**

Josiane Le Guennec

Document de travail



Institut National de la Statistique et des Études Économiques

Institut National de la Statistique et des Études Économiques
N°H2012/03

Série des documents de travail
de la Direction de la Diffusion et de l'Action Régionale

Estimations régionales dans l'enquête nationale
Transport et Déplacements 2007-2008

JOSIANE LE GUENNEC (*)

Document de travail
Avril 2012

**Au moment de la rédaction de ce document, Josiane Le Guennec faisait partie de la Direction régionale de l'INSEE de Bretagne. L'auteur remercie Jimmy Armoogum et Pascal Ardilly pour leurs avis et conseils.*

Ces documents de travail ne reflètent pas la position de l'INSEE et n'engagent que leurs auteurs.
Working-papers do not reflect the position of INSEE but only their authors'views.

Résumé

L'enquête nationale sur les Transports et Déplacements a été réalisée sur une période de 12 mois, de 2007 à 2008. Conçue pour obtenir une information sur l'ensemble du territoire national, elle ne permet pas des estimations sur des niveaux géographiques plus fins par exploitation directe, en raison de tailles d'échantillons régionaux insuffisantes. Cette étude teste plusieurs méthodes d'estimation par régression sur « petits domaines », utilisant une information auxiliaire connue dans chaque région grâce au recensement de population. La méthode fondée sur une régression mixte dans laquelle le lieu de résidence est spécifié en effet aléatoire s'avère celle qui réduit le mieux le risque de biais de modélisation.

Mots-clés : enquête, transport, méthode et outil, région, conditions de vie

Summary

Regional Estimates in the National Survey on Transport and Travel

The national survey on transport and travel was conducted during 12 months partially covering the years 2007 and 2008. Since it was designed to obtain information concerning the whole national territory, a direct processing of this survey cannot provide estimates concerning smaller areas according to the occasionally insufficient size of regional samples. This study tests several estimation methods through regressions on “small areas”. Such methods use additional information on each region which is derived from the population census. The method based on a mixed regression in which the place of residence is specified as a random effect turns out to be the more appropriate to reduce the risk of model bias.

Enquête nationale sur les transports et déplacements 2007-2008

Estimations régionales

1	L'enquête Transport de 2007	7
1.1	<i>Le questionnaire</i>	7
1.2	<i>Le plan de sondage</i>	7
1.3	<i>Les échantillons régionaux</i>	9
2	Les objectifs de l'étude	10
2.1	<i>La problématique du transport en région</i>	10
2.2	<i>Les indicateurs régionaux mesurés</i>	10
3	Les déplacements : un phénomène déterminé	11
3.1	<i>Quelques facteurs de la mobilité locale</i>	11
3.2	<i>Le modèle complet</i>	15
3.3	<i>La construction des marges régionales</i>	21
4	Les méthodes d'estimation sur « petits domaines »	22
4.1	<i>Précision d'un estimateur de domaine avec le sondage aléatoire simple</i>	23
4.2	<i>Les grands types de modèles</i>	24
5	Les méthodes d'estimation testées	27
5.1	<i>Les variables estimées</i>	27
5.2	<i>Traitement de la non-réponse totale</i>	28
5.3	<i>Les modalités de calage</i>	28
6	Les résultats comparés	29
6.1	<i>Comparaison des méthodes, à modèle identique</i>	29
6.2	<i>Résultats avec un modèle détaillé</i>	31
6.3	<i>Conclusion</i>	32
	Annexe 1 : effectif de l'échantillon par région.....	70
	Annexe 2 : population estimée au 1er janvier 2008 par région	71
	Annexe 3 : modèle simplifié du calage régional, nomenclature des variables de calage	72
	Annexe 4 : modèle détaillé, nomenclature des variables explicatives.....	73
	Annexe 5 : estimateur direct par calage sur des totaux régionaux	75
	Annexe 6 : les estimateurs régionaux initiaux, par calage, modélisés.....	76
	Annexe 7 - Estimation de la variance de l'estimateur direct.....	110

1 L'enquête Transport de 2007

Tous les dix ans environ, le Ministère de l'Équipement effectue une enquête nationale sur la pratique des déplacements et l'usage des moyens de transport par la population. La dernière édition de cette enquête, réalisée de la mi-2007 à la mi-2008 en six vagues, en collaboration entre le SOES et l'INSEE, est désormais à disposition des utilisateurs. Elle succède à l'enquête de 1993-1994. Conçue pour fournir des résultats nationaux, elle ne permet pas des estimations régionales aussi précises, en raison de tailles d'échantillon généralement trop faibles à ce niveau. C'est pourquoi l'on a tenté de mettre en œuvre des méthodes d'estimation fondées sur des modèles, afin d'améliorer la fiabilité des estimations dans des sous populations restreintes.

L'enquête s'est déroulée en deux visites. Pendant la semaine séparant les deux visites, un carnet relatant tous ses déplacements était renseigné par une personne du ménage, ainsi qu'un carnet décrivant l'usage d'un véhicule du ménage.

1.1 Le questionnaire

Le questionnaire de la première visite comprend un tronc commun à tous les ménages interrogés, permettant de décrire les caractéristiques socio démographiques du ménage, son logement dans son environnement, le parc de véhicules à sa disposition. Il décrit ensuite les déplacements contraints quotidiens, du domicile au lieu de travail, au lieu d'études ou au lieu de garde pour les jeunes enfants, de chaque individu du ménage.

Le questionnaire de deuxième visite était renseigné par un seul individu sélectionné au hasard dans le ménage, parmi ses membres âgés de 6 ans ou plus. A l'aide du carnet renseigné entre les deux visites, il décrit tous les déplacements de la personne pendant un jour de la semaine choisi au hasard, du lundi au vendredi, et tous ses déplacements effectués pendant le week-end de la semaine de référence.

Le dernier volet du questionnaire de deuxième visite décrit les voyages de longue distance (plus de 80 km à vol d'oiseau) effectués au cours des 13 semaines précédant l'enquête.

Le carnet véhicule renseigné entre les deux visites a également servi de support à la description de tous les déplacements effectués avec l'un des véhicules du ménage.

On dispose ainsi d'une information sur la propension générale à la mobilité : part de la population qui se déplace régulièrement en semaine, le week-end, ou au contraire ne se déplace jamais, sur les moyens de transport utilisés, les distances parcourues, les durées des déplacements liés aux activités quotidiennes (travail, études, achats, autres activités) en semaine, aux loisirs le week-end, aux voyages occasionnels dans l'année, et sur l'état du parc de véhicules utilisés par les ménages.

1.2 Le plan de sondage

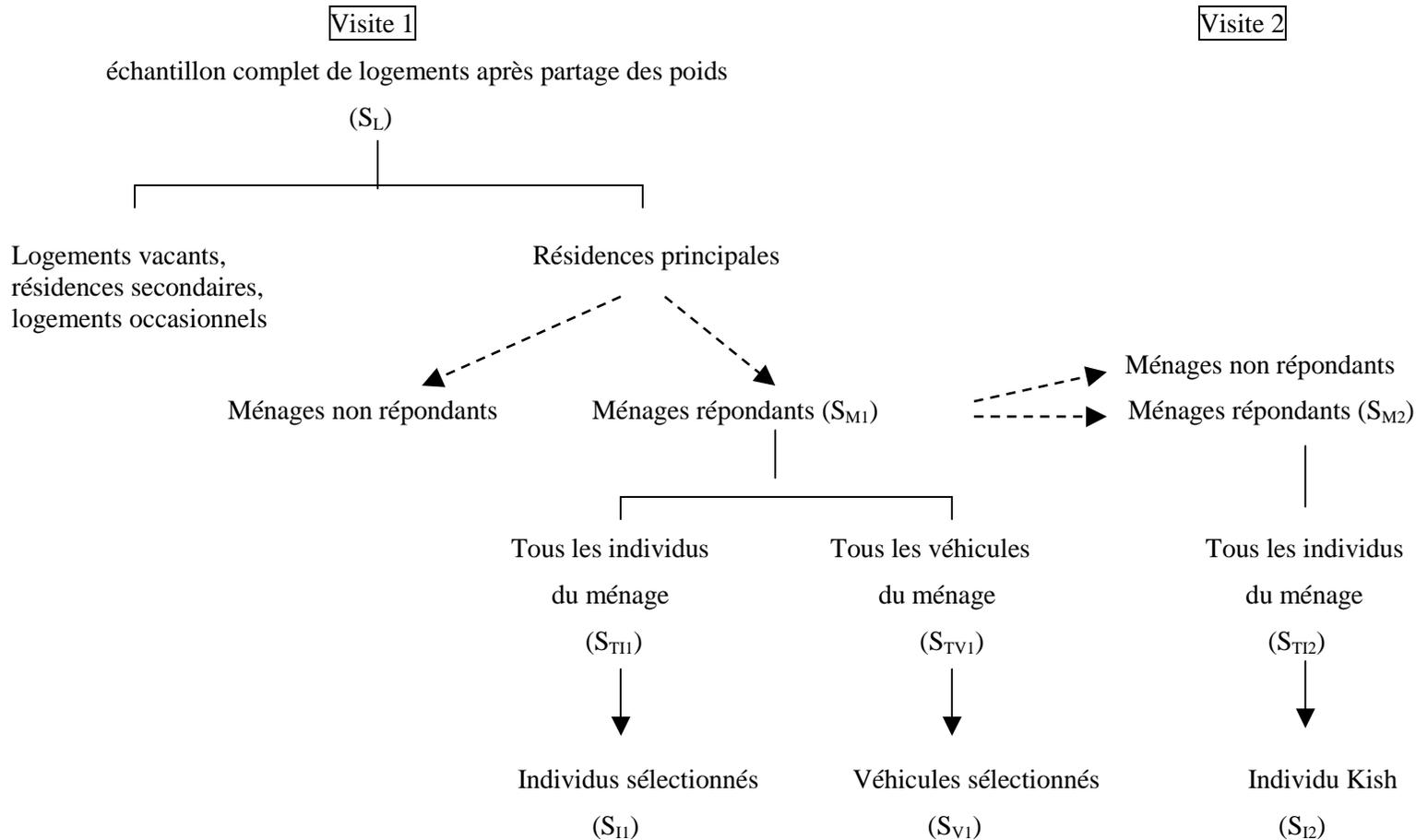
Comme la plupart des enquêtes de l'INSEE auprès des ménages, l'enquête Transport est le résultat d'un échantillonnage de logements à plusieurs degrés, dans lequel le premier degré est constitué d'un échantillon aréolaire de communes, appelé « échantillon-maître ». Celui-ci a été sélectionné en 1999, et utilisé pendant toute la période intercensitaire séparant le dernier recensement général de population du premier recensement rénové publié en 2009. Dans cinq régions ayant financé des renforcements d'échantillon, les logements ont été tirés dans un échantillon de communes élargi, l'EMEX, complémentaire de l'échantillon maître national pour les extensions régionales. Rappelons que l'échantillon maître (EM) et l'EMEX sont des échantillons aréolaires de zones du territoire, stratifiées par grande région¹ et tranche de taille des unités urbaines. Selon la strate, une unité primaire est un canton rural, une unité urbaine complète, un îlot d'une agglomération de plus de 100000 habitants.

Le tirage des logements constitue le deuxième degré de l'échantillonnage national, qui a mobilisé trois bases de sondage : le recensement de population (RP) de 1999, et pour les logements construits depuis,

¹ Huit zones avaient été constituées, par regroupement de régions limitrophes, lors du tirage de l'échantillon maître de 1999.

Figure 1. L'architecture de l'enquête

Conditionnellement à l'échantillon initial de logements, le déroulement de l'enquête peut être représenté selon les phases suivantes.



Lecture

Chaque flèche indique une procédure de sélection : \longrightarrow déterminée par le plan de sondage

\dashrightarrow déterminée par le comportement de réponse

la base de sondage des logements neufs (BSLN) dans les zones appartenant à l'échantillon maître national, la base de logements neufs (BLN) extraite du fichier SITADEL des permis de construire dans les communes de l'EMEX.

Conditionnellement à l'échantillon-maître et à l'EMEX, le tirage des logements a été réalisé en deux phases, afin de permettre une surreprésentation des ménages ayant plusieurs voitures en 1999 et une surreprésentation du milieu rural. Comme dans la plupart des enquêtes ménages de l'INSEE, on a également appliqué un coefficient de sous-représentation aux logements recensés en 1999 et déclarés alors résidences secondaires ou occasionnelles.

En plus d'une extension régionale dans l'EMEX, la région Pays de la Loire a également renforcé son échantillon dans le département de Loire Atlantique, avec une stratification visant une représentativité du bassin de Nantes-Saint-Nazaire. Cette extension locale a été sélectionnée selon un plan de sondage à un seul degré, par tirage direct des logements dans le recensement de 1999 et la BLN.

Au total, 29874 logements ont été sélectionnés dans l'échantillon initial, en France métropolitaine. Après élimination des logements hors champ (résidences secondaires ou vacantes au moment de l'enquête, logements disparus depuis la mise à jour de la base de sondage) et des logements non répondants, et ajout des logements résultant d'un éclatement, on a recueilli les réponses d'environ 20100 ménages en première visite, dont 18632 ont accepté de poursuivre l'enquête en deuxième visite.

En visite 1, le module « déplacements » concernait tous les individus du ménage. Cependant, dans les ménages comprenant plus de deux enfants de 6 ans ou plus inactifs scolarisés, deux d'entre eux seulement, sélectionnés par sondage aléatoire simple, répondaient à l'enquête.

A la fin de la première visite, un individu, dit individu « Kish », était sélectionné parmi les membres de 6 ans ou plus du ménage, avec une probabilité variable selon le nombre de déplacements de plus de 100 km réalisés au cours des trois mois précédant la première visite. Un carnet de déplacements lui était remis, à renseigner, pendant la semaine suivant le passage de l'enquêteur, par tous ses déplacements réalisés dans l'intervalle. Il lui était également demandé de remplir un carnet-véhicule, à chaque déplacement réalisé avec le véhicule du ménage, ou l'un des véhicules sélectionné au hasard dans les ménages en ayant plusieurs. Dans ce dernier cas, la sélection s'est faite avec des probabilités variables selon le type de véhicule (voiture ou deux-roues) et sa fréquence d'utilisation au cours des deux mois précédant l'enquête.

Ces deux carnets servaient de support à la réponse au questionnaire de deuxième visite, renseigné par l'individu Kish unique dans chaque ménage ayant répondu en première visite.

1.3 Les échantillons régionaux

Le nombre de ménages ayant répondu aux deux visites, par région, est indiqué en annexe 1. En Ile de France et en Pays de la Loire, où l'échantillon dépasse 2500 logements, l'estimateur habituel « Horvitz-Thomson », obtenu après sommation des observations dilatées par les poids de sondage redressés, a une précision suffisante, dès lors qu'on ne s'intéresse pas à des catégories trop fines.

C'est également le cas de quelques grands indicateurs synthétiques dans les trois autres régions ayant bénéficié d'une extension d'échantillon (Languedoc-Roussillon, Bretagne, Midi-Pyrénées), où le nombre de ménages répondants avoisine ou dépasse le millier. Dans ces cinq régions, les coefficients de variation estimés restent inférieurs à 10 % pour un certain nombre de résultats-clés, tels le nombre moyen de déplacements par personne et la part modale des déplacements quotidiens.

Il n'en va pas de même dans les autres régions, où l'on ne dispose que de quelques centaines de ménages répondants. Estimer la part modale des déplacements domicile-travail ou domicile-études, par exemple, devra, dans ces régions, faire appel à des méthodes spécifiques pour réduire autant que faire se peut, la variance du résultat.

2 Les objectifs de l'étude

2.1 La problématique du transport en région

Les collectivités territoriales sont très impliquées dans l'organisation des schémas de transport en région, dont elles assument une part importante du financement. La prise de conscience de la nécessité de maîtriser la consommation d'énergie autant que les émissions de dioxyde de carbone dans l'environnement, conduisent peu à peu à encourager l'usage des transports en commun en substitution à la voiture individuelle. Nombre de grandes agglomérations développent des infrastructures de transport réputées peu polluantes (tramways, métros) et adoptent des mesures destinées à limiter l'usage de la voiture en centre-ville. Le financement par les conseils régionaux de tarifs subventionnés en faveur des utilisateurs réguliers des compagnies de transport de voyageurs (SNCF sur les lignes de TER, autobus interurbains) va également dans ce sens.

La tendance à l'étalement urbain, lié à la concentration des emplois en milieu urbain et à la recherche de prix du foncier plus abordables, pèse à la fois sur la facture énergétique et sur les conditions de vie des ménages : allongement des distances parcourues dans les déplacements domicile-travail, du temps passé quotidiennement dans les transports. La fréquence des déplacements par mode de transport, par motif, leur durée moyenne, peuvent donc légitimement apparaître comme des indicateurs-clés dans l'appréciation de l'efficacité des politiques locales en œuvre ou dans l'orientation des choix à venir.

2.2 Les indicateurs régionaux mesurés

Le champ de l'étude est constitué de la mobilité locale en semaine, du lundi au vendredi, motivée par les activités de la vie quotidienne. Pour cette raison, les informations de l'enquête utilisées ont été limitées aux déplacements effectués dans un rayon de 80 km autour du domicile par l'individu Kish interrogé, pendant la journée sélectionnée dans la semaine séparant les deux visites², dit « jour de référence ». La notion retenue du domicile inclut le domicile familial principal occupé la majorité de l'année, mais aussi le domicile occasionnel occupé pour son travail par un adulte actif exerçant son emploi à une distance trop importante de son lieu de résidence familial pour y revenir chaque jour.

L'étude vise à estimer quelques indicateurs clés de cette mobilité par région. Les domaines sont constitués de chacune des régions métropolitaines. On a donc 22 domaines.

Les variables estimées sont les suivantes :

- population mobile le jour de semaine de référence : effectif et pourcentage
- fréquence des déplacements quotidiens un jour de semaine :
 - nombre total de déplacements par personne (population totale au dénominateur)
 - nombre de déplacements par mode de transport : voiture particulière, transport en commun, marche, autre mode
 - nombre de déplacements par motif
- nombre de personnes s'étant rendues au travail (ou au lieu d'étude), par mode de transport
- distance parcourue pendant la journée au cours des déplacements :
 - total journalier par personne
 - distance moyenne d'un déplacement par mode de transport
- temps passé quotidiennement dans les déplacements :
 - total journalier par personne
 - durée moyenne d'un déplacement par mode de transport

² Voir en annexe 7 le détail du plan de sondage.

L'estimation directe de ces variables, à l'aide des pondérations nationales, fait apparaître des différences entre les régions attribuables, dans des proportions variables, à des différences réelles de comportements, aux différences de structure de la population, mais aussi à l'aléa de sondage, d'autant plus élevé que le nombre de ménages répondants est plus faible. C'est l'objet de la modélisation de réduire autant que possible la largeur de cette « fourchette d'estimation » afin de ne plus laisser subsister que les écarts réels de niveaux dans les estimations régionales.

3 Les déplacements : un phénomène déterminé

Du lundi au vendredi, la journée de semaine est rythmée par le travail, celui des adultes actifs comme le travail étudiant des jeunes scolarisés, mais aussi par les nécessités de l'approvisionnement domestique, par les activités de loisir ou associatives, sans oublier les démarches diverses. Autant de raisons de quitter le domicile, qui varient selon l'âge, la situation d'activité, le lieu de résidence, et d'autres facteurs qui ne doivent pas tout au hasard. Les premiers résultats nationaux ont montré l'influence de l'âge sur la fréquence des déplacements par individu, du mode d'urbanisation sur les distances parcourues et l'usage des divers modes de transport. On peut donc s'attendre à pouvoir expliquer certaines différences interrégionales en matière de déplacements par celles des structures sociodémographiques et des répartitions de population dans l'espace.

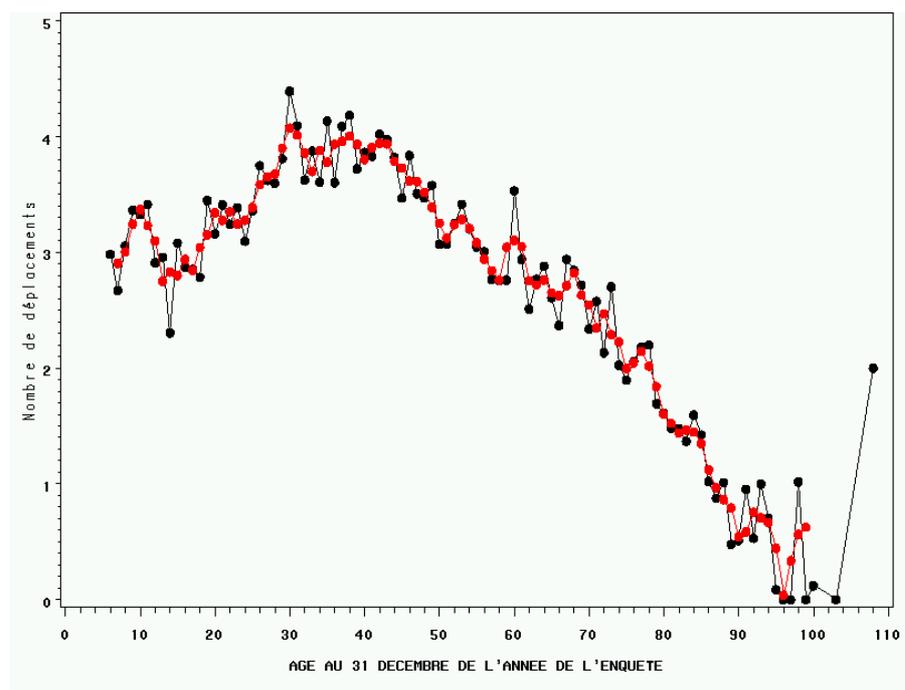
3.1 Quelques facteurs de la mobilité locale

L'usage des transports en commun et la durée passée dans les déplacements opposent très nettement l'Ile de France à la province. C'est en région parisienne que le recours aux transports collectifs est le plus généralisé : 20 % des déplacements se font par ce moyen, contre à peine 7 % dans l'ensemble de l'hexagone, et que les temps de transport domicile-travail sont les plus longs. C'est pourquoi, selon les critères, l'on distinguera la France de province et l'Ile de France dans ce qui suit.

3.1.1 La fréquence des déplacements selon l'âge

Le changement de cycle scolaire jusqu'à 25 ans, l'entrée dans la vie active entre 25 et 30 ans, le passage de la cinquantaine précédant la sortie de la vie active à 60 ans, puis le vieillissement marquent autant de paliers dans la mobilité individuelle. La courbe du nombre journalier de déplacements par personne et par âge détaillé à partir de 6 ans a conduit à retenir, dans les modèles d'analyse de la variance, les tranches d'âge suivantes : 6-13 ans, 14-18 ans, 19-25 ans, 26-29 ans, 30-39 ans, 40-49 ans, 50-60 ans, 61-69 ans, 70-79 ans, 80 ans et plus.

Figure 2 - Nombre de déplacements par personne selon l'âge



Lecture : la courbe rouge est le résultat d'un lissage en moyennes triennales
 Champ : France métropolitaine, population âgée de 6 ans ou plus

3.1.2 *Les déplacements selon la zone de résidence*

La zone de résidence, caractérisée par la taille de l'agglomération et sa proximité par rapport au centre ville, influence la fréquence des déplacements journaliers, mais aussi les distances parcourues et les temps de transport (figure 3).

Si on se déplace plus souvent dans les centres des grandes agglomérations -- de plus de 50000 habitants -- que dans leur périphérie, en milieu urbain qu'en milieu rural, les distances parcourues quotidiennement sont plus élevées en zone rurale et dans les couronnes périurbaines que dans les centres-villes, et décroissent avec la taille de l'agglomération. Ces comparaisons ne valent que pour les déplacements dont les points de départ et d'arrivée sont situés dans des communes distinctes, les distances étant ici mesurées à vol d'oiseau entre les centres des communes. L'Ile de France s'oppose très fortement au reste du pays par la durée totale passée dans les transports et déplacements un jour de semaine, très élevée en région parisienne.

Type de commune de résidence

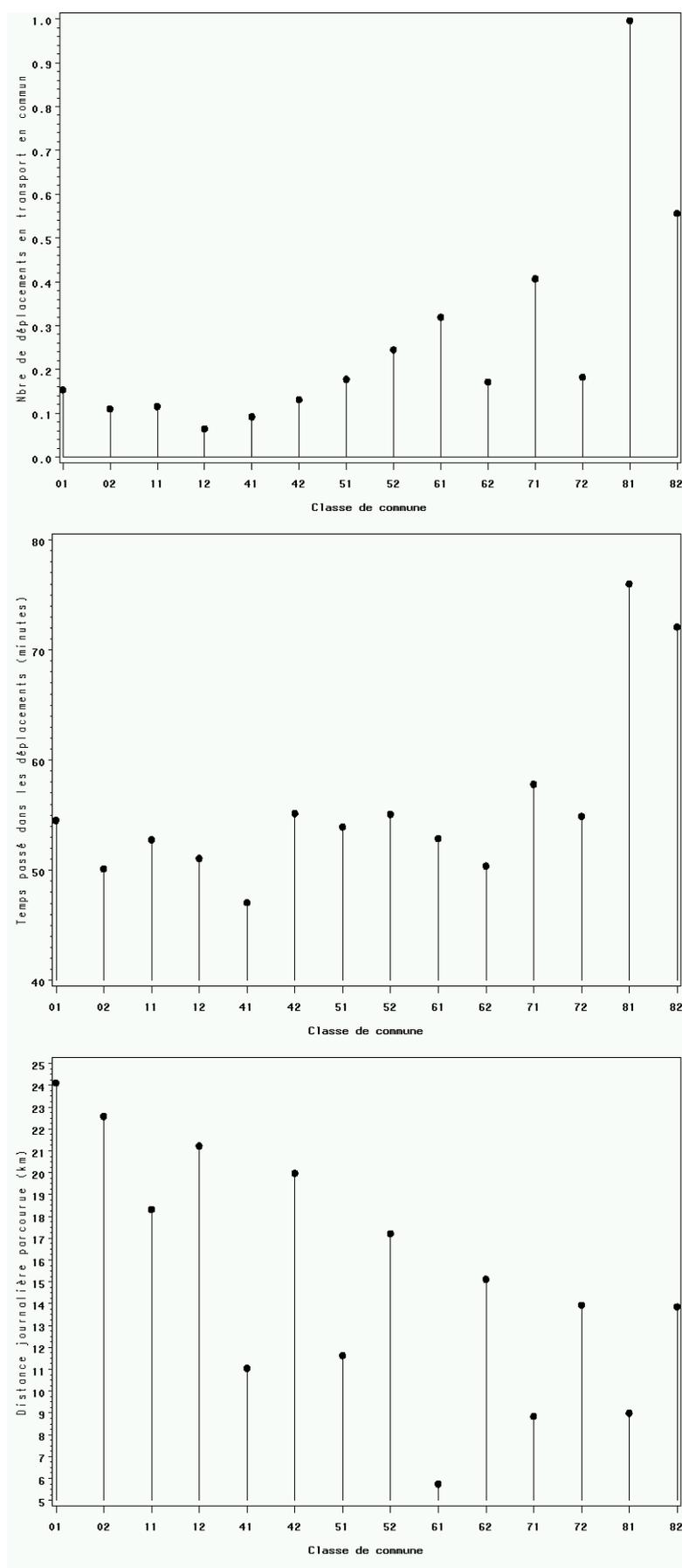
01	rurale dans un espace urbain
02	rurale dans un espace à dominante rurale
11	agglomération de moins de 20000 habitants - ville centre
12	agglomération de moins de 20000 habitants - banlieue
41	agglomération de 20000 à 50000 habitants - ville centre
42	agglomération de 20000 à 50000 habitants - banlieue
51	agglomération de 50000 à 100000 habitants - ville centre
52	agglomération de 50000 à 100000 habitants - banlieue
61	agglomération de 100000 à 200000 habitants - ville centre
62	agglomération de 100000 à 200000 habitants - banlieue
71	agglomération de 200000 à 2 M d'habitants - ville centre
72	agglomération de 200000 à 2 M d'habitants - banlieue
81	Paris
82	Banlieue parisienne

3.1.3 *Les déplacements selon la catégorie sociale*

Plus encore que par le milieu social, la mobilité est influencée par le type d'activité, distinguant les personnes ayant un emploi des inactifs, et parmi ces derniers, le type d'inactivité, retraités ou étudiants. Fréquence totale des déplacements et mode de transport utilisé apparaissent corrélés au type de profession exercée. Les retraités se déplacent moins que les actifs et les étudiants. Les professions dites « intermédiaires », dans lesquelles sont classées des activités très peu sédentaires (VRP, techniciens de maintenance), sont celles qui se déplacent le plus. Comme les professions libérales, ce sont aussi ces catégories qui font le plus de déplacements en voiture (figure 4).

L'effet est plus accusé en Ile de France qu'en province. Dans cette région, la plus grande propension des classes moyennes à utiliser les transports en commun peut s'expliquer par leur plus grande présence dans Paris intra-muros, où la densité des transports collectifs est maximum.

Figure 3 - Influence de la catégorie de commune de résidence sur les déplacements

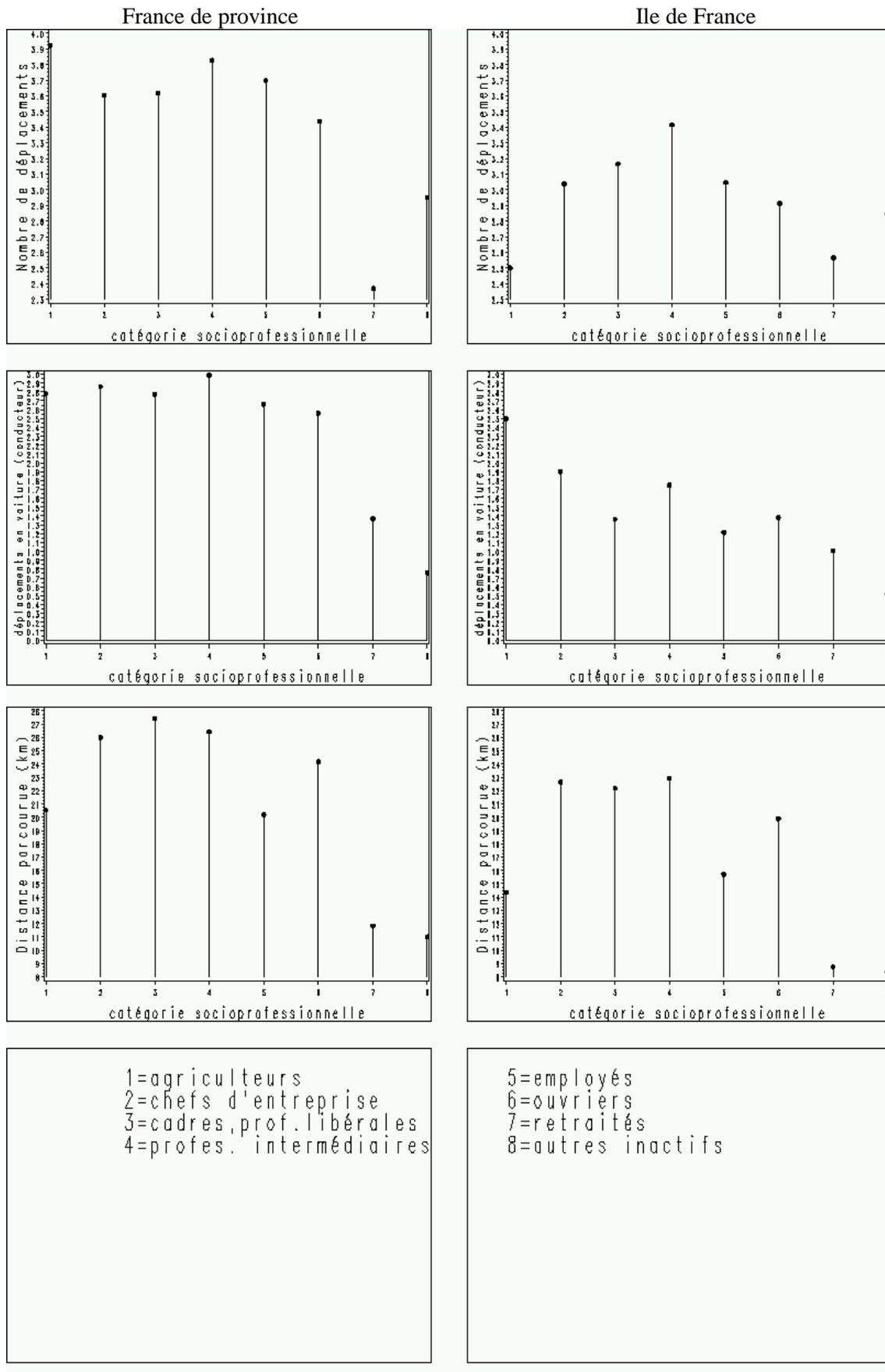


Nombre de déplacements par
personne et par jour

Durée totale passée dans les
déplacements, par personne et
par jour (minutes)

Distance totale à vol d'oiseau
parcourue par personne et par
jour (km)

Figure 4 - Les déplacements par catégorie socioprofessionnelle



3.2 Le modèle complet

Deux approches sont possibles pour modéliser la mobilité locale. La première consiste à rechercher un modèle de comportement individuel. Le nuage de points analysé est alors l'échantillon des unités ayant répondu à l'enquête, dont on cherche les caractéristiques les plus corrélées à la fréquence de leurs déplacements, aux distances parcourues et aux durées de transport. Dans cette enquête, on pouvait s'intéresser aux déplacements effectués par un individu, ou au total estimé des déplacements effectués par les membres d'un ménage échantillonné. C'est cette seconde méthode qui a été adoptée ici.

La seconde approche consiste à analyser les régions elles-mêmes, en recherchant les facteurs structurels : pyramide des âges, structure d'activité, de qualification..., les plus discriminants en matière de déplacements. Elle ne pouvait être retenue ici, en raison du faible nombre de points à analyser par régression (21 en excluant la Corse, dont l'échantillon est de taille trop faible pour participer à un modèle de ce type). C'est pourquoi on a procédé à une analyse de la variance du nuage des ménages.

Les régressions ont été pondérées par les poids de sondage après partage des poids entre échantillon national et extensions locales et après redressement de la non-réponse, mais avant calage.

3.2.1 *Un modèle de comportement individuel*

L'unité observée est l'individu Kish ayant répondu en deuxième visite. La population de référence est donc constituée des personnes de 6 ans ou plus au moment de l'enquête.

Au niveau individuel, la propension à la mobilité un jour de semaine est résumée par une variable dichotomique valant 1 si la personne a effectué au moins un déplacement le jour de référence, et 0 sinon. Les autres variables analysées sont des variables numériques : nombre total de déplacements effectués dans la journée, nombre de déplacements par mode de transport, nombre de déplacements par motif, distance totale parcourue, temps passé dans les déplacements pendant la journée de référence.

Les variables explicatives testées ont été :

- L'âge de l'individu, regroupé en 10 tranches : 6-13 ans, 14-18 ans, 19-25 ans, 26-29 ans, 30-39 ans, 40-49 ans, 50-60 ans, 61-69 ans, 70-79 ans, 80 ans et plus
- Le sexe
- La catégorie socioprofessionnelle en 7 postes, les retraités étant reclassés dans leur ancienne activité
- Le type de commune de résidence
- La structure familiale du ménage auquel appartient l'individu
- Le nombre de voitures dont dispose le ménage auquel appartient l'individu

Dans une phase exploratoire, la distance entre lieu de résidence et lieu de travail ou d'étude d'une part, le moyen de transport habituellement utilisé pour se rendre à son lieu de travail ou d'étude d'autre part, présents dans le questionnaire d'enquête, ont été introduits pour l'analyse des distances parcourues et de la durée des déplacements. Ce test a confirmé leur influence. Cependant, ces deux variables n'ont pu être retenues dans le modèle final d'estimation régionale car elles ne sont pas homogènes entre les deux sources : recensement et ENTND.

Dans l'enquête, un individu exerçant plusieurs emplois était interrogé sur chacun de ses lieux de travail, alors que le recensement retient seulement l'activité principale. En cas de pluri-activité, l'activité principale n'est pas non plus identifiable dans l'ENTND avec la même définition que dans le recensement.

Au recensement, les personnes ayant un emploi déclarent leur mode principal de transport pour se rendre au travail. Cette variable est apparue non homogène à la question correspondante dans l'enquête Transport. Elle n'a donc pu être utilisée pour un calage direct.

Toujours à titre exploratoire, l'influence de la densité des moyens de transport en commun autour du domicile, cernée par l'enquête, a également été testée. Le questionnaire relevait en effet le nombre de lignes de transport en commun par mode (bus, tramway, train) dans un rayon maximum autour du domicile de l'interviewé. Ni le recensement, ni les sources administratives ne permettant de construire des indicateurs de densité des transports collectifs par commune de résidence, cette information n'a pu être conservée dans la modélisation régionale définitive. On a cependant pu vérifier que ce facteur explicatif était assez correctement pris en compte par le type de zone de résidence, défini par son degré d'urbanisation et sa proximité par rapport à un centre-ville.

Dans le champ « France de province », la présence d'un effet régional résiduel a été recherché en ajoutant la région de résidence au modèle.

Les nomenclatures des variables explicatives sont indiquées en annexes 3 et 4.

Les facteurs explicatifs numériques ont été rendu catégoriels par regroupement de valeurs : c'est le cas de l'âge, exprimé en intervalles disjoints, du nombre de voitures, réduit à une modalité unique au-delà de 2. C'est donc par la technique de l'analyse de la variance, à l'aide de la procédure GLM du logiciel SAS, qu'on a recherché les facteurs les plus corrélés à la fréquence des déplacements individuels, tandis que la propension à la mobilité est modélisée par régression logistique.

Pour les raisons indiquées plus haut, toutes les régressions ont été réalisées sur deux champs géographiques distincts : Ile de France et France de province.

Propension à la mobilité le jour de référence

Programme :

```
%macro mobil(tab=base.indiv_modele,champ=prov,y=,x=,pond=);
  proc logistic data=&tab descending ;
    %if &champ=prov %then
      where rg ne '11' %str(;;);
    %if &champ=idf %then
      where rg='11' %str(;;);
    %if &pond=%str( ) %then
      weight pondki %str(;;); /* pondération nationale après redressement */
    %if &pond=sondage %then
      weight poidsind %str(;;); /* pondération initiale redressée pour non réponse */
    class rg(ref='52') sexe agmod(ref='40-49 ans') cs8(ref='6') cspr(ref='6') typmen5(ref='4')
      nbvoit (ref='1') clascom %if &champ=idf %then (ref='81'); %else (ref='61'); %str(;;);
    model &y = &x / selection=stepwise ;
    title3 %if &champ=prov %then "Champ : France de province";
      %else %if &champ=idf %then "Champ : Ile de France";
      %else "Champ : France entière"; %str(;;);
  run;
  quit;
%mend;

%mobil(champ=prov, x=agmod cs8 clascom typmen5 nbvoit sexe, pond=sondage);
%mobil(champ=idf, x=agmod cs8 clascom typmen5 nbvoit sexe, pond=sondage);
%mobil(champ=france,x=agmod cs8 clascom typmen5 nbvoit sexe, pond=sondage);

%mobil(champ=prov, x=agmod cs8 clascom typmen5 nbvoit sexe rg,pond=sondage);
%mobil(champ=france,x=agmod cs8 clascom typmen5 nbvoit sexe rg,pond=sondage);
```

Résultats

La part des personnes qui ne se déplacent pas du tout le jour de référence varie de 10 % à 25 % d'une région à l'autre. Les six variables testées paraissent significatives au seuil de 5 %. L'âge est le facteur le plus discriminant, suivi par la catégorie socioprofessionnelle de l'individu et la zone de résidence. La catégorie sociale du chef de ménage, significative au seuil de 5 % du test de Fischer, apporte peu au score d'ensemble dans un modèle contenant déjà la CS de l'individu interrogé. La population

« immobile » se recrute plutôt après 60 ans, parmi les inactifs, résidant hors des grands centres urbains et ne disposant pas de voiture.

Lorsqu'on ajoute à ces six facteurs, dans le champ de la France de province, la région de résidence en variable explicative, celle-ci paraît significative. Il reste donc un effet régional résiduel que les autres facteurs ne suffisent pas à éliminer. Son introduction modifie cependant peu l'efficacité globale du modèle, le nombre de paires concordantes passant de 73 % à 73,6 %.

Fréquence des déplacements effectués par individu le jour de référence, distance totale parcourue et durée totale consacrée aux déplacements

Programme :

```
%macro test(tab=base.indiv_modele,champ=,x=,y=,pond=);
proc GLM data=&tab ;
  %if &champ=prov %then
    where rg ne '11' %str();
  %if &champ=idf %then
    where rg='11' %str();
  %if &pond=%str( ) %then
    weight pondki %str(); /* pondération nationale après redressement */
  %if &pond=sondage %then
    weight poidsind %str(); /* pondération initiale redressée pour non réponse */
  class rg sexe agq ag6 agmod agmodb agmodc zoneresfin clascom zone cs8 cs8b cs_k cspr
    typmen5 nbvoit ;
  model &y = &x / solution ;
  title3 %if &champ=prov %then "Champ : France de province";
    %else %if &champ=idf %then "Champ : Ile de France";
    %else "Champ : France entière"; %str();
run; quit;
%mend;

title "Nombre de déplacements journaliers par individu";
%test(tab=base.indiv_modele,champ=prov,pond=sondage,y=ndep,
  x=sexe agmod*cs8*clascom typmen5 nbvoit cspr);
%test(tab=base.indiv_modele,champ=idf,pond=sondage,y=ndep,
  x=sexe*agmod*cs8 clascom typmen5 nbvoit cspr );

title "Nombre de déplacements professionnels par individu";
%test(tab=base.indiv_modele,champ=prov,pond=sondage,y=depprofes,
  x=sexe agmod*cs8 nbvoit clascom typmen5 cspr);
%test(tab=base.indiv_modele,champ=idf,pond=sondage,y=depprofes,
  x=sexe agmod cs8 clascom typmen5 nbvoit cspr);

title "Nombre de déplacements en voiture (conducteur) par individu";
%test(tab=base.indiv_modele,champ=prov,pond=sondage,y=ndepvoitc,
  x= agmod nbvoit cs8*clascom*typmen5 sexe cspr );
%test(tab=base.indiv_modele,champ=idf,pond=sondage,y=ndepvoitc,
  x=nbvoit agmod cs8*clascom*typmen5 cspr sexe);

title "Nombre de déplacements en transport en commun par individu";
%test(tab=base.indiv_modele,champ=prov,pond=sondage,y=ndeptc,
  x=sexe*agmod cs8*clascom*typmen5 nbvoit cspr mtc1000b );
%test(tab=base.indiv_modele,champ=idf,pond=sondage,y=ndeptc,
  x=sexe*agmod cs8*clascom typmen5*nbvoit*cspr );

title "Distance totale parcourue au cours d'une journée par individu";
%test(tab=base.indiv_modele,champ=prov,pond=sondage,y=disttot,
  x=sexe agmod*cs8*clascom cspr typmen5 nbvoit);
%test(tab=base.indiv_modele,champ=idf,pond=sondage,y=disttot,
  x=sexe agmod*nbvoit*clascom cs8*cspr typmen5 );

title "Durée passée dans les déplacements au cours d'une journée par individu";
%test(tab=base.indiv_modele,champ=prov,pond=sondage,y=tempstot,
```

```
x=agmod*clascom*cs8 sexe cspr typmen5 nbvoit );  
%test(tab=base.indiv_modele,champ=idf,pond=sondage,y=tempstot,  
x=agmod*cspr*nbvoit clascom*cs8 sexe*typmen5 );
```

Le tableau 1 résume les résultats obtenus, selon que l'on retient ou non l'Ile de France

Tableau 1 - Modélisation des individus

Variable expliquée	Régions de province		Ile de France	
	Effets significatifs	R2	Effets significatifs	R2
Nombre total de déplacements par personne le jour de référence	Zone de résidence *Age*CS de l'individu Équipement des ménages en voitures Sexe Type de ménage CS du chef de ménage	0,19	Sexe*Age*CS de l'individu Équipement des ménages en voitures Type de ménage CS du chef de ménage	0,10
Nombre de déplacements professionnels	CS de l'individu*Age Sexe Équipement des ménages en voitures Zone de résidence	0,35	CS de l'individu Équipement des ménages en voitures Sexe Type de ménage	0,39
Nombre de déplacements en transport en commun par personne	Sexe*Age Zone de résidence *CS de l'individu*type de ménage Équipement des ménages en voitures CS du chef de ménage	0,20	Sexe*Age Équipement des ménages en voitures *CS du chef de ménage*Type de ménage Zone de résidence*CS de l'individu	0,19
Nombre de déplacements en voiture par personne	Équipement des ménages en voitures Age Zone de résidence *CS de l'individu*Type de ménage Sexe CS du chef de ménage	0,29	Équipement des ménages en voitures Age Zone de résidence *CS de l'individu*Type de ménage CS du chef de ménage	0,28
Distance totale parcourue par personne le jour de référence	Zone de résidence *Age*CS de l'individu Équipement des ménages en voitures Sexe CS du chef de ménage	0,21	Sexe CS du chef de ménage*CS de l'individu Zone de résidence *Age*Nbre de voitures du ménage	0,25
Temps total passé dans les déplacements par personne	Zone de résidence *Age*CS de l'individu Sexe CS du chef de ménage	0,13	Age*CS du chef de ménage*Nbre de voitures du ménage Zone de résidence*CS de l'individu	0,15

Les migrations quotidiennes de travail, qu'elles soient motivées par l'emploi salarié ou les études, sont les plus prévisibles des déplacements. Leur modélisation par des facteurs démographiques et d'activité est donc plus efficace que pour l'ensemble de tous les déplacements, incluant ceux occasionnés par les loisirs ou les achats. Dans le premier cas, le modèle explique entre 35 % et 40 % de la variance entre individus selon le champ géographique, dans le second, 19 % de la variance totale hors Ile de France. Les interactions entre zone de résidence et âge ou catégorie sociale exercent, en province, une forte influence sur la fréquence des déplacements par individu. En Ile de France, le type d'activité et le milieu social, lui-même corrélé au nombre de voitures du ménage, sont déterminants.

La fréquence des déplacements en voiture et en transport en commun est fortement influencée par le niveau d'équipement des ménages en voitures. Quel que soit l'indicateur choisi : nombre moyen de déplacements en transport en commun dans la journée, ou propension générale à prendre les transports en commun (variable dichotomique valant 1 quand l'individu s'est déplacé au moins une fois en transport en commun le jour de référence), la présence d'équipements de transport collectif à moins d'un kilomètre du domicile favorise le recours à ces moyens. Lorsqu'on introduit la zone de résidence et les caractéristiques individuelles des personnes, l'influence des indicateurs de densité d'équipement est relayée par les autres facteurs.

La distance parcourue au cours d'un déplacement est mesurée par la distance à vol d'oiseau entre les centres des communes de départ et d'arrivée. La distance totale parcourue dans la journée est le cumul des distances de tous les déplacements effectués le jour de référence. Tous les déplacements effectués à l'intérieur d'une même commune sont affectés d'une distance nulle. Il en résulte une certaine imprécision du résultat.

Zone de résidence et nombre de voitures du ménage sont déterminantes sur la distance parcourue au cours des déplacements journaliers, surtout en province. A cela s'ajoute une corrélation avec la catégorie socioprofessionnelle, les cadres se déplaçant sur de plus longues distances. La distance domicile-travail, lorsqu'elle est déclarée par la personne interrogée, apparaît déterminante, mais elle ne pourra pas être reprise dans une procédure d'estimation régionale, les informations disponibles sur l'ensemble de la population n'étant pas homogènes à celles de l'enquête.

3.2.2 *Approche simultanée des niveaux ménages et individus*

Dans chaque ménage répondant, on calcule le nombre d'individus du ménage par sexe-âge et par catégorie socioprofessionnelle, que l'on introduit en variables explicatives, à côté des caractéristiques du ménage, communes à chacun de ses membres : zone de résidence, équipement en voitures, catégorie sociale du chef de ménage, structure familiale. On cherche alors à expliquer le cumul des déplacements des membres du ménage, estimé par les déplacements pondérés de l'individu interrogé.

Programme

```
options mprint;

%MACRO REGRESMEN(tab=synt.mendep,champ=,y=,x=);
  PROC GLM data=&tab ;
    %if &champ=idf %then
      where rg='11' %str(;) ;
    %else %if &champ=prov %then
      where rg ne '11' %str(;) ;
    class clascom cspr typmenb voit ;
    model &y = &x / solution;
    weight poidsmen;
  run; quit;
%MEND;

title "Modèle de niveau ménages";
title2 "Champ : France de province";
%REGRESMEN(champ=prov,
  y=ndep depprofes ndeptc ndepvoitc disttot tempstot,
  x=clascom cspr typmenb voit cs_k_1-cs_k_7 sexagereg_k_101-sexagereg_k_109
  sexagereg_k_201-sexagereg_k_209);
title2 "Champ : Ile de France";
```

```
%REGRESMEN(champ=idf,
            y=ndep depprofes ndeptc ndepvoitc disttot tempstot,
            x=clascom cspr typmenb voit cs_k_1-cs_k_7 sexagereg_k_101-sexagereg_k_109
            sexagereg_k_201-sexagereg_k_209);
```

Tableau 2 - Modélisation des ménages

Variable expliquée	R2	
	France de province	Ile de France
Nombre de déplacements par ménage le jour de référence	0,46	0,48
Nombre de déplacements professionnels	0,39	0,46
Nombre de déplacements en transport en commun	0,21	0,21
Nombre de déplacements en voiture	0,37	0,29
Distance totale parcourue le jour de référence	0,23	0,32
Temps total passé dans les déplacements	0,33	0,39

Le tableau ci-dessus résume les résultats obtenus.

3.3 La construction des marges régionales

Les enquêtes de recensement mettent désormais à disposition une information démographique et sociale riche mise à jour annuellement, et représentative au niveau régional. Au moment de la réalisation de cette étude, les estimations définitives au 1^{er} janvier 2006, basées sur un cycle quinquennal d'enquêtes de 2004 à 2008, ainsi que l'enquête « transversale » de 2009 étaient disponibles.

Le recensement nous donne la population des ménages et sa répartition selon les caractéristiques des individus (sexe, âge, activité, CSP, diplôme, mode de cohabitation). Il relève également le nombre de voitures à disposition du ménage, la commune d'étude ou de travail en plus de celle de résidence. Le fichier du RP est enrichi de codes géographiques, dont celui du zonage en aires urbaines et espace rural (ZAUER), qui permet de caractériser le lieu de résidence en terme de proximité à un centre urbain selon la taille de celui-ci.

Pour être utilisée dans un estimateur modélisé, l'information auxiliaire doit être mesurée à la date de l'enquête. Elle est constituée des valeurs individuelles $x_{j,k}$ des variables X_j pour chaque individu k de l'échantillon, et des totaux $X_{j,R}$ de ces variables dans chaque région R . En raison du décalage temporel entre le dernier recensement disponible (2006) et la date de référence de l'ENTD (janvier 2008), les valeurs $X_{j,R}$ mobilisées comprennent une part d'estimation.

Le recensement de 2006, résultat de la concaténation des 5 enquêtes annuelles de recensement de 2004 à 2008, nous donne la valeur des variables auxiliaires par région au 1^{er} janvier 2006. Elles ont été actualisées à 2008 en faisant l'hypothèse d'une constance de la structure de la population et des

ménages entre 2006 et 2008. Soit $P_{R,08}$ la population des ménages dans la région R au 1^{er} janvier 2008 :

$$\hat{X}_{j,R,08} = \frac{X_{j,R,06}}{P_{R,06}} \times P_{R,08}$$

On a évalué de même la répartition des ménages par critère.

La population totale par sexe et âge estimée par région était disponible au 1^{er} janvier 2008. On en a déduit une estimation de la population des ménages sous l'hypothèse d'une constance de leur part dans la population totale par âge détaillé, le ratio utilisé étant celui fourni par le recensement de 2006 :

$$\hat{P}_{ménages,08} = \sum_{\text{âge}} \left(\frac{P_{ménages,06}}{P_{tot,06}} \times P_{tot,08} \right)_{\text{âge}}$$

Le nombre des ménages au 1^{er} janvier 2008 a été estimé par le rapport entre population estimée des ménages et taille moyenne des ménages en 2008. Celle-ci a été évaluée, selon les régions, en prenant le résultat de l'enquête annuelle de recensement de 2008, ou la moyenne des résultats des trois EAR de 2006 à 2008 :

$$TAILMEN_{08} = \begin{cases} (EAR_{06} + EAR_{07} + EAR_{08}) / 3 & \text{(régions Picardie, Alsace, Lorraine, Nord - Pas - de - Calais, Limousin)} \\ EAR_{08} & \text{(autres régions)} \end{cases}$$

$$NB\hat{M}EN_{R,08} = \frac{\hat{P}_{ménages,08}}{TAILMEN_{08}}$$

4 Les méthodes d'estimation sur « petits domaines »

En théorie de l'estimation, un « domaine » est une sous-population dont les contours ne sont pas définis par les paramètres du plan de sondage, au contraire d'une strate quand l'échantillon est stratifié. Il en résulte que l'intersection de l'échantillon, tiré dans la population cible complète, avec ce domaine, a une taille n_d aléatoire. Le total (respectivement la moyenne) d'une variable d'intérêt dans le domaine est estimé sans biais par l'estimateur Horvitz-Thomson classique restreint à cette sous-population. Sa variance dépend en revanche de la taille N_d de cette sous-population et du nombre n_d d'unités échantillonnées dans le domaine, et non exclusivement de la taille totale n de l'échantillon. On obtient donc des estimateurs d'autant moins précis que le domaine est une sous-population plus petite, et que son intersection avec l'échantillon est également de faible taille.

Notations :

U est la population de référence, de taille N

$U_d \subset U$ est une sous population de U

s est l'échantillon sélectionné dans U, de taille n

$s_d = s \cap U_d$ est l'intersection entre le domaine U_d et l'échantillon s

\mathcal{Y} est une variable d'intérêt de l'enquête

X_j est une variable auxiliaire, connue par ailleurs

$Y_U = \sum_{i \in U} Y_i$ est le total de \mathcal{Y} dans la population, et $\bar{Y}_U = \frac{1}{N} \sum_{i \in U} Y_i$ sa moyenne

$Y_d = \sum_{i \in U_d} Y_i$ est le total de \mathcal{Y} dans le domaine U_d , et $\bar{Y}_d = \frac{1}{N_d} \sum_{i \in U_d} Y_i$ sa moyenne dans U_d

$X_j = \sum_{i \in U} X_{j,i}$ est le total de \mathcal{X}_j dans la population, $\bar{X}_j = \frac{1}{N} \sum_{i \in U} X_{j,i}$ sa moyenne

$X_{jd} = \sum_{i \in U_d} X_{j,i}$ est le total de \mathcal{X}_j dans le domaine U_d , $\bar{X}_{j,d} = \frac{1}{N_d} \sum_{i \in U_d} X_{j,i}$ sa moyenne

$\mathbf{X} = \begin{bmatrix} X_1 \\ \dots \\ X_j \\ \dots \\ X_J \end{bmatrix}$ est le vecteur des totaux dans U , de J variables auxiliaires \mathcal{X}_j

\mathbf{X}_d est le vecteur des totaux dans U_d des J variables auxiliaires \mathcal{X}_j

$\pi_i = \text{Prob}\{i \in s\}$ est la probabilité d'inclusion de l'unité i dans l'échantillon s

4.1 Précision d'un estimateur de domaine avec le sondage aléatoire simple

Dans le cas du sondage aléatoire simple, le problème peut se formaliser de la façon suivante. On a sélectionné, par sondage aléatoire simple, un échantillon s de taille n dans la population U , et l'on estime la moyenne \bar{Y}_U d'une variable d'intérêt \mathcal{Y} . Une unité i de U a la probabilité $\pi_i = \frac{n}{N}$ d'être incluse dans s . Avec un sondage aléatoire simple, \bar{Y}_U est estimé sans biais par la moyenne empirique dans l'échantillon :

$$\hat{\bar{Y}}_U = \frac{\hat{Y}}{\hat{N}} = \frac{\sum_{i \in s} \frac{y_i}{\pi_i}}{\sum_{i \in s} \frac{1}{\pi_i}} = \frac{1}{n} \sum_{i \in s} y_i = \bar{y}$$

$$V(\hat{\bar{Y}}_U) = (1-f) \frac{S^2}{n} \quad \text{où : } S^2 = \frac{1}{N-1} \sum_{i \in U} (Y_i - \bar{Y}_U)^2$$

L'intersection s_d de s avec la sous population U_d contient n_d unités.

Le total Y_d est estimé sans biais par :

$$\hat{Y}_d = \sum_{i \in s_d} \frac{N}{n} y_i = N \frac{n_d}{n} \bar{y}_d$$

qui a pour variance :

$$V(\hat{Y}_d) = N^2 \frac{(1-f)}{n} \frac{1}{N-1} [(N_d - 1)S_d^2 + N_d(1-P_d)\bar{Y}_d^2]$$

$$V(\hat{Y}_d) \approx N^2 \frac{(1-f)}{n} P_d [S_d^2 + (1-P_d)\bar{Y}_d^2] = N_d^2 \left(\frac{1-E(f_d)}{E(n_d)} \right) [S_d^2 + (1-P_d)\bar{Y}_d^2] \quad (a)$$

$$\text{où : } \begin{cases} P_d = \frac{N_d}{N} = E\left(\frac{n_d}{n}\right) \\ f_d = \frac{n_d}{N_d} \\ \bar{Y}_d = \frac{1}{N_d} \sum_{i \in U_d} Y_i \\ S_d^2 = \frac{1}{N_d - 1} \sum_{i \in U_d} (Y_i - \bar{Y}_d)^2 \end{cases}$$

Lorsque la taille N_d du domaine est connue, la moyenne peut être estimée par : $\hat{Y}_d = \frac{\hat{Y}_d}{N_d}$, qui a alors pour variance :

$$V(\hat{Y}_d) = \frac{1}{N_d^2} V(\hat{Y}_d) \approx \left(\frac{1 - E(f_d)}{E(n_d)} \right) [S_d^2 + (1 - P_d) \bar{Y}_d^2] \quad (\text{b})$$

Lorsque la taille N_d du domaine est inconnue, la moyenne de \mathcal{Y} dans U_d est estimée approximativement sans biais par la moyenne empirique du domaine dans l'échantillon :

$$\hat{Y}_d = \frac{\hat{Y}_d}{\hat{N}_d} = \frac{1}{N \frac{n_d}{n}} \sum_{i \in s_d} \left(\frac{N}{n} y_i \right) = \bar{y}_d$$

La linéarisation de Taylor nous donne l'erreur quadratique moyenne approchée de la moyenne estimée :

$$EQM(\bar{y}_d) \approx (1 - f) \frac{S_d^2}{n \times P_d} = (1 - f) \frac{S_d^2}{E(n_d)} \quad (\text{c})$$

On voit d'après (b) et (c) que la variance de la moyenne est en $1/E(n_d)$. Dans le cas (b), elle comprend un terme supplémentaire en $(1 - P_d)$. Elle augmente par conséquent d'autant plus que l'effectif échantillonné n_d , lui-même dépendant de la proportion P_d du domaine dans la population, est plus faible. Pour une valeur donnée du paramètre θ à estimer, qu'il s'agisse du total dans la population ou de sa moyenne, l'erreur relative de l'estimation, exprimée par son coefficient de variation $CV(\hat{\theta}_d) = \hat{\theta}_d / \sqrt{V(\hat{\theta}_d)}$, s'accroît en raison inverse de la taille du domaine dans l'échantillon.

4.2 Les grands types de modèles

Dans son « panorama des méthodes d'estimation sur petits domaines », P. Ardilly distingue trois grandes familles d'estimateurs.

4.2.1 Les estimateurs directs

Ils ne font appel qu'aux observations réalisées dans le domaine. L'estimateur Horvitz-Thomson « spontané », obtenu avec les pondérations du sondage, se range dans cette catégorie. Il peut être amélioré par régression ou par calage sur des informations auxiliaires propres au domaine. Le gain de précision sera d'autant plus important que les variables de calage seront mieux corrélées aux variables d'intérêt. Dans la pratique, le calcul d'un estimateur régional redressé par calage repose sur la disponibilité d'une information auxiliaire pertinente et précise à ce niveau. Néanmoins, la variance d'un tel estimateur est une expression en $1/n_d$: sa précision reste tributaire de la taille de l'échantillon dans le domaine.

L'estimateur calé sur des données régionales a pour expression :

$$\boxed{\hat{Y}_{d,cal} \approx \hat{Y}_{d,reg} = \hat{Y}_{d,HT} + \hat{\beta}_d' (\mathbf{X}_d - \hat{\mathbf{X}}_{d,HT})} \quad (1)$$

où \mathbf{X}_d est le vecteur des totaux de J variables auxiliaires, dans U_d , $\hat{\mathbf{X}}_{d,HT}$ son estimateur Horvitz-Thomson, et $\hat{\beta}_d$ le vecteur des coefficients de la régression linéaire de \mathcal{Y} sur \mathcal{X} à l'intérieur du domaine, estimé dans le sous-échantillon s_d .

$$\hat{\beta}_d = (\mathbf{X}_d' \mathbf{D} \mathbf{X}_d)^{-1} \mathbf{X}_d' \mathbf{D} \mathbf{Y}_d \quad \text{avec : } \mathbf{D} = \text{diag} \left(\frac{1}{\pi_i} \right)$$

\mathbf{X}_d est la matrice (n_d,J) des observations des variables X_j sur les unités de l'échantillon appartenant au domaine U_d .

Une variante consiste à estimer le coefficient $\hat{\mathbf{b}}$ dans l'échantillon complet, et non plus dans sa restriction s_d au domaine, afin d'en améliorer la robustesse :

$$\boxed{\tilde{Y}_{d,regsyn} = \hat{Y}_{d,HT} + \hat{\mathbf{b}}' (\mathbf{X}_d - \hat{\mathbf{X}}_d)} \quad (2)$$

où :

$$\hat{\mathbf{b}} = (\mathbf{X}' \mathbf{D} \mathbf{X})^{-1} \mathbf{X}' \mathbf{D} \mathbf{Y} \quad \text{avec : } \mathbf{D} = \text{diag} \left(\frac{1}{\pi_i} \right)$$

\mathbf{X} étant la matrice (n,J) des observations des variables X_j sur toutes les unités de l'échantillon s.

4.2.2 Les estimateurs indirects avec modélisation implicite

Ces estimateurs reposent sur une hypothèse de similitude de comportement entre les unités appartenant au domaine et l'ensemble de la population.

L'estimateur synthétique sans autre information auxiliaire que la taille du domaine N_d est le plus simple de ces estimateurs : le total Y dans la région est estimé sous l'hypothèse d'égalité de la moyenne entre la région et la population totale.

$$\hat{Y}_{d,syn} = \hat{Y}_U = \frac{\hat{Y}_U}{\hat{N}} \Rightarrow \boxed{\hat{Y}_{d,syn} = \hat{Y}_U \times N_d} \quad (3)$$

L'estimateur synthétique de type régression est fondé sur un modèle de super-population. A condition de disposer d'information auxiliaire sur la population complète, on postule l'existence d'une relation linéaire entre la variable d'intérêt \mathcal{Y} à estimer et un vecteur de J variables X_j , observées dans U. On fait l'hypothèse que cette relation est également vérifiée à l'intérieur du domaine, ce qui permet de calculer un nouvel estimateur par régression :

$$\boxed{\hat{Y}_{d,regsyn} = \hat{\mathbf{b}}' \mathbf{X}_d} \quad (4)$$

où :

$$\hat{\mathbf{b}} = (\mathbf{X}' \mathbf{D} \mathbf{X})^{-1} \mathbf{X}' \mathbf{D} \mathbf{Y} \quad \text{avec : } \mathbf{D} = \text{diag} \left(\frac{1}{\pi_i} \right)$$

\mathbf{X} est ici la matrice (n,J) des observations des variables X_j sur toutes les unités de l'échantillon s.

Contrairement à l'estimateur direct (1), le coefficient de régression $\hat{\mathbf{b}}$ est estimé dans l'échantillon complet s, ce qui le rend plus robuste. En contrepartie, on crée un biais si l'hypothèse sous-jacente d'identité de comportement entre le domaine et l'ensemble de la population n'est pas vérifiée.

L'estimateur post-stratifié est une déclinaison de ce modèle, lorsque les variables auxiliaires explicatives X_j sont de type catégoriel. Supposons que le croisement des modalités des variables X_j conduise à partitionner la population en H classes. Dans le domaine, N_{dh} est la taille de la classe h. L'estimateur synthétique par régression devient alors :

$$\hat{Y}_{d,regsyn} = \sum_{h=1}^H N_{dh} \frac{\hat{Y}_h}{\hat{N}_h} \quad (5)$$

L'hypothèse sous-jacente est une égalité des moyennes de classe \bar{Y}_h dans le domaine et la population entière. L'estimateur est redressé par les structures réelles connues du domaine, relativement aux critères de définition de la classification retenue.

Les estimateurs composites consistent en une combinaison linéaire d'un estimateur direct dans le domaine et d'un estimateur synthétique tel que ceux décrits précédemment. Ils sont donc de la forme :

$$\hat{Y}_{d,comp} = \alpha_d \hat{Y}_{d,Direct} + (1 - \alpha_d) \hat{Y}_{d,syn} \quad (6) \quad \text{où : } 0 < \alpha_d < 1$$

Ces estimateurs se distinguent par la façon dont le coefficient α est défini. En particulier, l'estimateur composite optimum utilise un coefficient qui minimise l'erreur quadratique moyenne, sous une hypothèse simplificatrice :

$$\hat{Y}_{d,compopti} = \alpha_{d,opti} \hat{Y}_{d,Direct} + (1 - \alpha_{d,opti}) \hat{Y}_{d,syn}$$

$$\alpha_{d,opti} = \frac{1}{1 + F_d} \quad F_d = \frac{EQM(\hat{Y}_{d,Direct})}{EQM(\hat{Y}_{d,syn})}$$

4.2.3 Les estimateurs indirects avec modélisation explicite

On distingue les modèles qui se situent d'emblée au niveau du domaine (les observations modélisées sont les agrégats mesurés dans chacun des domaines) et les modèles qui s'appuient sur des comportements individuels.

Le modèle de Fay-Herriot est du premier type. L'estimation porte d'emblée sur l'ensemble de tous les domaines (par exemple, les 90 départements ou les 22 régions françaises). La méthode prend en compte un aléa dû au modèle et un aléa dû au sondage. Elle s'appuie sur les techniques de régression avec effet aléatoire, dont les paramètres sont estimés par la méthode du maximum de vraisemblance. Le modèle a pour formulation générale :

$$\vartheta_d = \beta' z_d + b_d v_d \quad (A)$$

$$\text{où : } \vartheta_d = g(\bar{Y}_d)$$

ϑ_d est soit un transformé de la moyenne dans le domaine de la variable d'intérêt, soit la moyenne elle-même (g peut être la fonction identité). C'est la vraie valeur, inobservée, du paramètre ϑ dans le domaine U_d , tandis que $\hat{\vartheta}_d$ en est l'estimateur direct. Cette équation postule une relation linéaire entre le paramètre et un vecteur de variables auxiliaires connues z_d . b_d est un réel connu qui dépend du domaine. v_d est l'aléa du modèle, qui suit une loi d'espérance nulle et de variance σ_v^2 inconnue. Il prend en compte la variabilité inter-domaines non décrite dans z_d .

Par ailleurs, l'estimateur direct dans le domaine s'écrit :

$$\hat{\vartheta}_d = \vartheta_d + e_d \quad (B)$$

où e_d représente l'aléa de sondage.

En combinant (A) et (B), on obtient la forme générale complète de l'estimateur :

$$\hat{\vartheta}_d = \boldsymbol{\beta}' \mathbf{z}_d + \mathbf{b}_d \mathbf{v}_d + e_d \quad (8)$$

Les modèles de niveau individu ont pour forme générale :

$$Y_{d,i} = \boldsymbol{\beta}' \mathbf{X}_{d,i} + \nu_d + e_{d,i} \quad (9)$$

De façon classique, $\mathbf{X}_{d,i}$ est un vecteur de variables auxiliaires observées sur l'individu i appartenant au domaine d , corrélées à la variable numérique continue à estimer \mathcal{Y} , dont elles « expliquent » les variations. Le paramètre ν_d est un scalaire exprimant l'effet spécifique, aléatoire, du domaine d sur la variable \mathcal{Y} , et $e_{d,i}$ est le résidu du modèle, d'espérance nulle et de variance inconnue $\sigma_{d,e}^2$. Il s'agit d'un modèle mixte, combinant un effet aléatoire, l'appartenance au domaine, et des effets déterministes, les variables explicatives X_j .

5 Les méthodes d'estimation testées

On a testé quatre méthodes d'estimation : un estimateur traditionnel par calage sur des informations régionales, un estimateur direct modifié par régression dans l'échantillon national complet, un estimateur synthétique de type régression, et un estimateur par modélisation explicite. Ce sont les estimateurs (1), (2), (4), et (9) décrits plus haut.

Les estimateurs (2) et (4) ont chacun été calculés selon deux hypothèses :

- le modèle de comportement est unique sur toute la France, le coefficient de régression \mathbf{b} est donc estimé dans l'échantillon complet France entière.
- il existe un comportement spécifique francilien distinct d'un modèle provincial ; on estime alors un coefficient de régression $\boldsymbol{\beta}$ dans l'échantillon national privé des ménages de l'Ile de France.

Le modèle (9) a été estimé par la méthode du maximum de vraisemblance, avec la procédure MIXED du logiciel SAS, dans laquelle le code région est spécifié en effet aléatoire, et les variables de calage en effets déterministes. Afin de rester le plus proche possible du modèle sous-jacent au calage régional (1), l'unité k retenue est le ménage, auquel sont associées ses caractéristiques propres (variables X « de niveau » ménage) et le total estimé dans le ménage des variables de niveau individu de

l'échantillon Kish : $\hat{X}_m = \sum_{i \in s_m} \frac{x_i}{\pi_{i/m}} = \tilde{x}_{i,m}$, i étant l'indice de l'individu Kish sélectionné dans le

ménage m , et $\pi_{i/m}$ sa probabilité conditionnelle de sélection à l'intérieur du ménage.

5.1 Les variables estimées

Les indicateurs calculés sont des effectifs, des proportions, des moyennes par personne basées sur la population totale, des ratios par déplacement. Tous sont calculés pour une journée du lundi au vendredi, et ont pour cible les déplacements locaux effectués par un individu âgé de 6 ans ou plus, dans un rayon de moins de 80 km autour de son domicile. Le champ de la population mobile est plus large, car elle inclut tous les individus de 6 ans ou plus ayant effectué au moins un déplacement dans la journée de référence, quelle que soit la distance parcourue.

Le moyen de transport utilisé est celui déclaré par l'individu comme le moyen principal pour ce déplacement. Les déplacements en transport en commun comprennent le ramassage scolaire et le transport collectif organisé par l'employeur.

Un déplacement est considéré comme une navette domicile-travail si le motif d'origine ou de destination est le travail habituel ou les études. La ventilation des personnes selon le mode de transport utilisé (voiture ou transport en commun) fait donc référence aux déplacements de travail effectivement observés le jour de référence, et non au mode de transport déclaré comme habituel par la personne interrogée en première visite.

La distance parcourue est la distance à vol d'oiseau entre les centres des communes (ou des arrondissements à Paris, Lyon, Marseille) d'origine et de destination. Par convention, un déplacement interne à une commune (à un arrondissement) a donc une distance nulle.

5.2 Traitement de la non-réponse totale

L'échantillon a été préalablement repondéré pour non-réponse totale avant toute estimation régionale.

Chacune des deux visites de l'enquête a occasionné de la non-réponse. Les deux phases de non-réponse ont été analysées par modèle logistique et redressées séparément. Dans chaque cas, le coefficient de repondération était égal à l'inverse de la probabilité de réponse estimée de l'individu, issue des régressions. Les variables corrélées au comportement de réponse et disponibles dans la base de sondage ont été, en première visite : le type de logement (individuel/collectif), la taille du logement (moins de 3 pièces, 3 pièces, 4 pièces, 5 pièces ou plus), le nombre de personnes du logement (une ou plusieurs), la région (Ile de France, autre région), la vague de collecte³. En deuxième visite, les probabilités de réponse sont une fonction de la catégorie socioprofessionnelle (CS) de l'individu, du croisement âge-sexe et du nombre de voitures du ménage.

5.3 Les modalités de calage

Afin d'apprécier l'effet de la méthode sur le résultat, tous les estimateurs ont d'abord été construits avec les mêmes variables auxiliaires. Dans un second temps, le modèle a été amélioré par la prise en compte de modalités supplémentaires et les estimateurs indirects de type régression ont été calculés sous ce modèle détaillé.

Le calage sur des totaux régionaux est la méthode qui contraignait le plus à limiter le nombre de modalités dans le modèle, en raison des faibles tailles d'échantillon par région. C'est ce qui a déterminé le choix des nomenclatures des variables auxiliaires retenues pour cette comparaison.

L'estimateur direct par calage régional est basé sur un calage simultané des échantillons de ménages et d'individus Kish ayant répondu en visite 2. Celle-ci ayant pour cible les déplacements des seuls individus de 6 ans ou plus, l'échantillon régional a été calé sur le nombre total de ménages et sur la population âgée de six ans ou plus de la région. Les tailles des échantillons régionaux n'ont pas permis d'ajouter un troisième niveau de calage de l'échantillon complet des membres des ménages, y compris en-dessous de 6 ans, sur la population totale des ménages.

Les variables de calage ont été les suivantes.

Niveau ménages :

- zone de résidence en 6 classes (1=communes rurales en espace rural, 2=communes rurales en espace urbain, 3=communes péri-urbaines d'unités urbaines de moins de 100000 habitants, 4=ville-centres d'agglomérations de moins de 100000 habitants, 5=communes péri-urbaines d'agglomérations de 100000 habitants ou plus, 6=ville-centres d'unités urbaines de 100000 habitants ou plus)

³ La collecte s'est déroulée sur une année complète en six vagues successives à intervalles de deux mois, afin de tenir compte des effets saisonniers. Chaque vague a été réalisée auprès d'un sixième de l'échantillon, sélectionné aléatoirement.

- type de ménage en 4 modalités (personne seule ou ménage sans famille, famille monoparentale, couple sans enfants, couple avec enfants)
- nombre de voitures du ménage (ménages sans voiture, ménages ayant une voiture, ménages ayant 2 voitures ou plus)
- catégorie socioprofessionnelle de la personne de référence en 4 modalités (professions indépendantes incluant les agriculteurs, cadres et professions intermédiaires, ouvriers et employés, inactifs)

Niveau individus :

- Sexe
- Âge en 5 groupes (6-24 ans, 25-34 ans, 35-49 ans, 50-64 ans, 65 ans et plus)
- catégorie socioprofessionnelle de l'individu en six postes, les retraités étant reclassés dans leur profession d'origine (agriculteurs et professions indépendantes en activité ou retraités, cadres et professions intermédiaires actifs, ouvriers et employés actifs, anciens cadres et professions intermédiaires, anciens ouvriers et employés, autres inactifs)
- Nombre de voitures du ménage auquel appartient l'individu (0,1, 2 ou plus)

Dans le calage des ménages, les agriculteurs ont été regroupés avec les autres professions indépendantes (artisans, commerçants, chefs d'entreprise) en raison de leur faible effectif.

Avec la fonction de calage logit, les rapports entre poids de calage obtenus et poids initiaux se situent le plus souvent dans une échelle 0,25-3.

Les modalités de calage retenues sont regroupées par rapport à celles utilisées dans le redressement national de l'enquête, y compris dans les régions ayant fait des extensions d'échantillon. Néanmoins, dans ces régions, les variables d'intérêt calculées avec ces nouvelles pondérations sont très proches de celles obtenues avec les poids nationaux, particulièrement en Ile de France et en Pays de la Loire.

6 Les résultats comparés

6.1 Comparaison des méthodes, à modèle identique

En Ile de France et en Pays de la Loire, où les échantillons régionaux comprenaient plus de 2500 ménages répondants, on peut considérer que l'estimateur direct obtenu avec les pondérations nationales fait référence. Les résultats de l'enquête ont été calés sur les populations de ces régions lors du redressement national, et la taille de ces échantillons régionaux leur confère une certaine robustesse.

Dans les trois autres régions ayant fait une extension d'échantillon, les pondérations nationales intègrent également un calage sur des totaux régionaux. Les tailles d'échantillon, proches du millier de répondants, confèrent aux estimations directes une précision intermédiaire entre celle obtenue dans les deux régions précédentes et chacune des 17 autres régions sans extension.

Dans ces cinq régions, et principalement en Ile de France et en Pays de la Loire, l'écart entre l'estimateur direct, et un estimateur modélisé utilisant les mêmes modalités des variables auxiliaires que le calage régional, donne par conséquent un aperçu du biais éventuellement introduit par chacune des méthodes mises en œuvre.

6.1.1 L'estimateur synthétique de type régression

Il s'agit du modèle (4) décrit en section 4.2.2. Les coefficients de régression ont été estimés successivement sur l'ensemble de l'échantillon national, et sur l'échantillon hors Ile de France. Ces deux versions donnent des résultats différents principalement pour les variables traduisant l'usage des transports en commun ou de la voiture, et le temps consacré aux déplacements un jour de semaine, points sur lesquels l'Ile de France se singularise le plus. L'estimateur calculé à partir d'un « modèle de province » crée pour ces variables un biais moins important que celui issu d'un « modèle France

entière ». Ils se confondent pour les autres indicateurs : nombre de déplacements par motif, distance parcourue, déplacements à pied ou en vélo.

L'estimateur synthétique de type régression apparaît biaisé dès que le comportement régional diffère sensiblement du comportement moyen national.

En Ile de France, où une personne mobile effectue moins de déplacements par jour qu'un habitant de province, le nombre de déplacements par personne est surestimé. Dans les Pays de la Loire et en Bretagne, où la situation est inverse, la méthode sous-estime la fréquence des déplacements par habitant. En Midi-Pyrénées et en Languedoc-Roussillon, où un habitant se déplace avec la même fréquence qu'en moyenne nationale, l'estimateur par régression donne un résultat proche de l'estimateur direct.

Pour les mêmes raisons, la mobilité en transport en commun est correctement estimée en Bretagne et en Pays de la Loire, très sous-estimée en Ile de France, surestimée dans les deux régions méridionales concernées par une extension (figure 5).

Les écarts sont de même nature dans les autres régions (figure 6). L'estimateur par régression se trouve en-dehors de l'intervalle de confiance de l'estimateur initial direct dans près d'un quart des cas dans les régions de province.

Selon que le modèle est estimé dans l'échantillon national complet, ou dans l'échantillon hors Ile de France, la méthode tend à recentrer les estimations régionales vers la moyenne nationale ou vers la moyenne de province. En conséquence, elle estime correctement un indicateur régional dans les régions situées dans la moyenne nationale ou dans la moyenne de province, elle donne des résultats biaisés ailleurs.

6.1.2 L'estimateur par modèle mixte avec effet régional aléatoire

Il s'agit du modèle (9) dans la section 4.2.3. Comme le précédent, ce modèle a été estimé dans le champ national complet et dans celui de la France de province. Dans les régions avec extension, les estimateurs obtenus selon ces deux variantes sont quasiment confondus (figures 7 et 8). Dans les autres régions, c'est aussi très généralement le cas. La présence d'un coefficient régional aléatoire permet de prendre en charge l'écart de comportement entre l'Ile de France et le reste du pays.

Cette méthode donne des résultats beaucoup plus proches de l'estimateur direct que l'estimateur synthétique de type régression. Dans les régions avec extension, l'estimateur modélisé reste très généralement à l'intérieur des bornes de l'intervalle de confiance à 95 % de l'estimateur direct. La fréquence des déplacements en transport en commun est la variable qui résiste le plus à la modélisation testée. En Midi-Pyrénées et en Languedoc-Roussillon, l'estimateur par modèle mixte reste en-dehors de l'intervalle de confiance de l'estimateur direct. Il en est de même dans 4 autres régions sans extension.

On a représenté dans la figure 9, à côté de l'estimateur direct initial, valeur de référence, les quatre estimateurs « petits domaines » calculés dans les régions avec extension d'échantillon : estimateur synthétique de type régression sur le champ hors Ile de France, estimateur par modèle mixte, champ national et champ hors Ile de France, estimateur direct modifié (modèle (2) de la section 4.2.1). L'estimateur synthétique de type régression apparaît le plus biaisé. Il sort très souvent de l'intervalle de confiance de l'estimateur direct. L'estimateur direct modifié (pastilles bleu clair dans le graphique), est le plus proche de l'estimateur direct : c'est celui qui minimise le risque de biais, mais on sait par ailleurs qu'il réduit peu la variance, par rapport à un estimateur direct par calage sur des totaux régionaux ([7]). L'estimateur par modèle mixte se situe le plus souvent entre les deux. Il pourrait donc présenter une solution de compromis.

6.2 Résultats avec un modèle détaillé

L'estimateur synthétique de type régression et l'estimateur par régression à effets mixtes ont été estimés avec un modèle utilisant un plus grand nombre de variables auxiliaires. Ce deuxième modèle repose sur les mêmes effets explicatifs principaux que le modèle du calage régional, mais certains facteurs sont détaillés en un plus grand nombre de modalités.

Les variables explicatives sont les suivantes.

Niveau ménages :

- catégorie de commune en 14 modalités (voir annexe)
- type de ménage en 5 modalités (personne seule, famille monoparentale, couple sans enfants, couple avec enfants, ménage sans famille)
- nombre de voitures du ménage (aucune, 1 voiture, 2 voitures ou plus)
- catégorie socioprofessionnelle de la personne de référence en 7 modalités (professions indépendantes incluant les agriculteurs, cadres, professions intermédiaires, ouvriers, employés, retraités, autres inactifs)
- type d'habitat (appartement dans un immeuble collectif, autre type de logement)

Niveau individus :

- population masculine par tranche d'âge (6-13 ans, 14-18 ans, 19-25 ans, 26-29 ans, 30-39 ans, 40-49 ans, 50-60 ans, 61-69 ans, 70-79 ans, 80 ans et plus)
- population féminine selon les mêmes tranches d'âge
- catégorie socioprofessionnelle de l'individu en huit postes, les retraités étant reclassés dans leur profession d'origine (agriculteurs actifs et retraités, professions indépendantes en activité ou retraités, cadres et professions intermédiaires actifs, ouvriers et employés actifs, anciens cadres et professions intermédiaires, anciens ouvriers et employés, élèves-étudiants, autres inactifs)
- nombre d'individus de 6 ans ou plus du ménage croisé respectivement avec la catégorie de commune de résidence, le type de ménage, le nombre de voitures dont il dispose.

6.2.1 L'estimateur synthétique de type régression

La nomenclature en 14 postes des communes distingue Paris et sa banlieue de toutes les autres zones de résidence. La spécificité francilienne se trouve ainsi prise en compte par le modèle estimé dans l'échantillon national complet. La part modale des déplacements en transport collectif et en voiture, tout comme le temps de transport journalier, qui atteignent des niveaux très spécifiques en Ile de France, sont sensiblement mieux estimés dans cette région avec ce nouveau modèle (figure 11). L'estimateur synthétique de la fréquence des déplacements en transport en commun, en voiture comme conducteur, et celui de la durée des déplacements par personne et par jour, se situe dans l'intervalle de confiance de l'estimateur de référence en Ile de France, contrairement à l'estimateur précédent. Dans les autres régions avec extension, l'amélioration est plus modeste et n'empêche pas un biais d'estimation par rapport à la valeur prise en référence.

Les deux modèles produisent des estimations quasiment identiques pour les variables les moins corrélées à la région Ile de France ou à l'âge des personnes. Le changement de modèle est en particulier sans effet sur l'estimation de la fréquence des déplacements par motif, lorsque celui-ci n'est pas professionnel.

6.2.2 *L'estimateur par modèle mixte avec un effet régional aléatoire*

La régression avec effets mixtes prend déjà en compte une partie de la singularité francilienne. Le modèle détaillé modifie par conséquent moins nettement le résultat obtenu précédemment avec la même méthode. L'introduction de tranches d'âge détaillées améliore néanmoins l'estimation de la fréquence des déplacements liés au travail habituel, dans les régions avec extension (figure 12), à l'exception de la région Midi-Pyrénées, proche de la moyenne nationale, où l'estimateur synthétique apparaît plus efficace.

6.3 Conclusion

Les figures 13 et 14 comparent les estimateurs indirects obtenus avec le modèle détaillé, respectivement sur l'échantillon national complet et sur l'échantillon hors Ile de France, à l'estimateur direct national, dans les cinq régions ayant fait une extension, à l'estimateur national et à l'estimateur direct par calage régional, dans les autres régions. Dans les régions avec extension, l'estimation par modèle mixte est celle qui s'écarte le moins de l'estimation de référence. Dans les autres régions, cette méthode produit aussi les estimations qui sortent le moins souvent de l'intervalle de confiance de l'estimation directe par les pondérations nationales.

Figure 5 - Estimateurs synthétiques par régression nationale et par régression hors Ile de France, modèle simplifié Résultats dans les régions avec extension d'échantillon

Lecture du graphique

Les valeurs représentées sont des moyennes journalières par habitant de 6 ans ou plus, qu'il soit mobile ou non le jour de référence.

Pour chaque variable représentée dans une région, le segment symbolise l'intervalle de confiance à 95 % autour du vrai total estimé par l'estimateur direct. Le centre du segment est la moyenne régionale estimée avec les pondérations nationales, identique à celles d'un calage régional dans les régions avec extension (modèle 1 dans la section 4.2.1).

Le modèle simplifié est celui décrit en section 6.1.

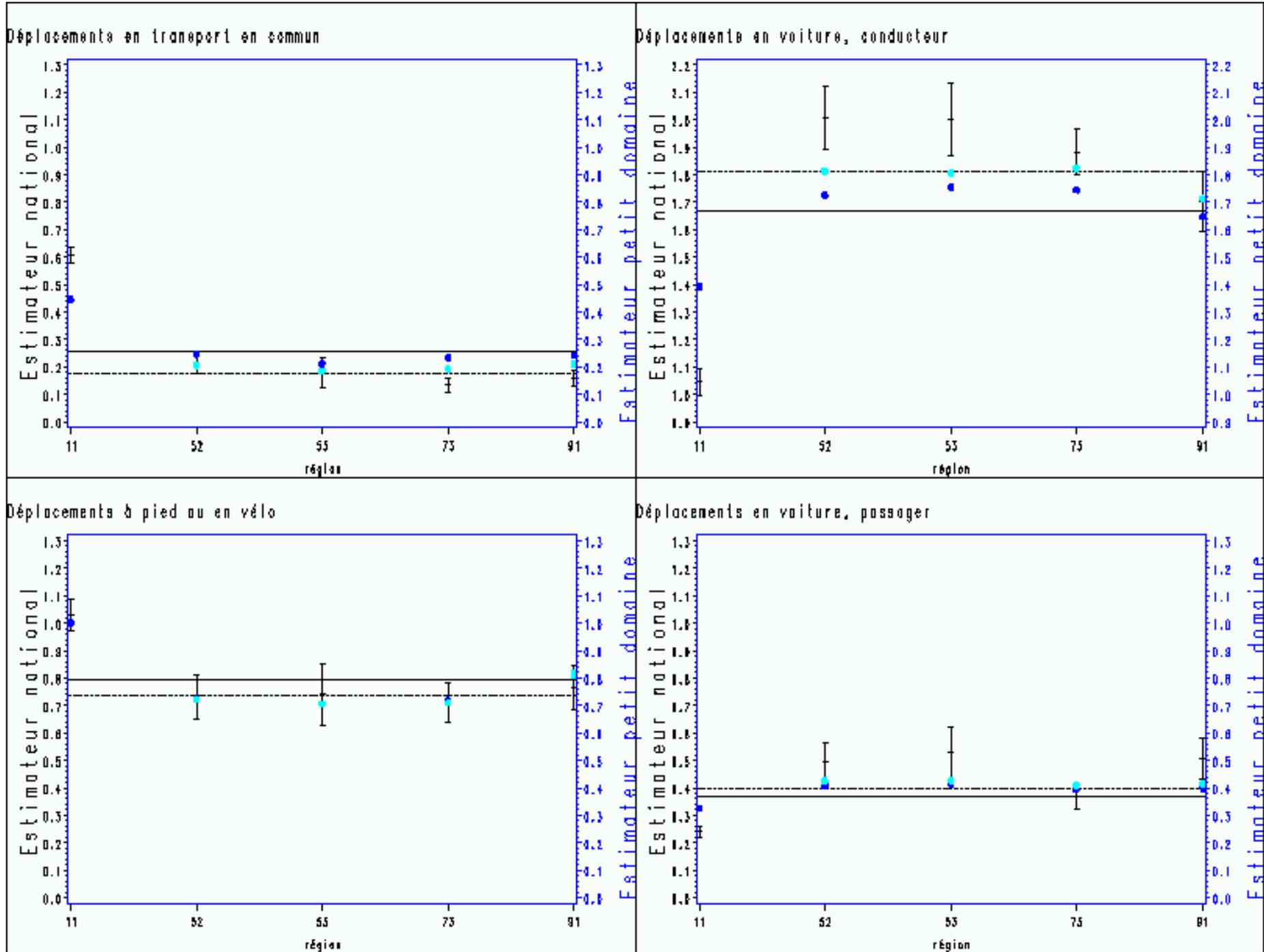
Les pastilles bleu sombre représentent l'estimateur synthétique de type régression (modèle 4 dans la section 4.2.2) dans chaque région, les coefficients **b** étant estimés dans l'ensemble de l'échantillon national.

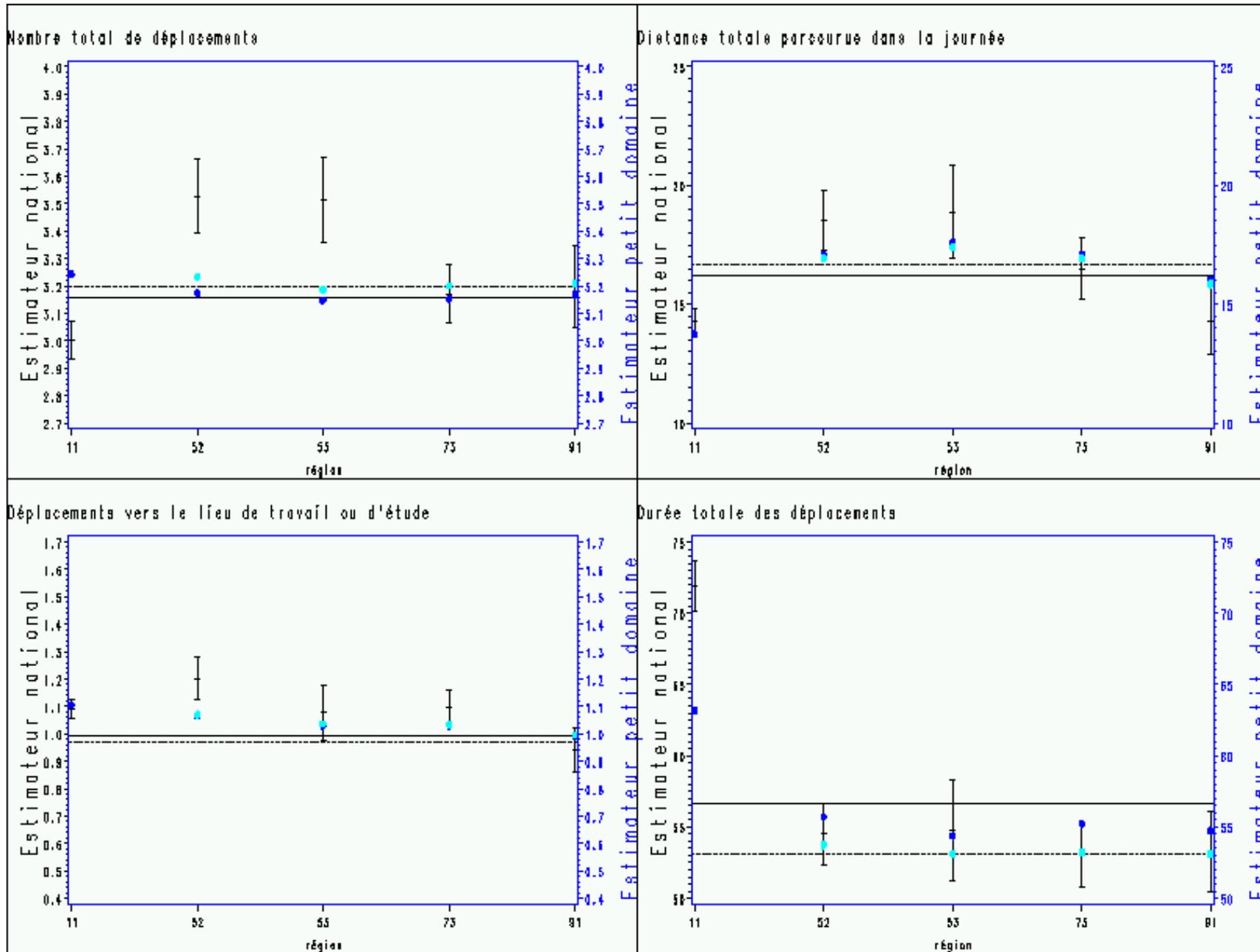
Les pastilles bleu ciel représentent l'estimateur du modèle 4 dans chaque région, les coefficients **b** étant estimés dans l'échantillon national hors Ile de France.

Les deux lignes horizontales noires représentent la moyenne nationale, respectivement en France métropolitaine (trait plein) ou en France de province (trait pointillé), calculée avec les pondérations nationales.

Liste des codes région

11	Ile de France
52	Pays de la Loire
53	Bretagne
73	Midi-Pyrénées
91	Languedoc-Roussillon





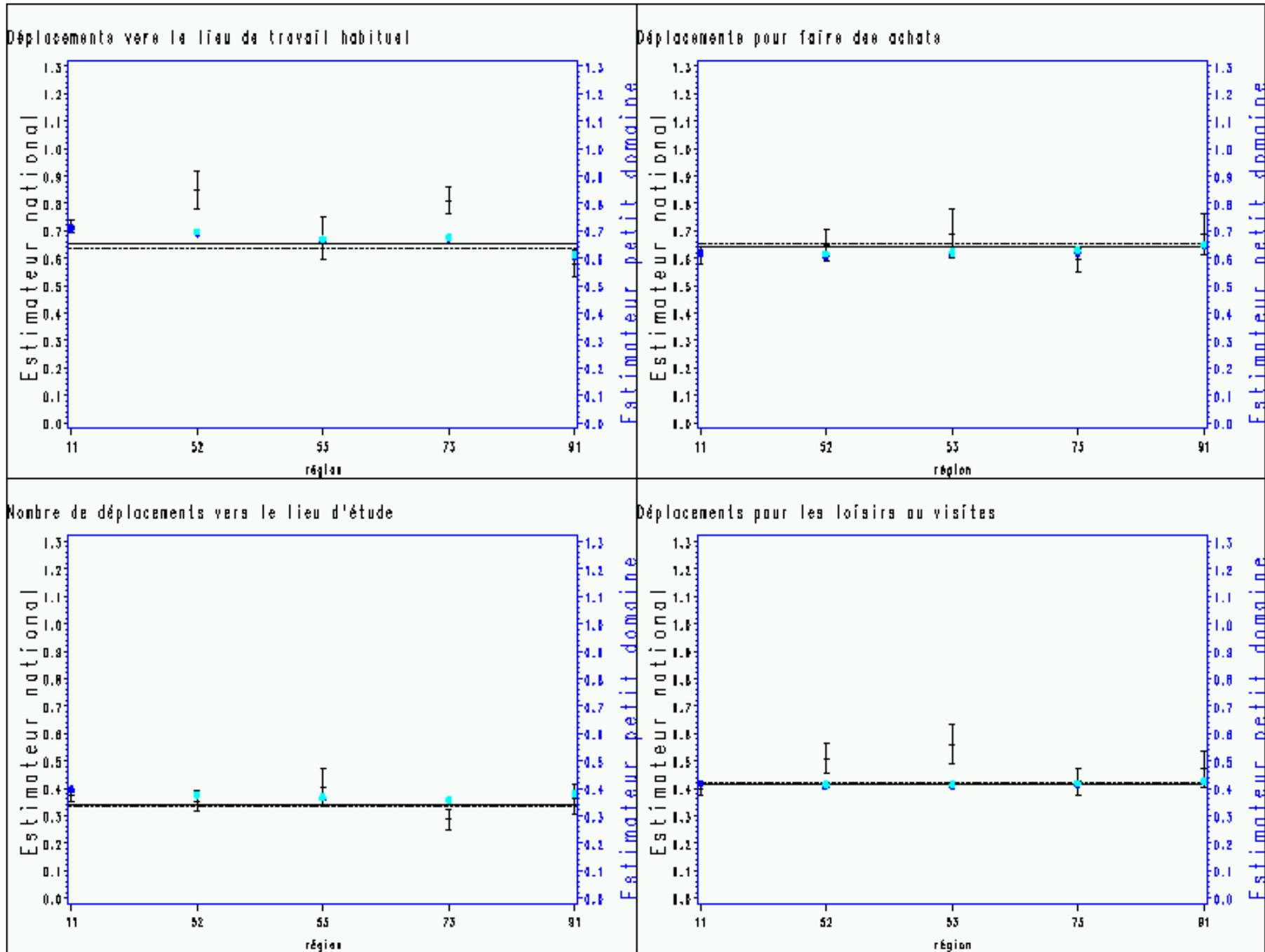


Figure 6 - Estimateurs synthétiques par régression nationale et par régression hors Ile de France, modèle simplifié
Résultats dans les régions sans extension d'échantillon

Lecture du graphique

Les valeurs représentées sont des moyennes par habitant de 6 ans ou plus, qu'il soit mobile ou non le jour de référence.

Pour chaque variable représentée dans une région, le segment symbolise l'intervalle de confiance à 95 % autour du vrai total estimé par l'estimateur direct. Le centre du segment est la moyenne régionale estimée avec les pondérations nationales.

Le modèle simplifié est celui décrit en section 6.1.

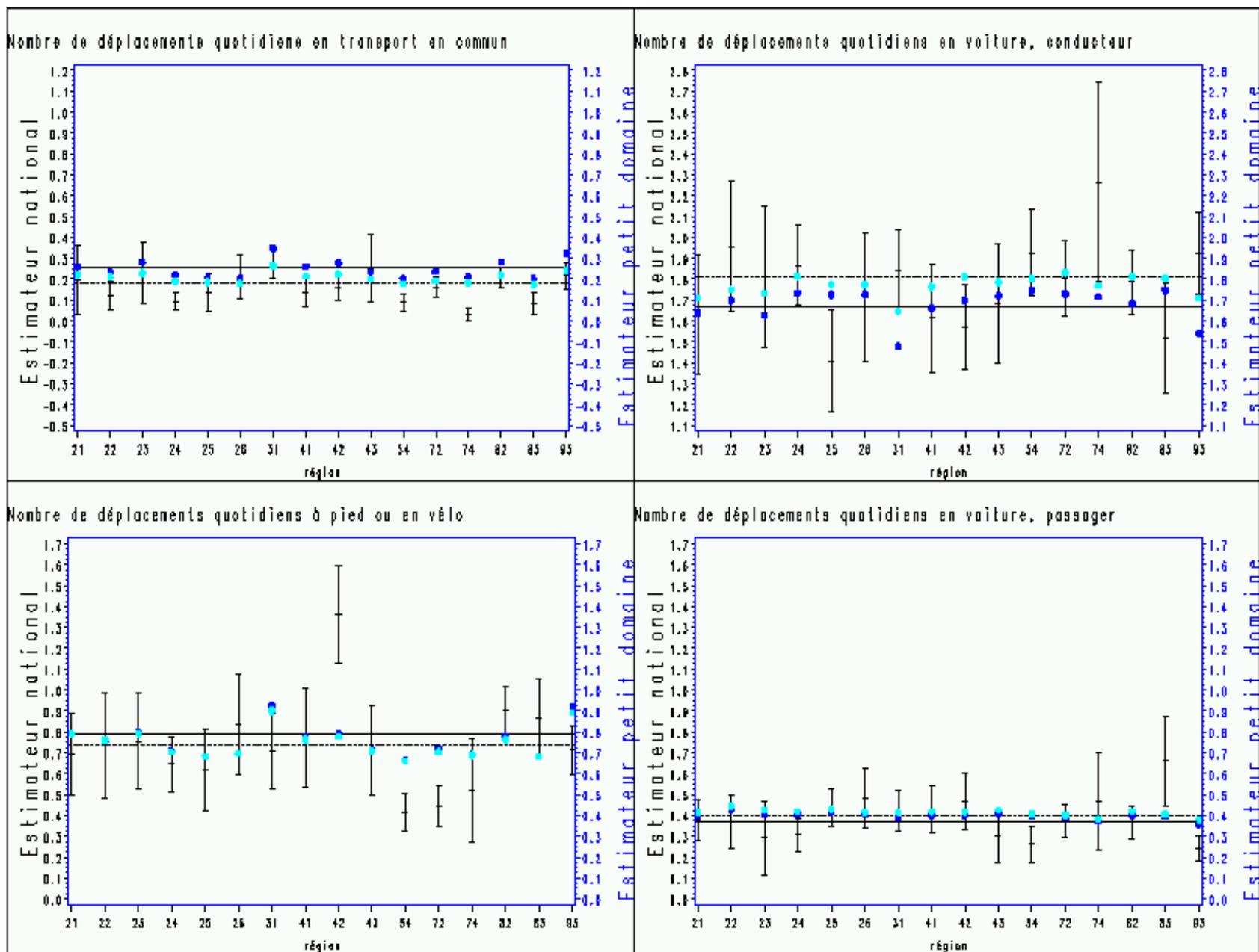
Les pastilles bleu sombre représentent l'estimateur synthétique de type régression (modèle 4 dans la section 4.2.2) dans chaque région, les coefficients **b** étant estimés dans l'ensemble de l'échantillon national.

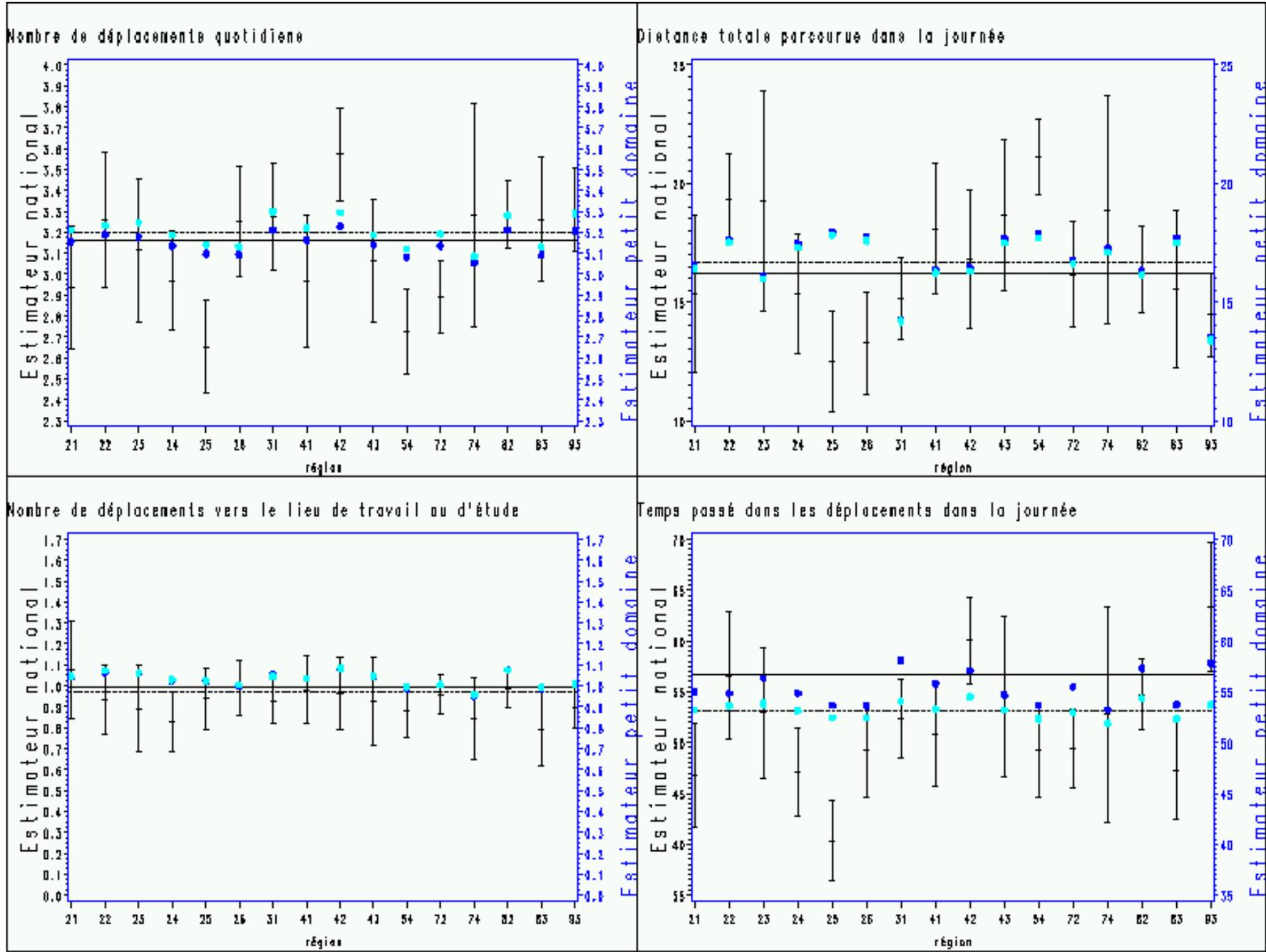
Les pastilles bleu ciel représentent l'estimateur du modèle 4 dans chaque région, les coefficients **b** étant estimés dans l'échantillon national hors Ile de France.

Les deux lignes horizontales noires représentent la moyenne nationale, respectivement en France métropolitaine (trait plein) ou en France de province (trait pointillé), calculée avec les pondérations nationales.

Liste des codes région

21	Champagne-Ardenne	42	Alsace
22	Picardie	43	Franche-Comté
23	Haute-Normandie	54	Poitou-Charentes
24	Centre	72	Aquitaine
25	Basse-Normandie	74	Limousin
26	Bourgogne	82	Rhône-Alpes
31	Nord	83	Auvergne
41	Lorraine	93	Provence Côte d Azur





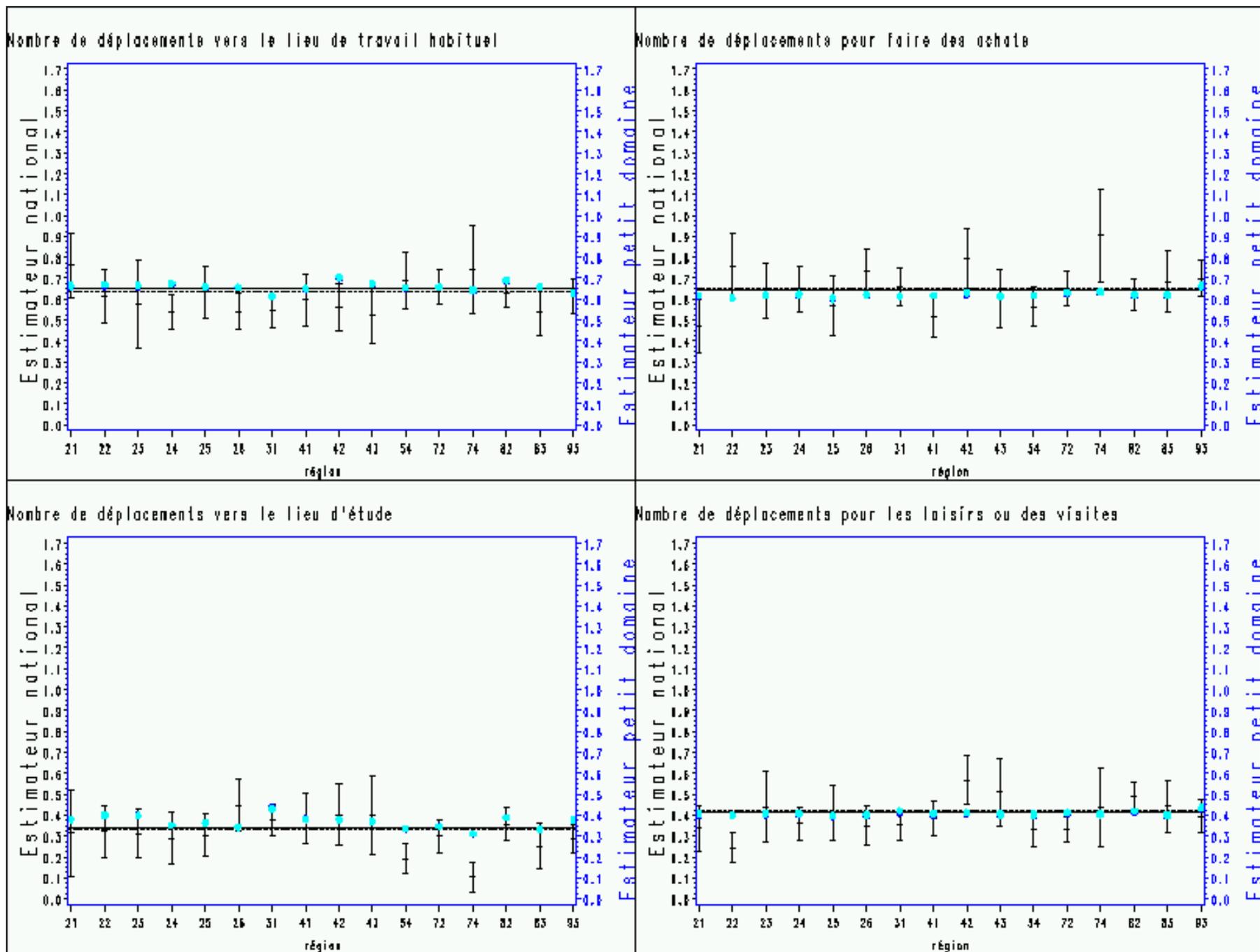


Figure 7 - Estimateur par modèle mixte avec effet régional aléatoire, modèle simplifié Régions avec extension

Lecture du graphique

Les valeurs représentées sont des moyennes par habitant de 6 ans ou plus, qu'il soit mobile ou non le jour de référence.

Pour chaque variable représentée dans une région, le segment symbolise l'intervalle de confiance à 95 % autour du vrai total estimé par l'estimateur direct. Le centre du segment est la moyenne régionale estimée avec les pondérations nationales, identique à celles d'un calage régional dans les régions avec extension (modèle 1 dans la section 4.2.1).

Le modèle simplifié est celui décrit en section 6.1.

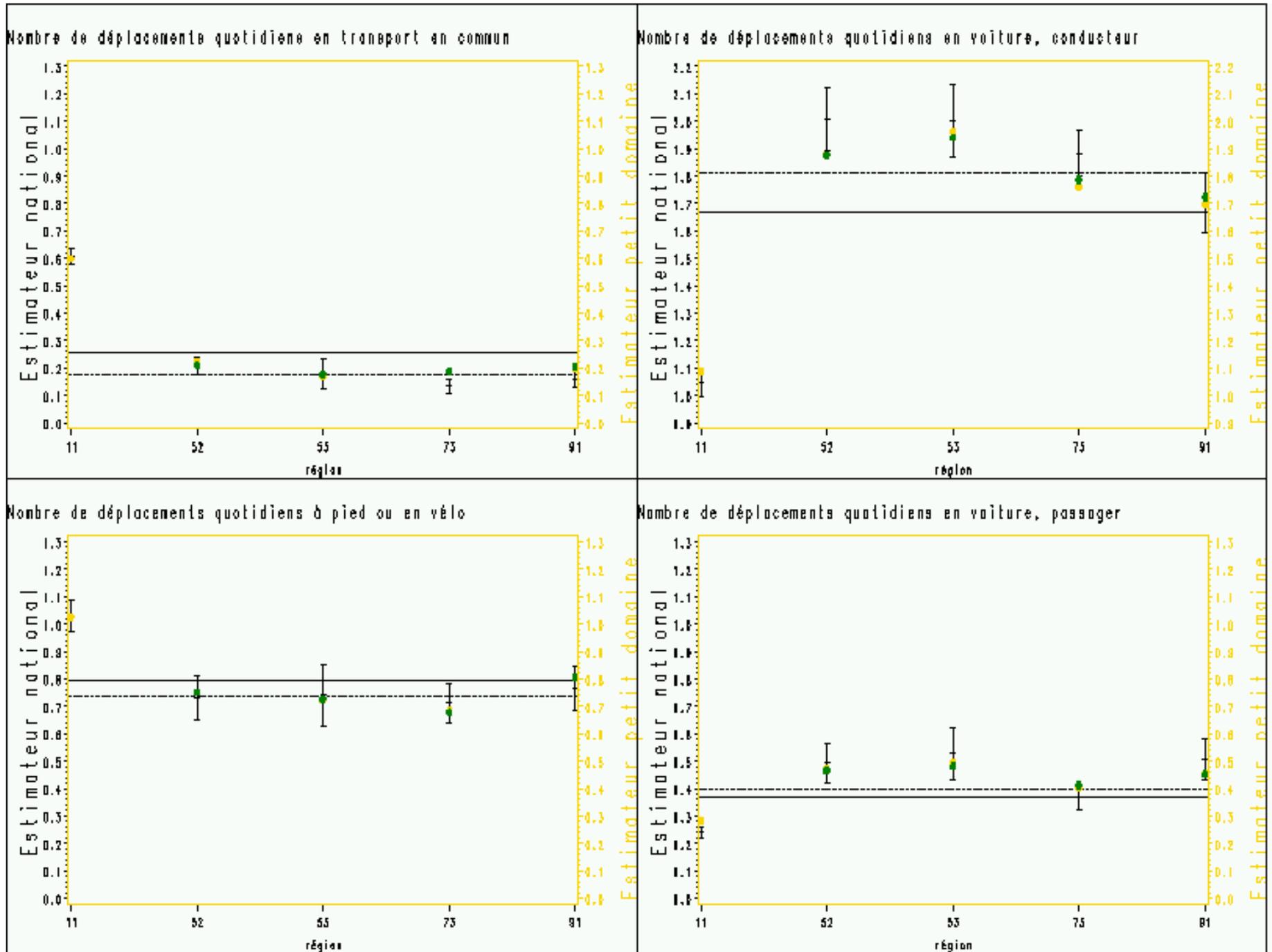
Les pastilles vertes représentent l'estimateur par modèle mixte dans chaque région (modèle 9 de la section 4.2.3), les coefficients **b** étant estimés dans l'ensemble de l'échantillon national.

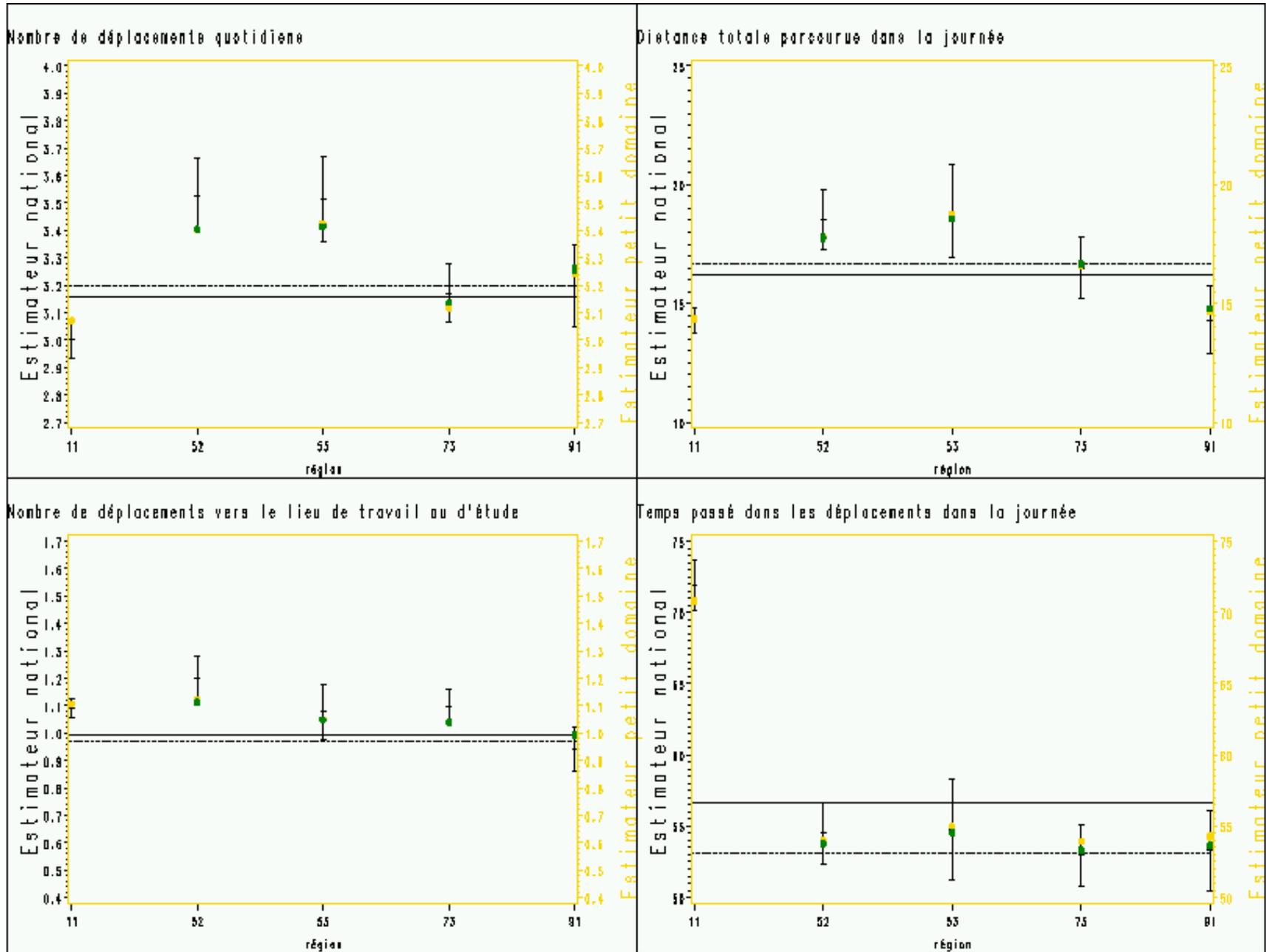
Les pastilles jaunes représentent l'estimateur régional du modèle 9, les coefficients **b** étant estimés dans l'échantillon national hors Ile de France.

Les deux lignes horizontales noires représentent la moyenne nationale, respectivement en France métropolitaine (trait plein) ou en France de province (trait pointillé), calculée avec les pondérations nationales.

Liste des codes région

11	Ile de France
52	Pays de la Loire
53	Bretagne
73	Midi-Pyrénées
91	Languedoc-Roussillon





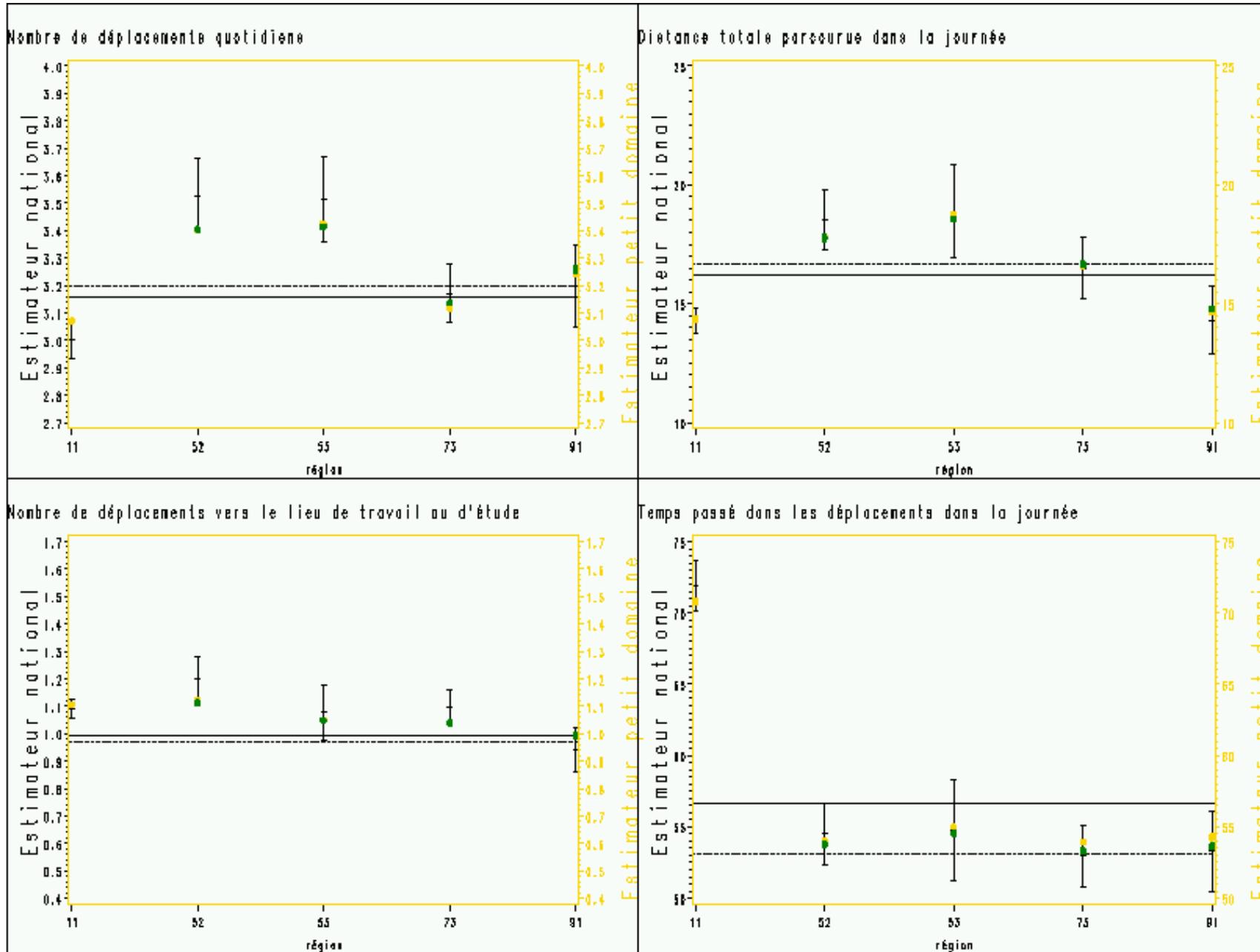


Figure 8 - Estimateur par modèle mixte avec effet régional aléatoire, modèle simplifié Régions sans extension d'échantillon

Lecture du graphique

Les valeurs représentées sont des moyennes par habitant de 6 ans ou plus, qu'il soit mobile ou non le jour de référence.

Pour chaque variable représentée dans une région, le segment symbolise l'intervalle de confiance à 95 % autour du vrai total estimé par l'estimateur direct. Le centre du segment est la moyenne régionale estimée avec les pondérations nationales.

Le modèle simplifié est celui décrit en section 6.1.

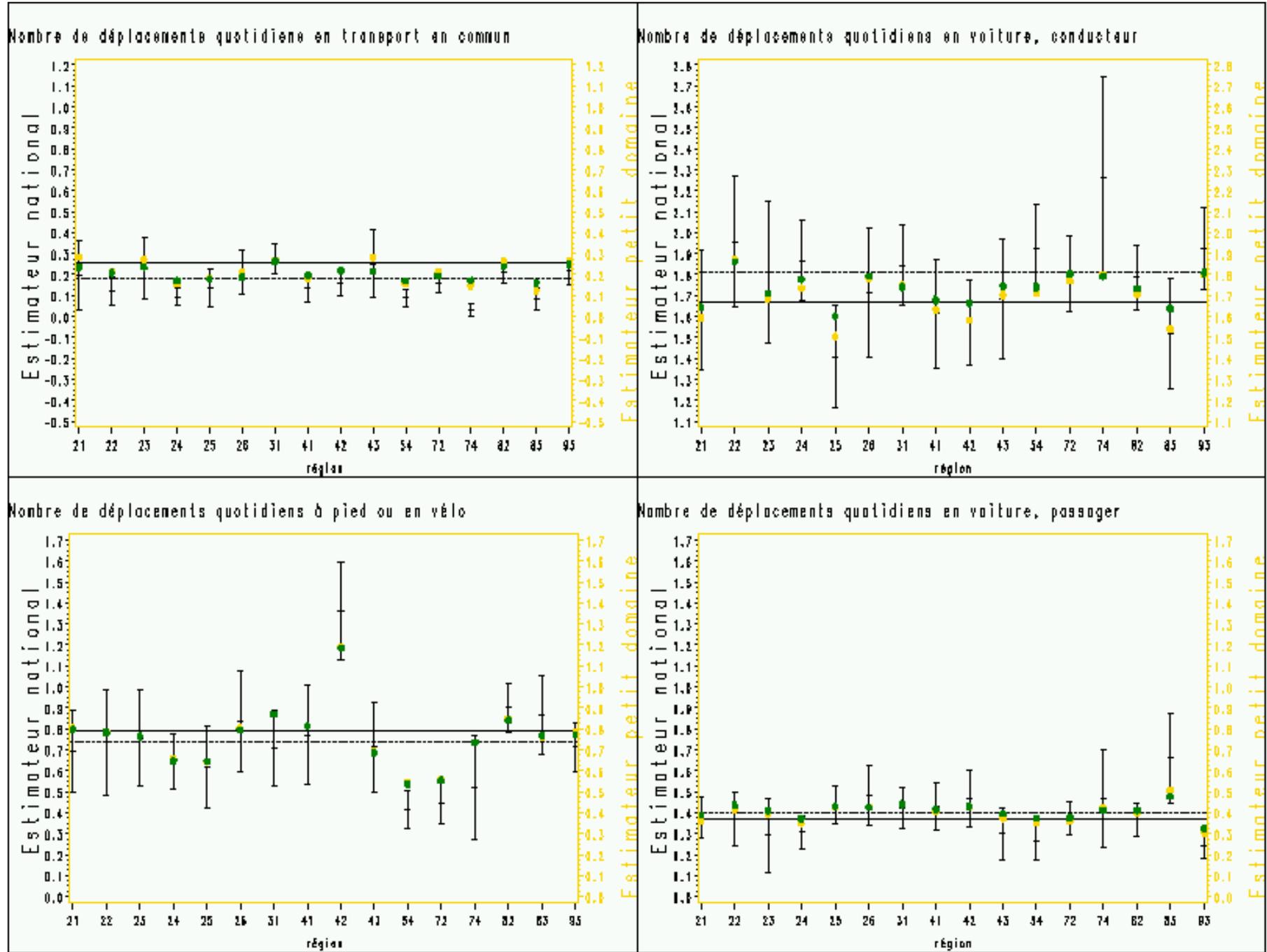
Les pastilles vertes représentent l'estimateur par modèle mixte dans chaque région (modèle 9 de la section 4.2.3), les coefficients **b** étant estimés dans l'ensemble de l'échantillon national.

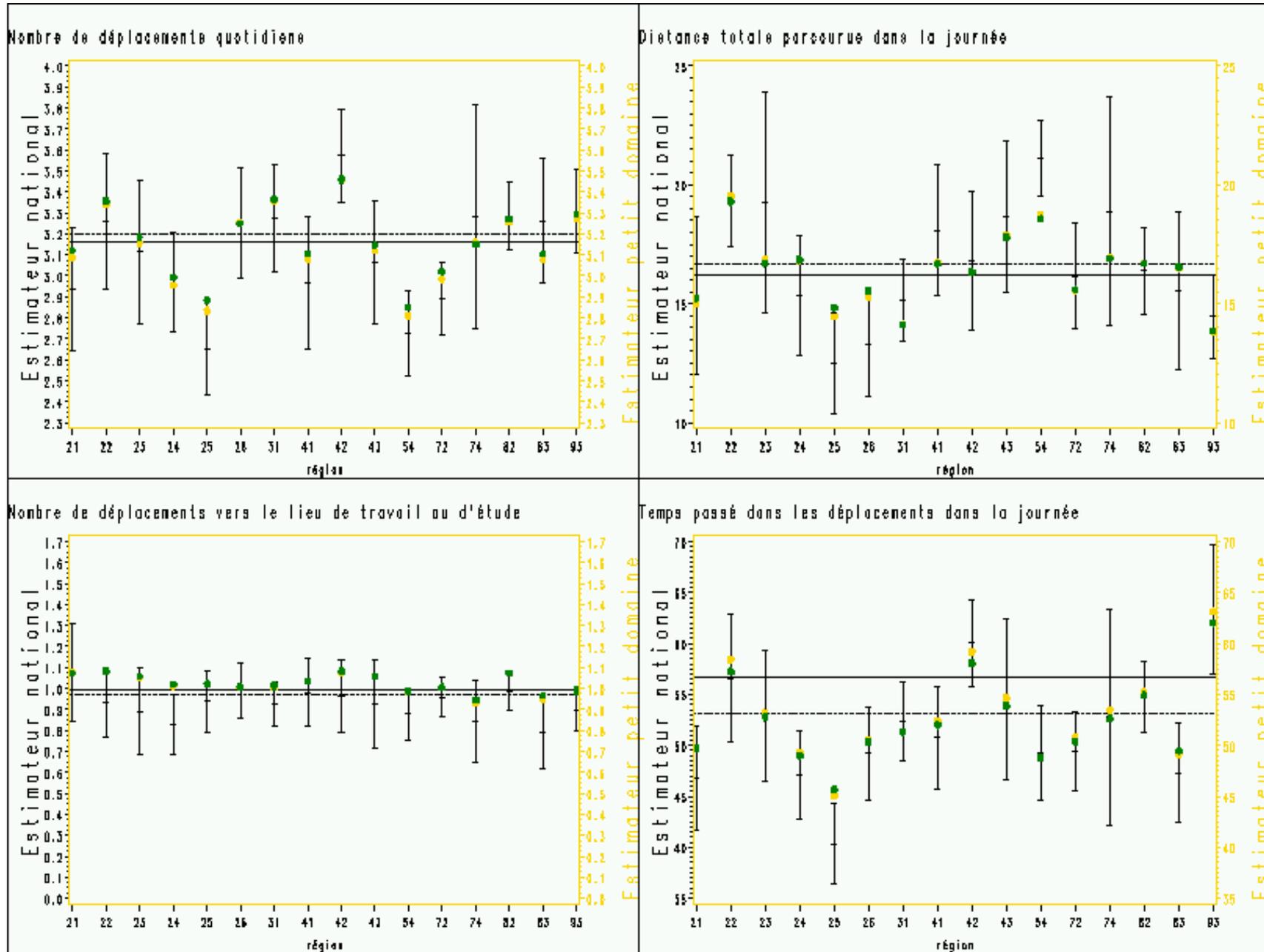
Les pastilles jaunes représentent l'estimateur régional du modèle 9, les coefficients **b** étant estimés dans l'échantillon national hors Ile de France.

Les deux lignes horizontales noires représentent la moyenne nationale, respectivement en France métropolitaine (trait plein) ou en France de province (trait pointillé), calculée avec les pondérations nationales.

Liste des codes région

21	Champagne-Ardenne	42	Alsace
22	Picardie	43	Franche-Comté
23	Haute-Normandie	54	Poitou-Charentes
24	Centre	72	Aquitaine
25	Basse-Normandie	74	Limousin
26	Bourgogne	82	Rhône-Alpes
31	Nord	83	Auvergne
41	Lorraine	93	Provence Côte d Azur





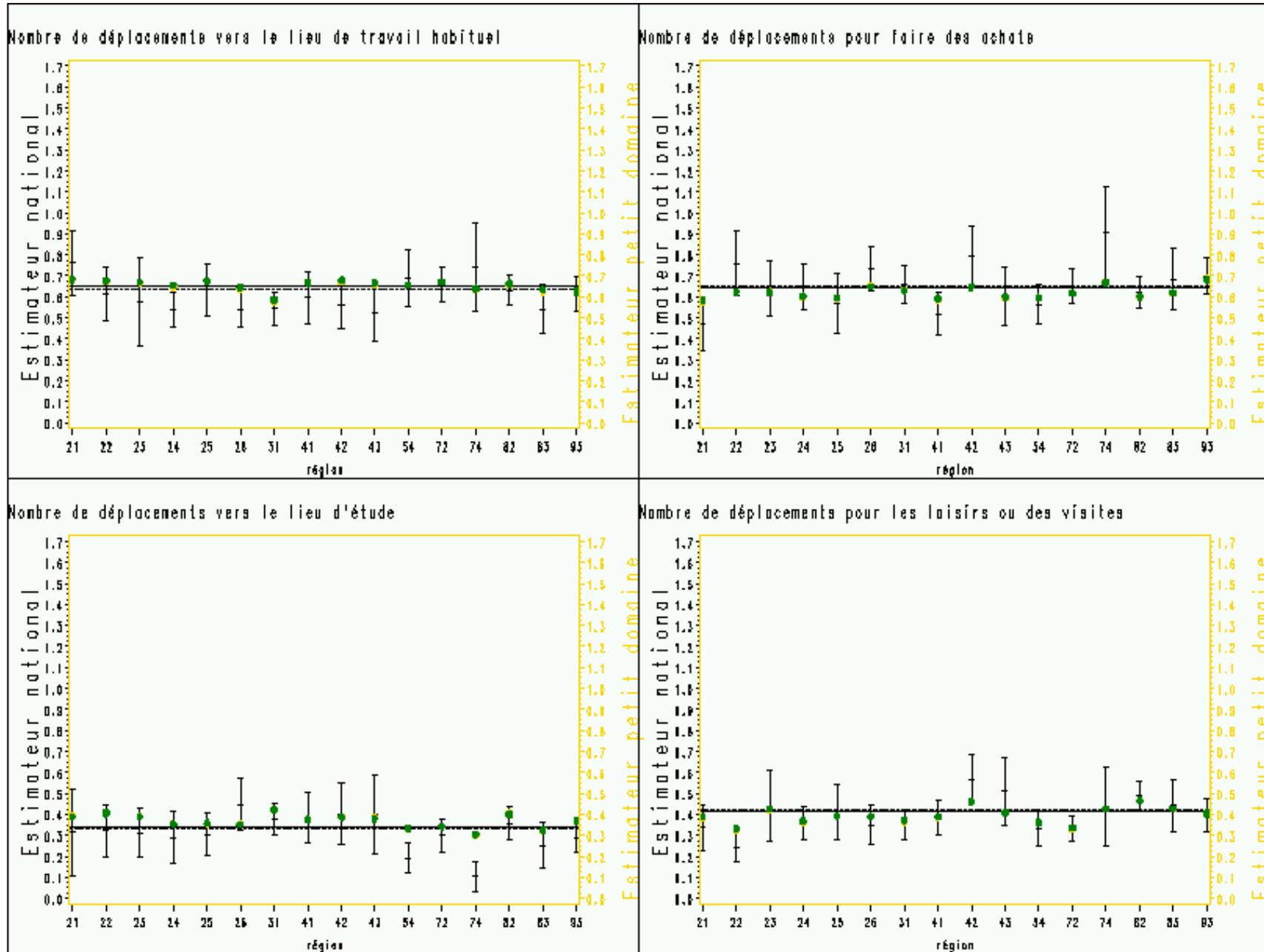


Figure 9 - Estimateur direct et estimateurs modélisés, modèle simplifié Régions avec extension d'échantillon

Lecture du graphique

Les valeurs représentées sont des moyennes par habitant de 6 ans ou plus, qu'il soit mobile ou non le jour de référence.

Pour chaque variable représentée dans une région, le segment symbolise l'intervalle de confiance à 95 % autour du vrai total estimé par l'estimateur direct. Le centre du segment est la moyenne régionale estimée avec les pondérations nationales, identique à celles d'un calage régional dans les régions avec extension (modèle 1 dans la section 4.2.1).

Le modèle simplifié est celui décrit en section 6.1.

Les pastilles bleu ciel représentent l'estimateur direct modifié (modèle 2 de la section 4.2.1), les coefficients étant estimés dans l'échantillon national hors Ile de France.

Les pastilles bleu sombre représentent l'estimateur synthétique de type régression (modèle 4 de la section 4.2.2), les coefficients étant estimés dans l'échantillon national hors Ile de France.

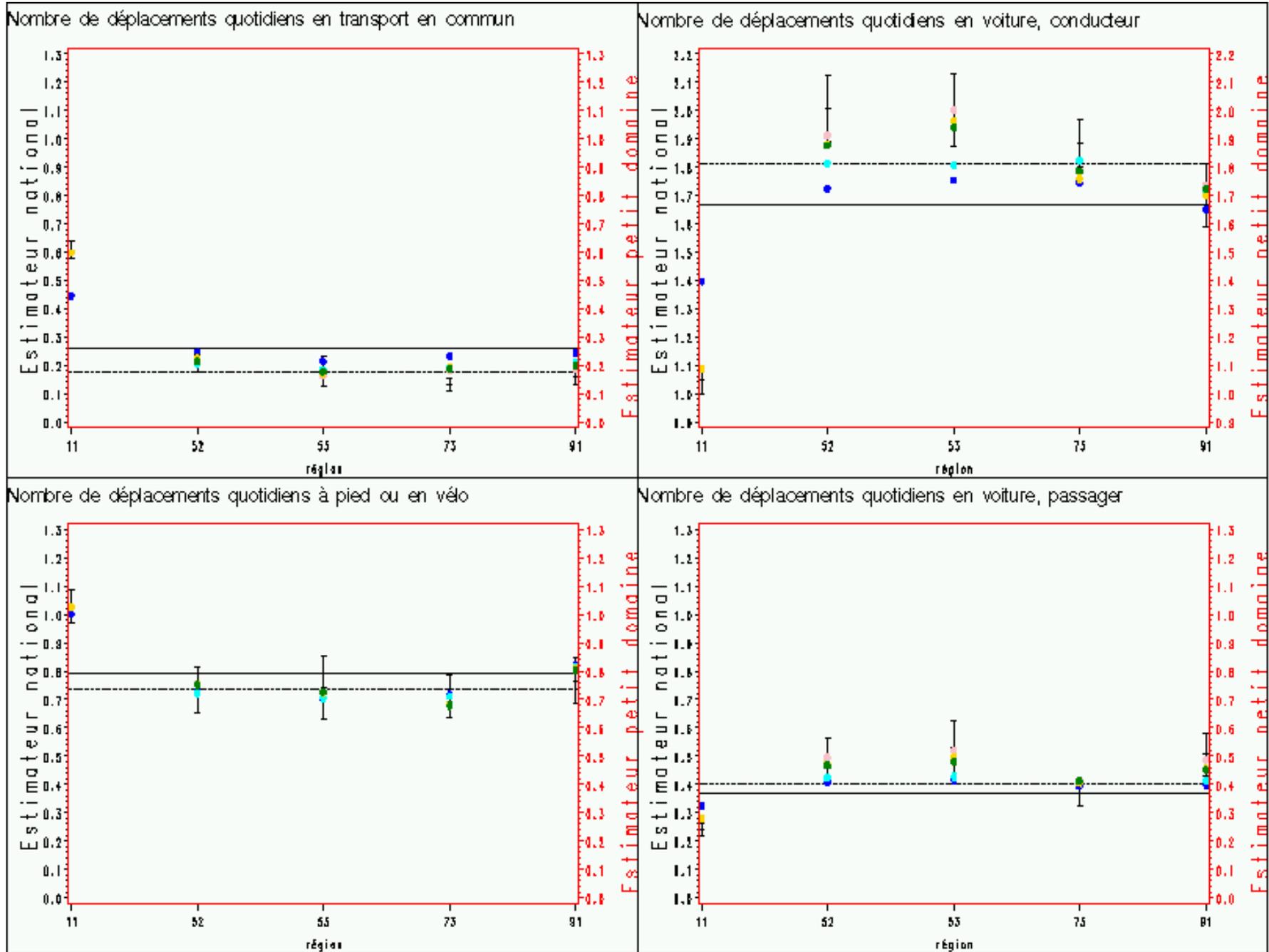
Les pastilles vertes représentent l'estimation par modèle mixte à effet régional aléatoire (modèle 9 de la section 4.2.3), les coefficients étant estimés dans l'échantillon national hors Ile de France.

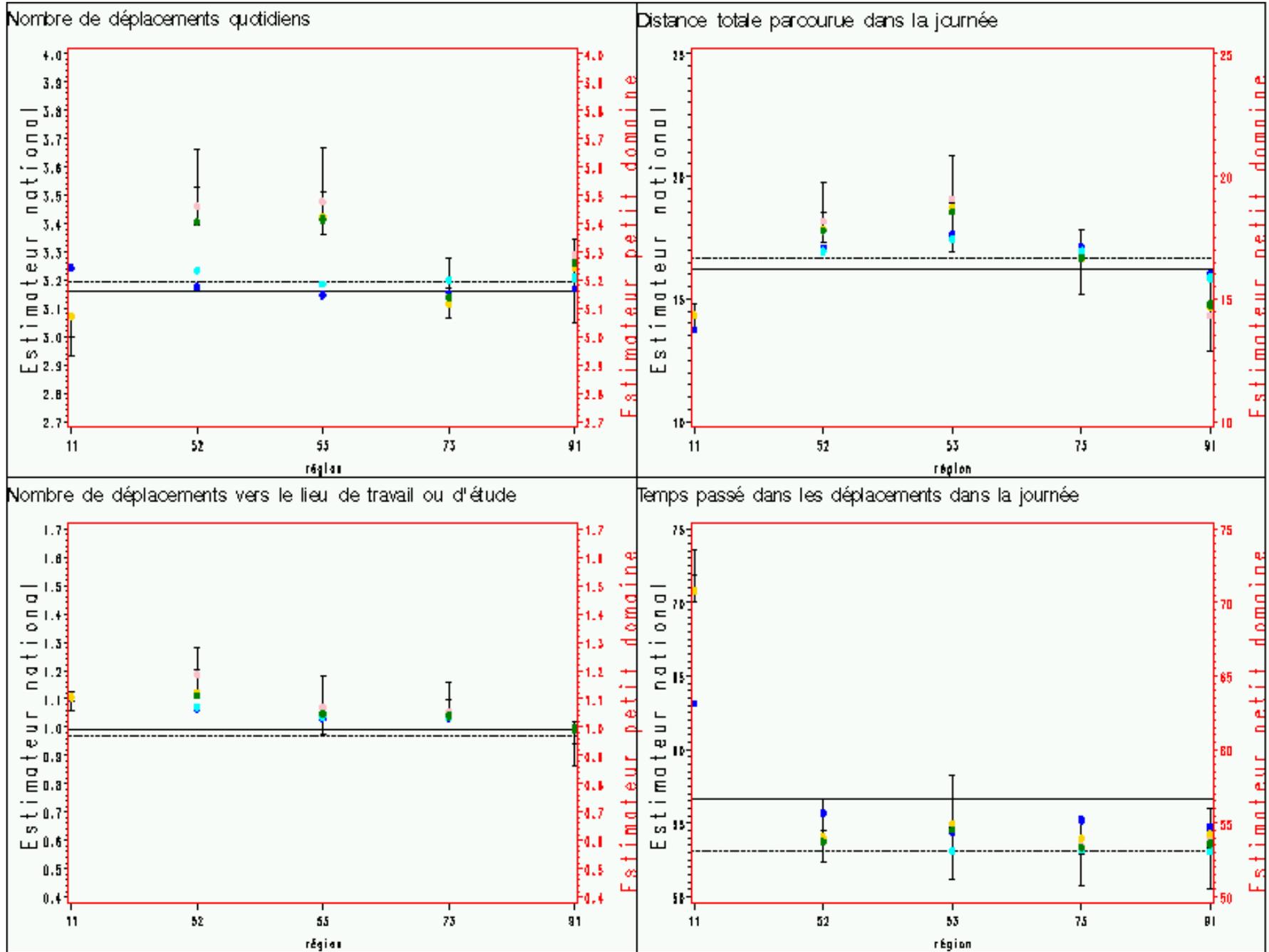
Les pastilles jaunes représentent l'estimation du modèle 9, les coefficients étant estimés dans l'échantillon national complet.

Les deux lignes horizontales noires représentent la moyenne nationale, respectivement en France métropolitaine (trait plein) ou en France de province (trait pointillé), calculée avec les pondérations nationales.

Liste des codes région

11	Ile de France
52	Pays de la Loire
53	Bretagne
73	Midi-Pyrénées
91	Languedoc-Roussillon





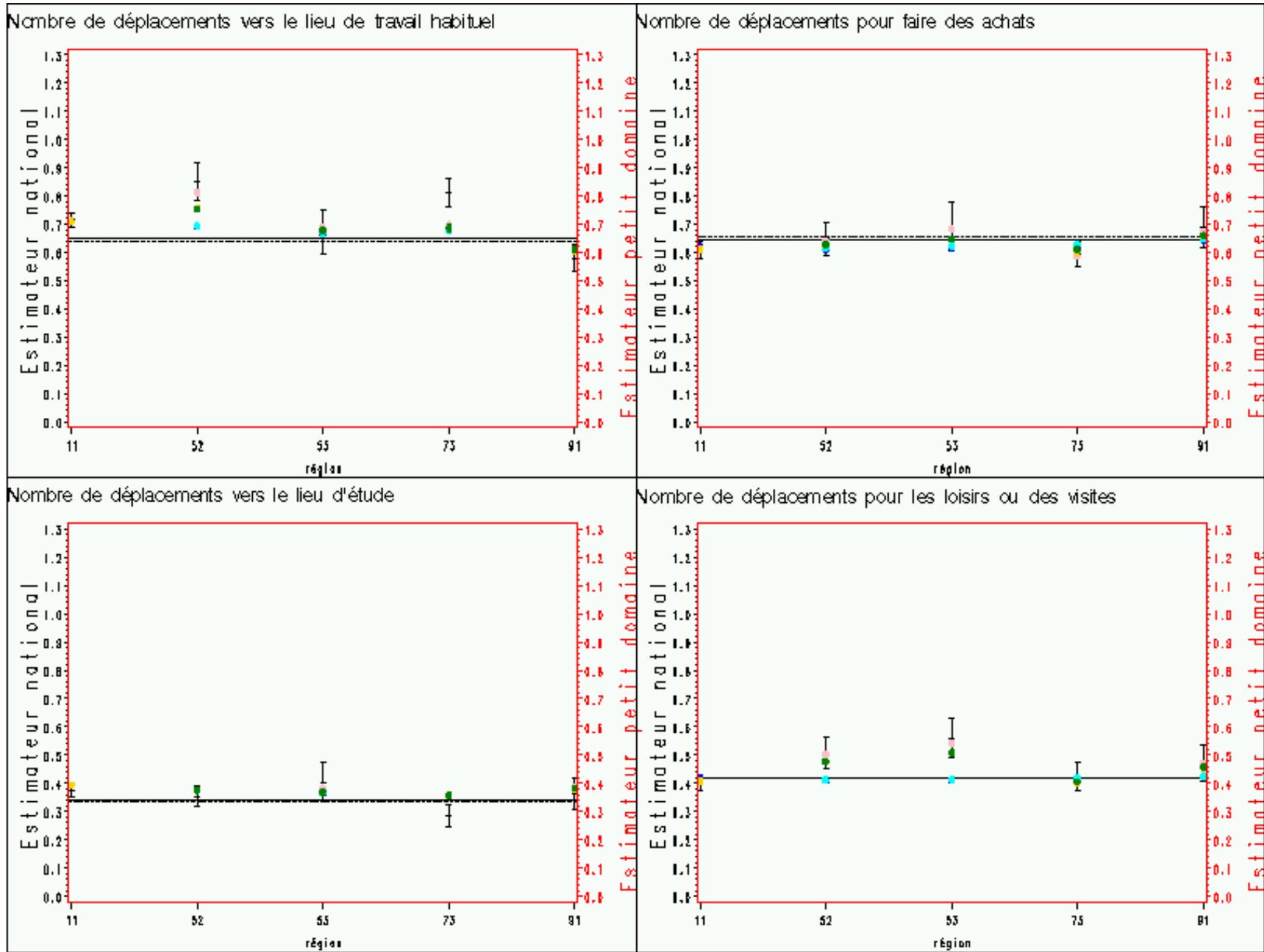


Figure 10 - Estimateur direct et estimateurs modélisés, modèle simplifié Régions sans extension d'échantillon

Lecture du graphique

Les valeurs représentées sont des moyennes par habitant de 6 ans ou plus, qu'il soit mobile ou non le jour de référence.

Pour chaque variable représentée dans une région, le segment symbolise l'intervalle de confiance à 95 % autour du vrai total estimé par l'estimateur direct. Le centre du segment est la moyenne régionale estimée avec les pondérations nationales.

Les pastilles rouges représentent l'estimation directe par calage séparé de chaque région sur des populations régionales (modèle 1 dans la section 4.2.1).

Les pastilles bleu ciel représentent l'estimateur direct modifié (modèle 2 de la section 4.2.1), les coefficients étant estimés dans l'échantillon national hors Ile de France.

Les pastilles bleu sombre représentent l'estimateur synthétique de type régression (modèle 4 de la section 4.2.2), les coefficients étant estimés dans l'échantillon national hors Ile de France.

Les pastilles vertes représentent l'estimation par modèle mixte à effet régional aléatoire (modèle 9 de la section 4.2.3), les coefficients étant estimés dans l'échantillon national hors Ile de France.

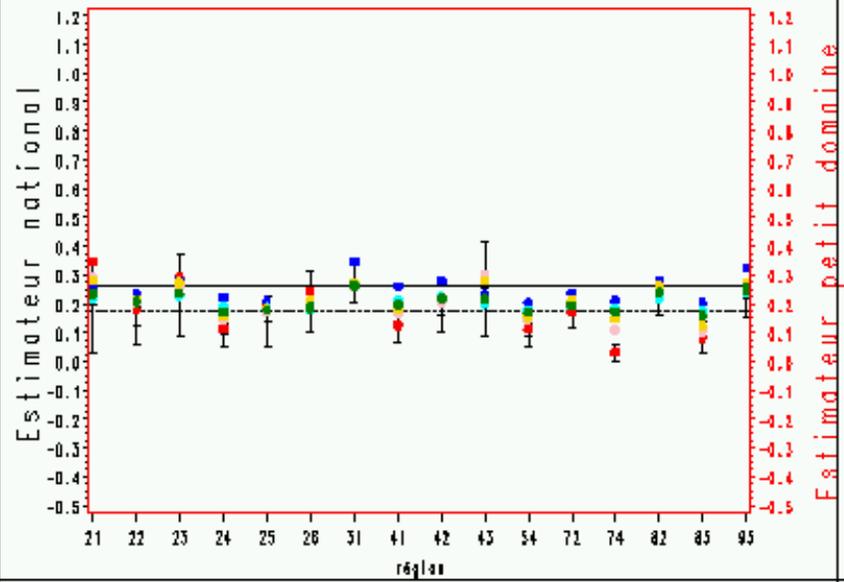
Les pastilles jaunes représentent l'estimation du modèle 9, les coefficients étant estimés dans l'échantillon national complet.

Les deux lignes horizontales noires représentent la moyenne nationale, respectivement en France métropolitaine (trait plein) ou en France de province (trait pointillé), calculée avec les pondérations nationales.

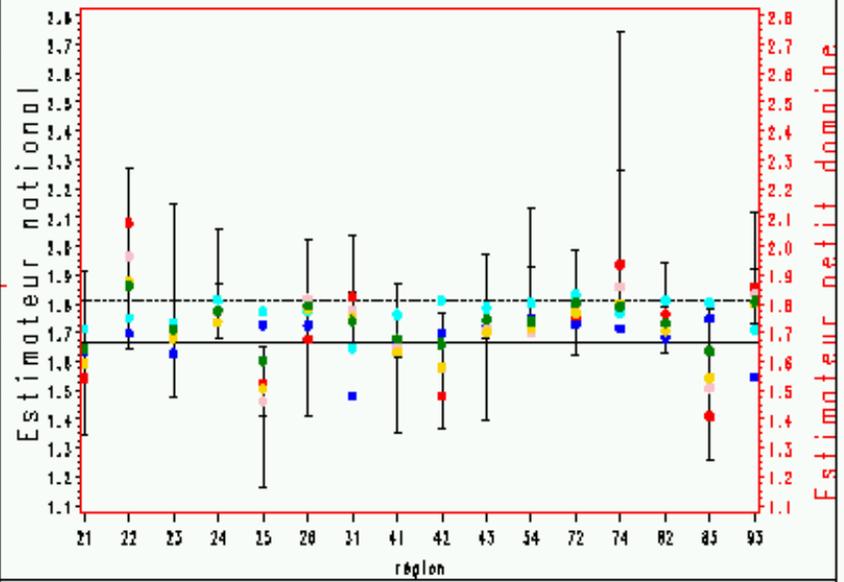
Liste des codes région

21	Champagne-Ardenne	42	Alsace
22	Picardie	43	Franche-Comté
23	Haute-Normandie	54	Poitou-Charentes
24	Centre	72	Aquitaine
25	Basse-Normandie	74	Limousin
26	Bourgogne	82	Rhône-Alpes
31	Nord	83	Auvergne
41	Lorraine	93	Provence Côte d Azur

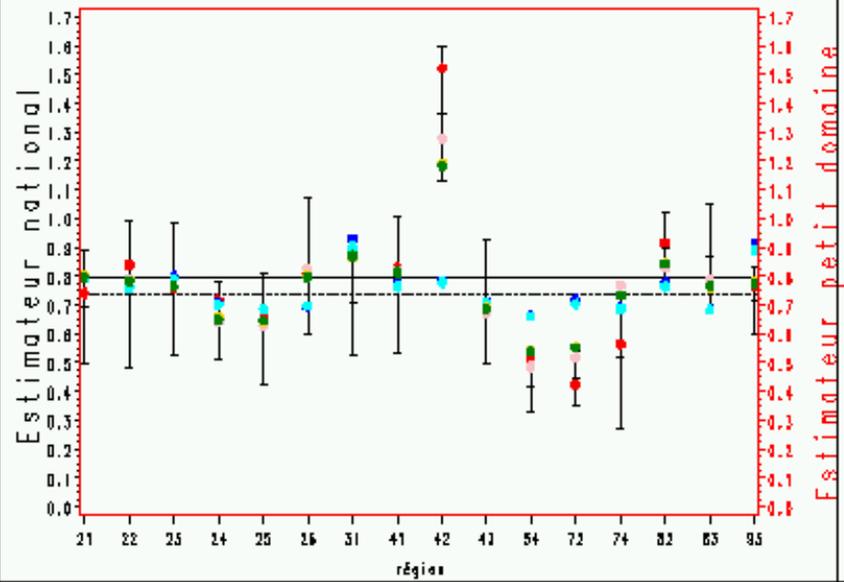
Déplacements en transport en commun



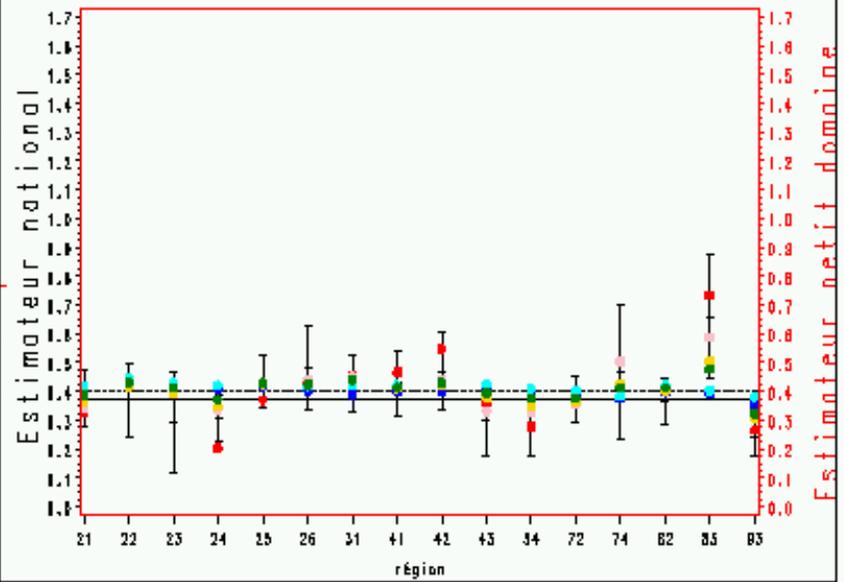
Déplacements en voiture, conducteur

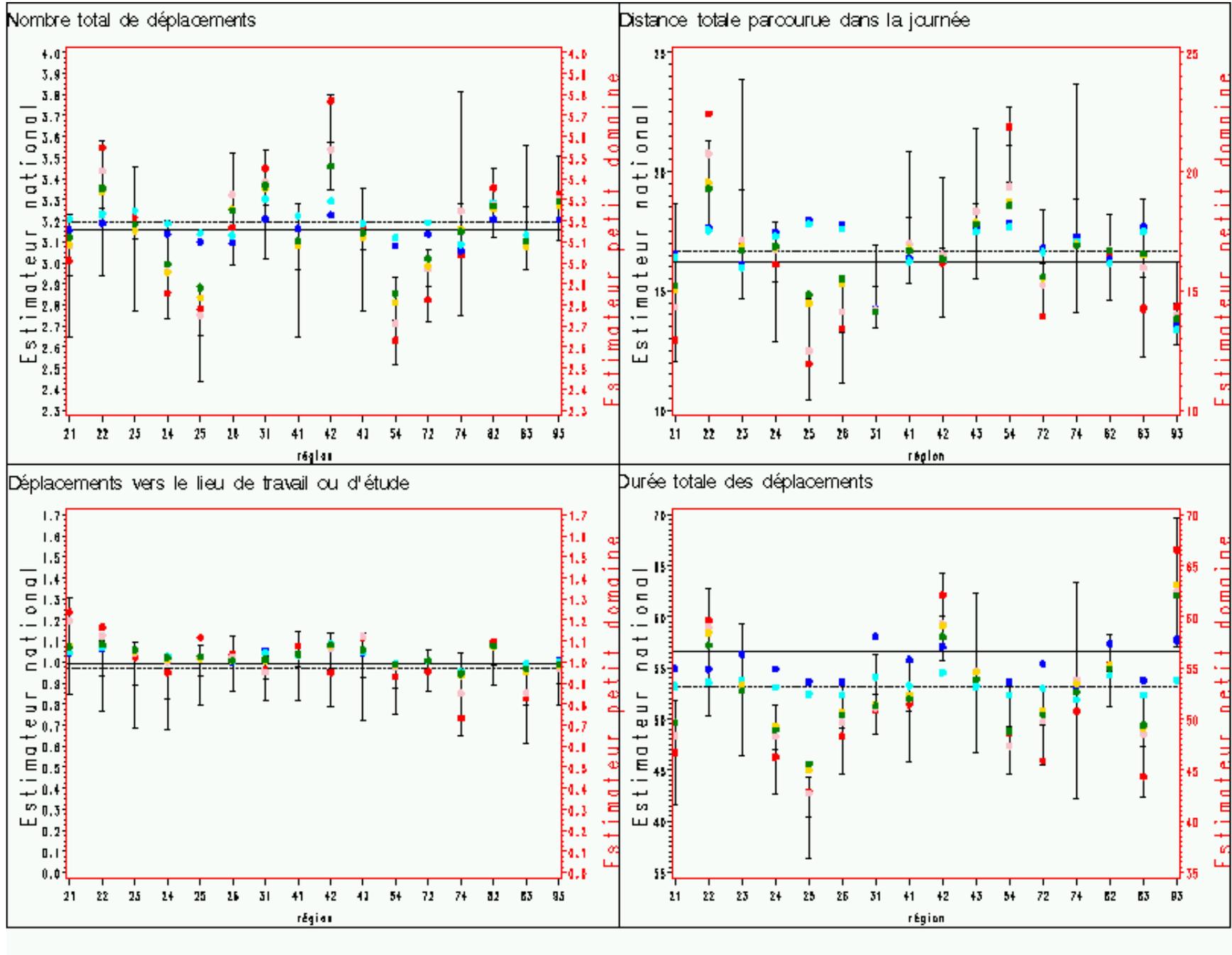


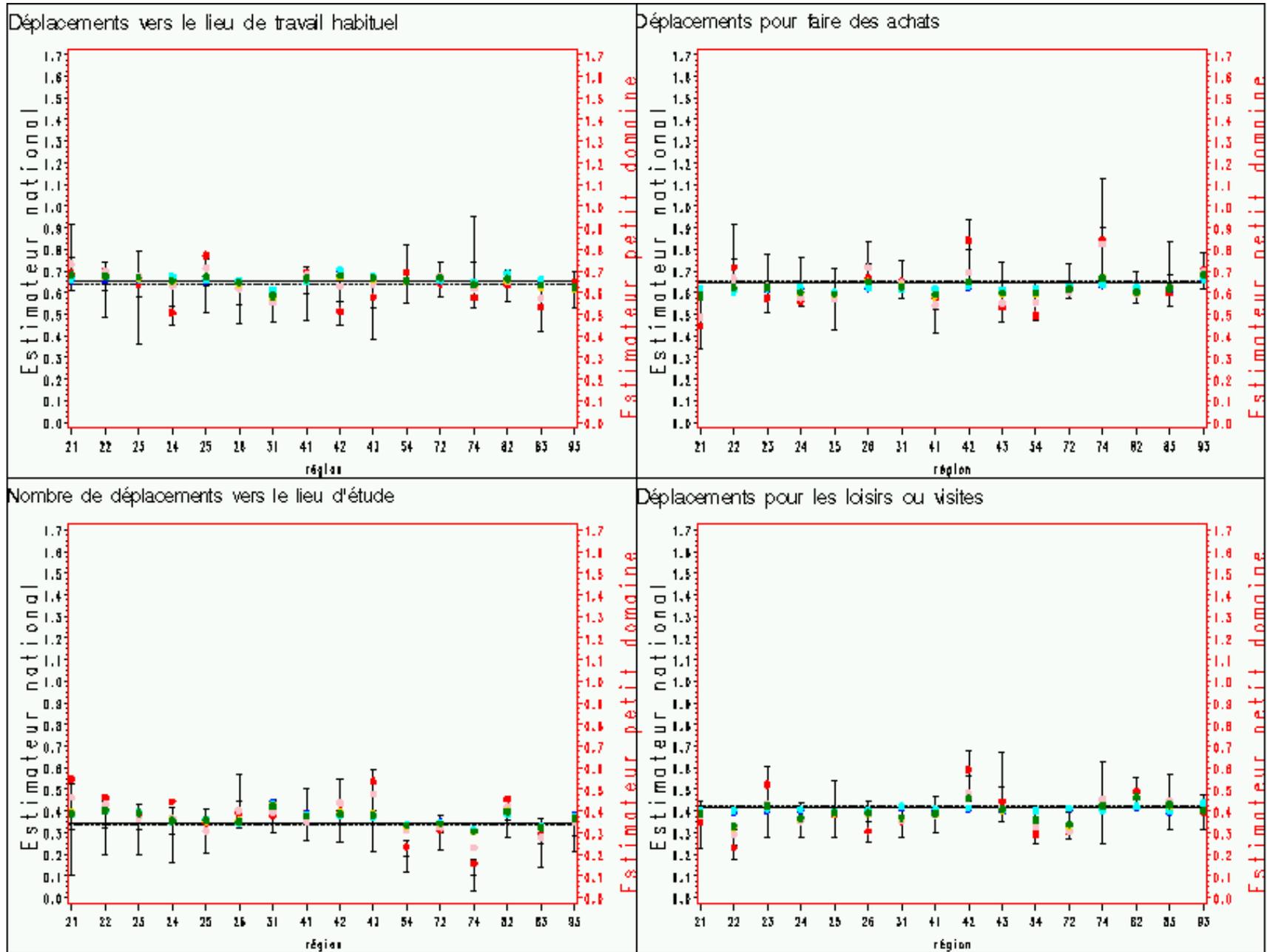
Déplacements à pied ou en vélo



Déplacements en voiture, passager







**Figure 11 - Estimateurs synthétiques de type régression selon le modèle explicatif
Résultats dans les régions avec extension d'échantillon**

Lecture du graphique

Les valeurs représentées sont des moyennes par habitant de 6 ans ou plus, qu'il soit mobile ou non le jour de référence.

Pour chaque variable représentée dans une région, le segment symbolise l'intervalle de confiance à 95 % autour du vrai total estimé par l'estimateur direct. Le centre du segment est la moyenne régionale estimée avec les pondérations nationales, identique à celles d'un calage régional dans les régions avec extension (modèle 1 dans la section 4.2.1).

Le modèle simplifié est celui décrit en section 6.1, le modèle détaillé est décrit en section 6.2.

Les pastilles bleu sombre représentent l'estimateur synthétique de type régression (modèle 4 de la section 4.2.2) avec le modèle détaillé.

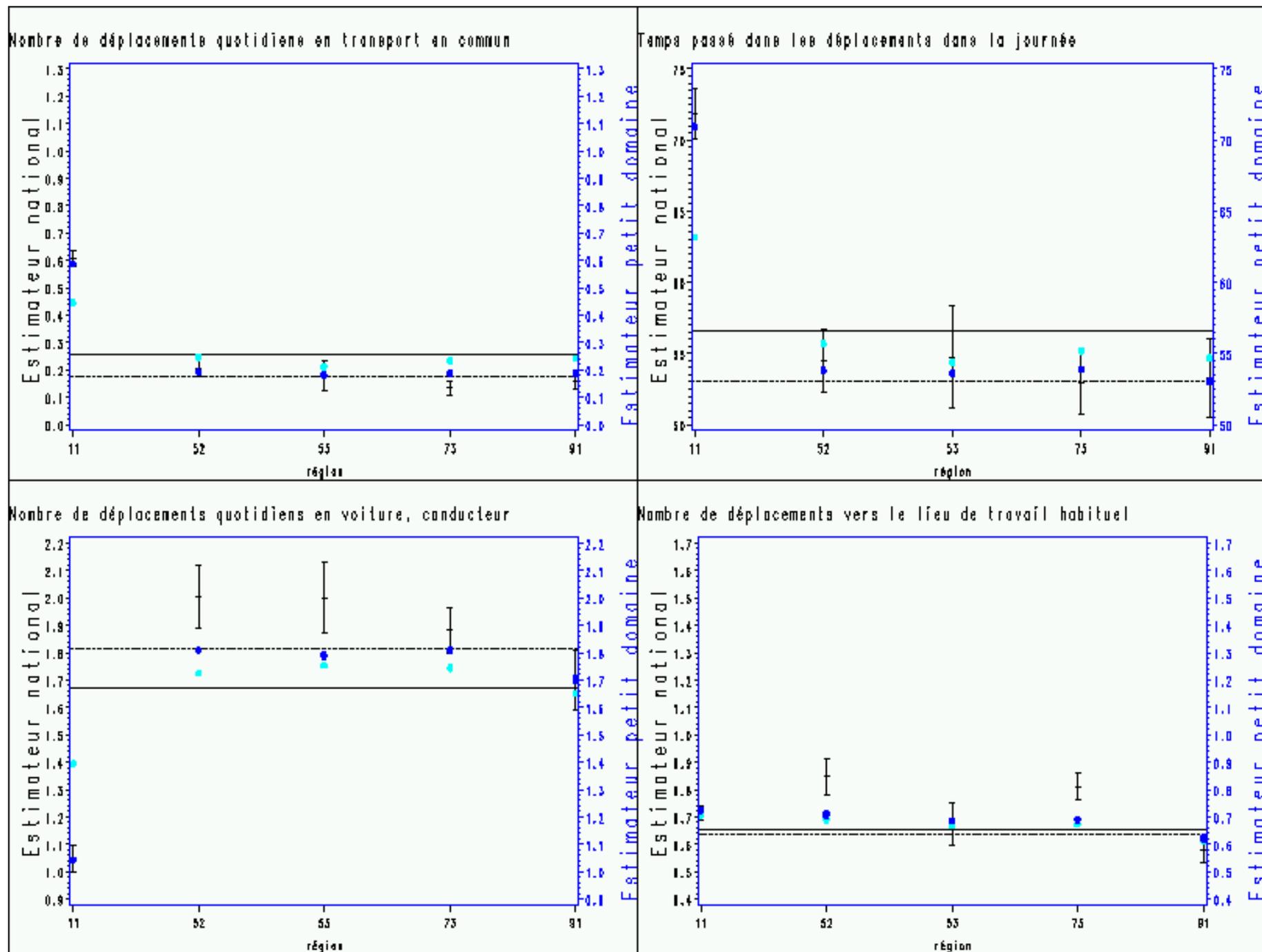
Les pastilles bleu ciel représentent l'estimateur par modèle 4 avec le modèle simplifié.

Les coefficients de régression sont estimés dans l'ensemble de l'échantillon national dans les deux cas.

Les deux lignes horizontales noires représentent la moyenne nationale, respectivement en France métropolitaine (trait plein) ou en France de province (trait pointillé), calculée avec les pondérations nationales.

Liste des codes région

11	Ile de France
52	Pays de la Loire
53	Bretagne
73	Midi-Pyrénées
91	Languedoc-Roussillon



**Figure 12 - Estimateurs par modèle mixte à effet régional aléatoire selon le modèle explicatif
Résultats dans les régions avec extension d'échantillon**

Lecture du graphique

Les valeurs représentées sont des moyennes par habitant de 6 ans ou plus, qu'il soit mobile ou non le jour de référence.

Pour chaque variable représentée dans une région, le segment symbolise l'intervalle de confiance à 95 % autour du vrai total estimé par l'estimateur direct. Le centre du segment est la moyenne régionale estimée avec les pondérations nationales, identique à celles d'un calage régional dans les régions avec extension (modèle 1 dans la section 4.2.1).

Le modèle simplifié est celui décrit en section 6.1, le modèle détaillé est décrit en section 6.2.

Les pastilles vertes représentent l'estimateur par modèle mixte (modèle 9 de la section 4.2.3) avec le modèle détaillé.

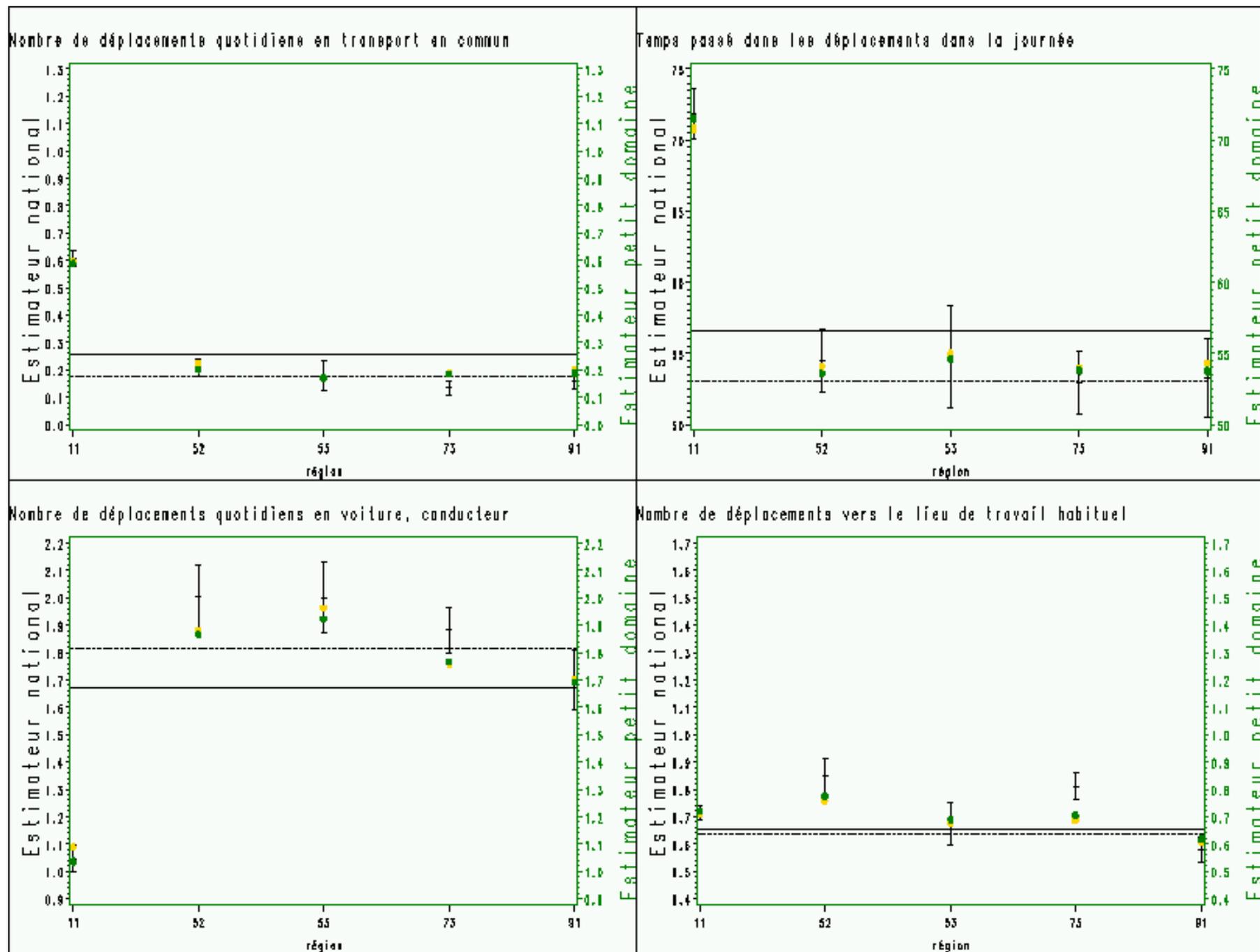
Les pastilles jaunes représentent l'estimateur du modèle 9 avec le modèle simplifié.

Les coefficients de régression sont estimés dans l'ensemble de l'échantillon national dans les deux cas.

Les deux lignes horizontales noires représentent la moyenne nationale, respectivement en France métropolitaine (trait plein) ou en France de province (trait pointillé), calculée avec les pondérations nationales.

Liste des codes région

11	Ile de France
52	Pays de la Loire
53	Bretagne
73	Midi-Pyrénées
91	Languedoc-Roussillon



**Figure 13 - Estimateur direct et estimateurs « petits domaines » avec le modèle explicatif détaillé
Résultats dans les régions avec extension d'échantillon**

Lecture du graphique

Les valeurs représentées sont des moyennes par habitant de 6 ans ou plus, qu'il soit mobile ou non le jour de référence.

Pour chaque variable représentée dans une région, le segment symbolise l'intervalle de confiance à 95 % autour du vrai total estimé par l'estimateur direct. Le centre du segment est la moyenne régionale estimée avec les pondérations nationales, identique à celles d'un calage régional dans les régions avec extension (modèle 1 dans la section 4.2.1).

Toutes les autres valeurs représentées ont été estimées avec le modèle détaillé décrit dans la section 6.2.

Les pastilles bleu clair représentent l'estimateur synthétique de type régression (modèle 4 de la section 4.2.2), les coefficients de régression étant estimés dans l'ensemble de l'échantillon national.

Les pastilles bleu sombres représentent l'estimateur du modèle 4, les coefficients de régression étant estimés dans l'échantillon national hors Ile de France.

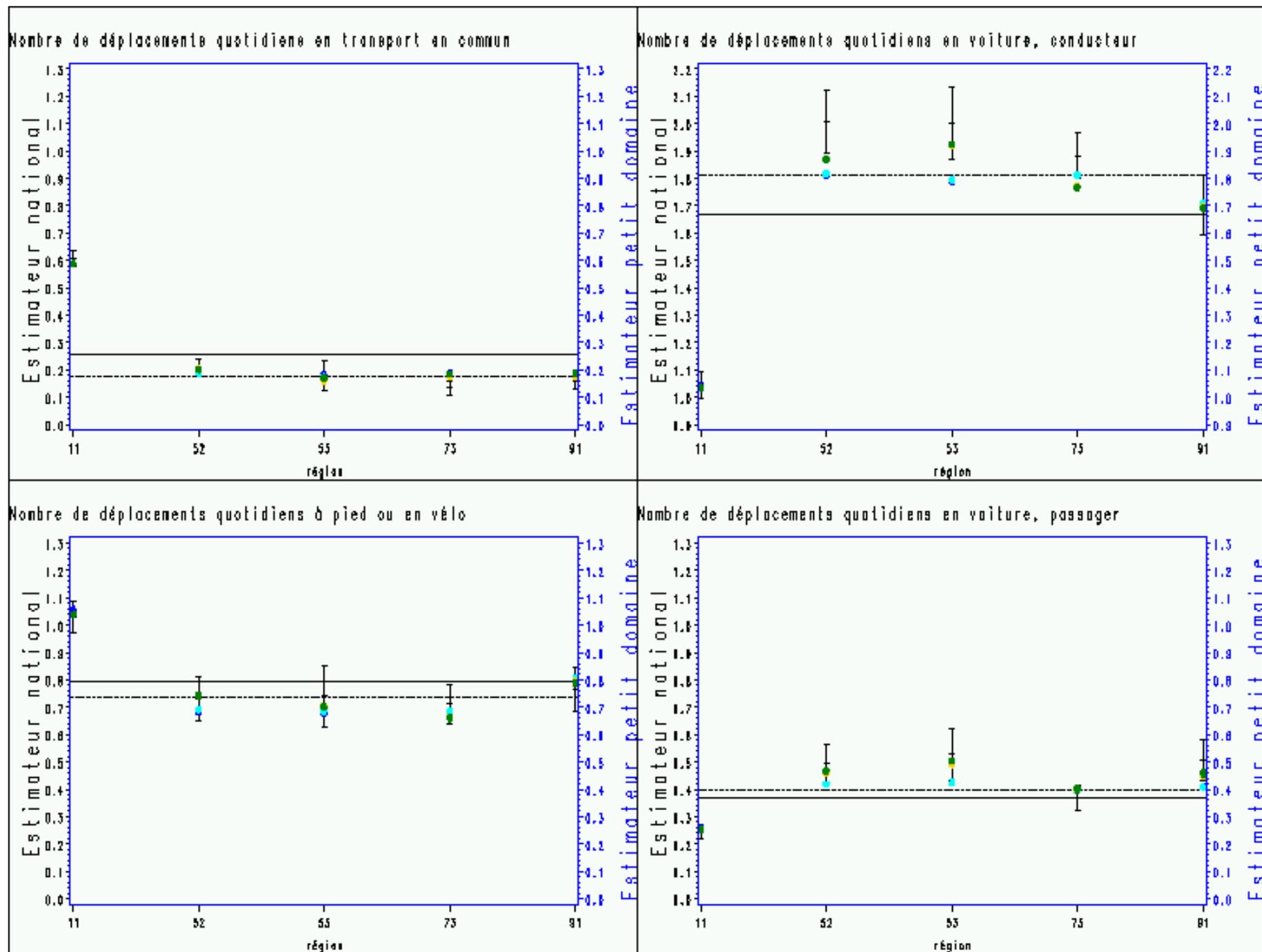
Les pastilles vertes représentent l'estimateur par modèle mixte (modèle 9 de la section 4.2.3), les coefficients de régression étant estimés dans l'ensemble de l'échantillon national.

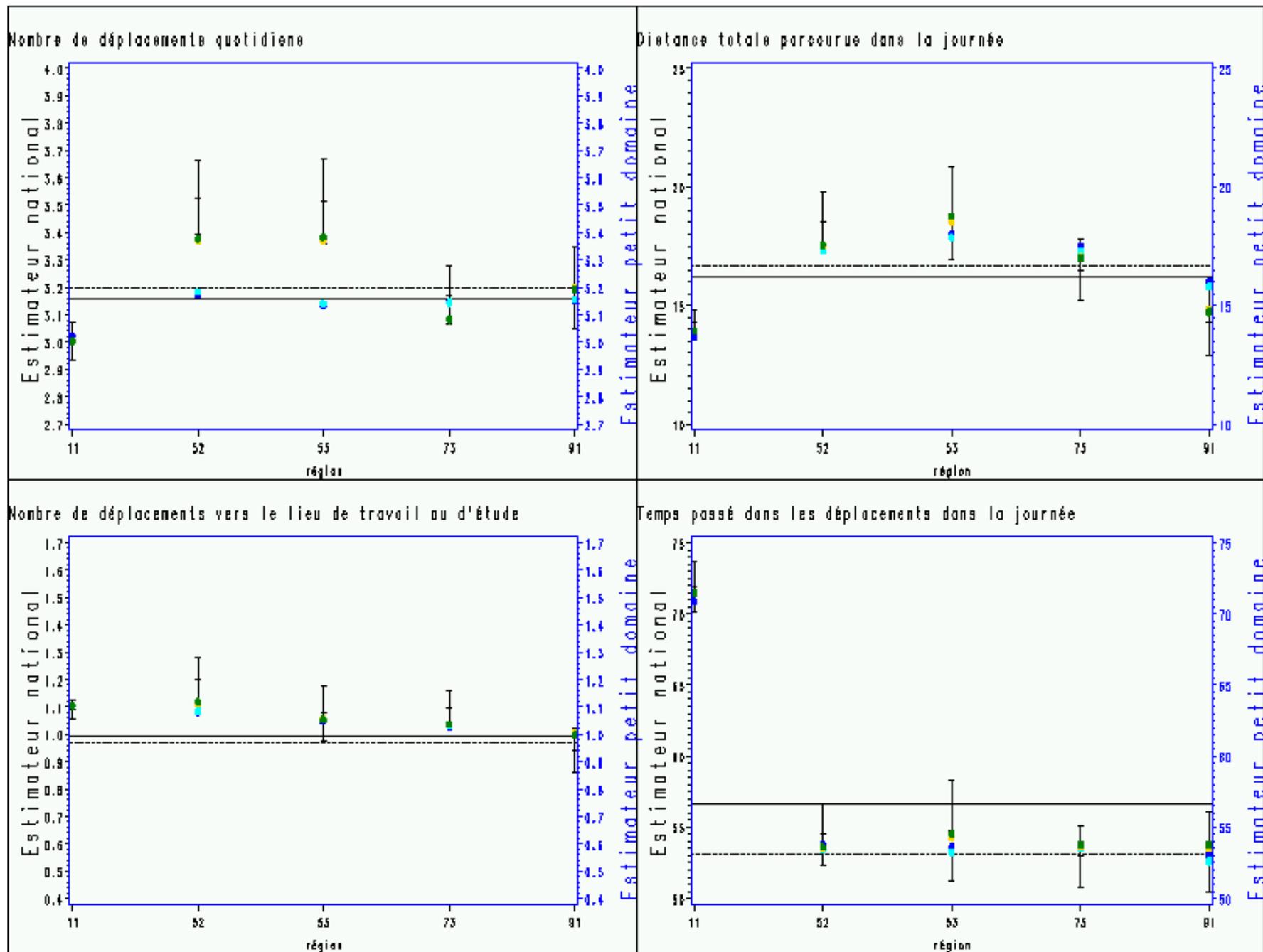
Les pastilles jaunes représentent l'estimateur du modèle 9, les coefficients de régression étant estimés dans l'échantillon national hors Ile de France.

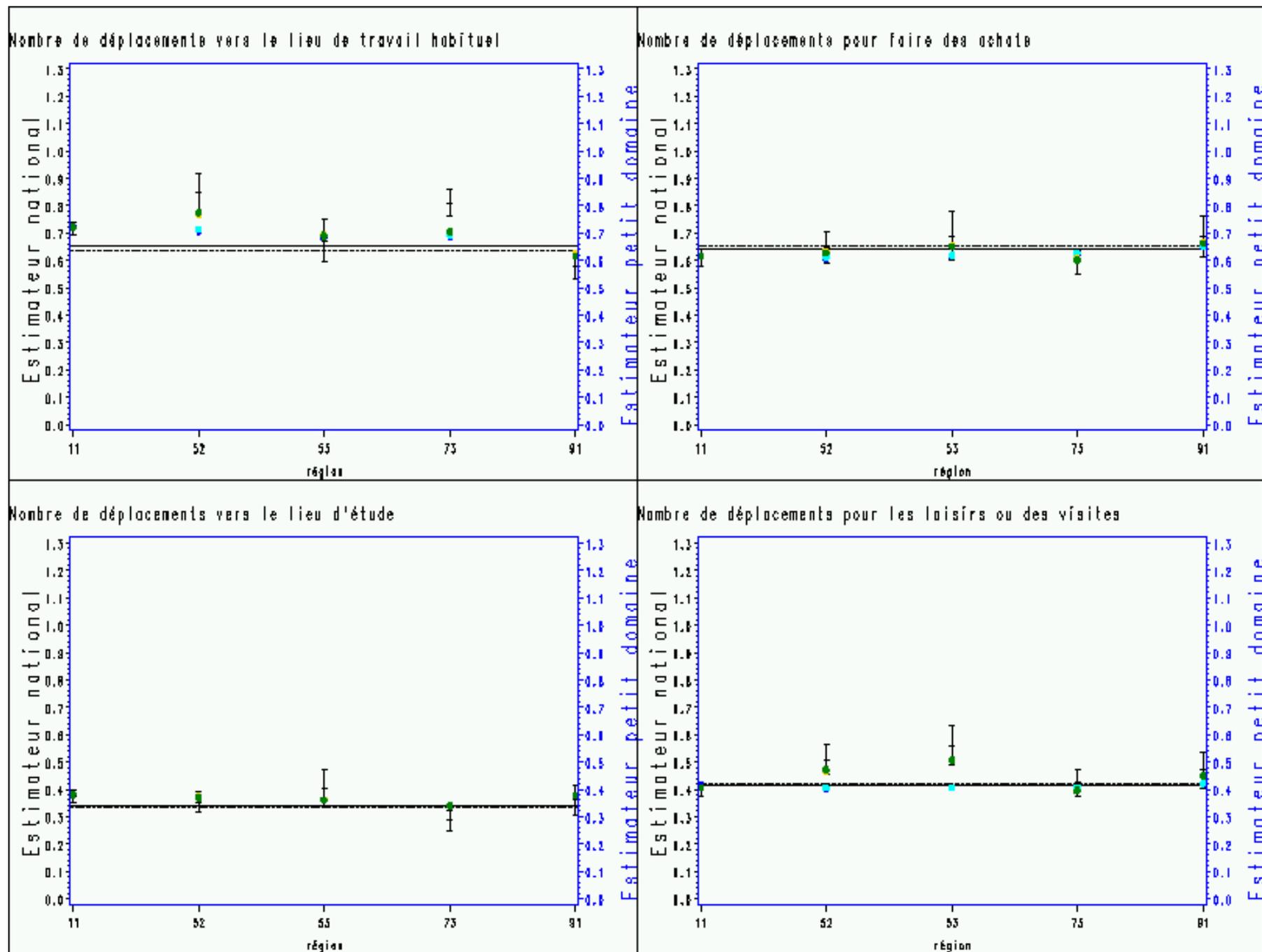
Les deux lignes horizontales noires représentent la moyenne nationale, respectivement en France métropolitaine (trait plein) ou en France de province (trait pointillé), calculée avec les pondérations nationales.

Liste des codes région

11	Ile de France
52	Pays de la Loire
53	Bretagne
73	Midi-Pyrénées
91	Languedoc-Roussillon







**Figure 14 - Estimateurs directs et estimateurs « petits domaines » avec le modèle explicatif détaillé
Résultats dans les régions sans extension d'échantillon**

Lecture du graphique

Les valeurs représentées sont des moyennes par habitant de 6 ans ou plus, qu'il soit mobile ou non le jour de référence.

Pour chaque variable représentée dans une région, le segment symbolise l'intervalle de confiance à 95 % autour du vrai total estimé par l'estimateur direct. Le centre du segment est la moyenne régionale estimée avec les pondérations nationales.

Les pastilles rouges représentent l'estimateur direct par calage sur des populations régionales (modèle 1 dans la section 4.2.1), avec les modalités simplifiées décrites en section 6.1.

Toutes les autres valeurs représentées ont été estimées avec le modèle détaillé décrit dans la section 6.2.

Les pastilles bleu clair représentent l'estimateur synthétique de type régression (modèle 4 de la section 4.2.2), les coefficients de régression étant estimés dans l'ensemble de l'échantillon national.

Les pastilles bleu sombre représentent l'estimateur du modèle 4, les coefficients de régression étant estimés dans l'échantillon national hors Ile de France.

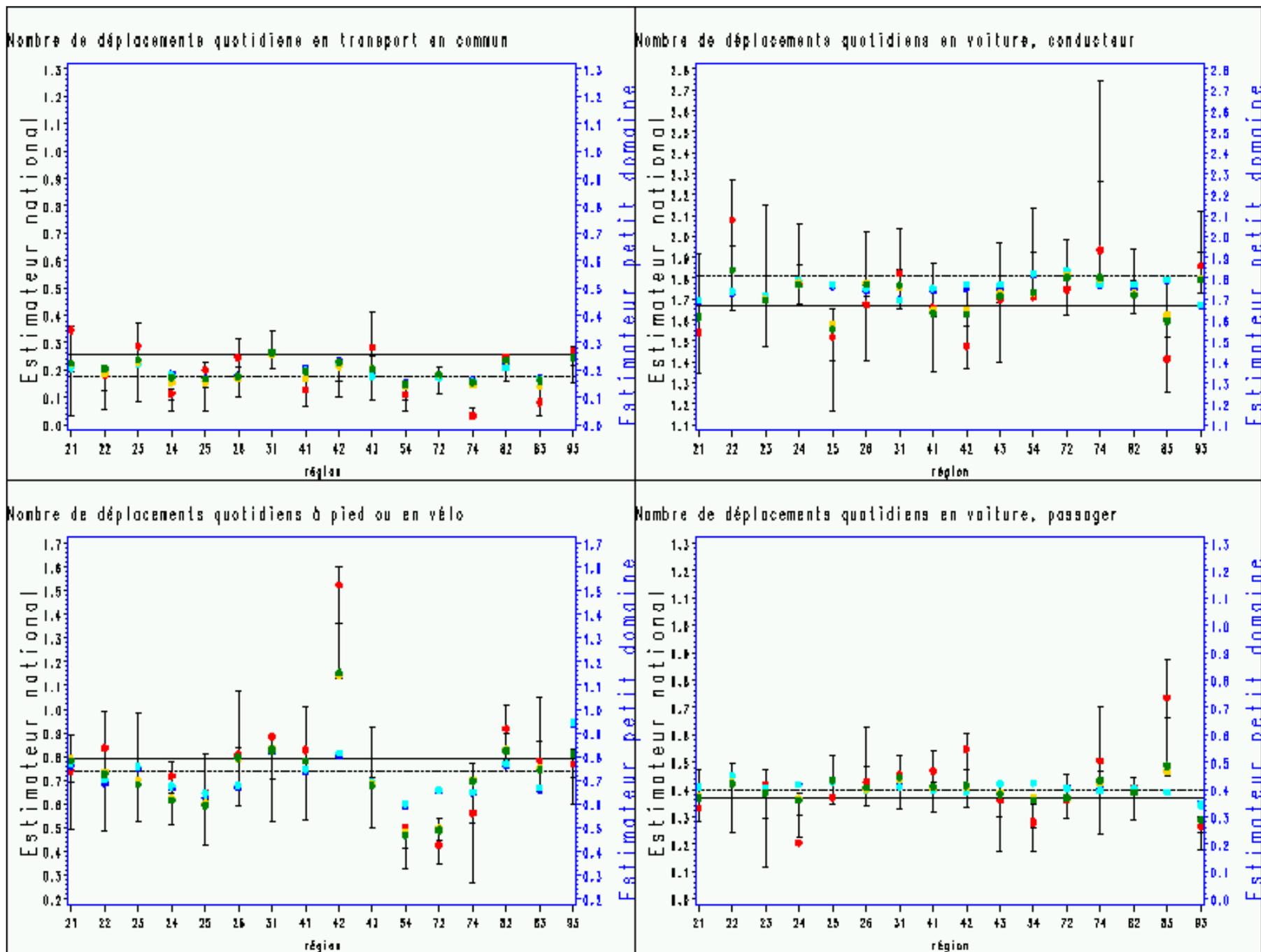
Les pastilles vertes représentent l'estimateur par modèle mixte (modèle 9 de la section 4.2.3), les coefficients de régression étant estimés dans l'ensemble de l'échantillon national.

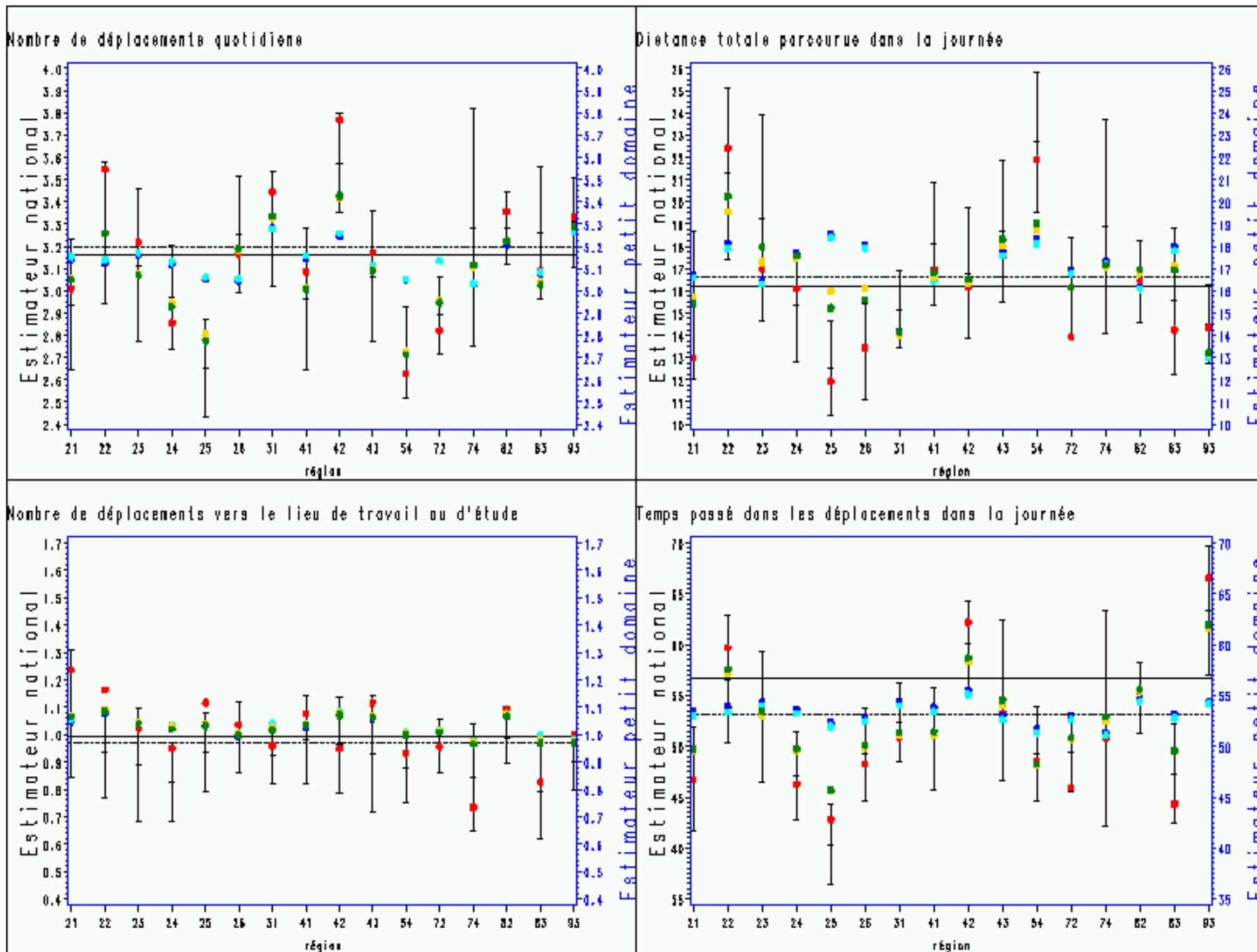
Les pastilles jaunes représentent l'estimateur du modèle 9, les coefficients de régression étant estimés dans l'échantillon national hors Ile de France.

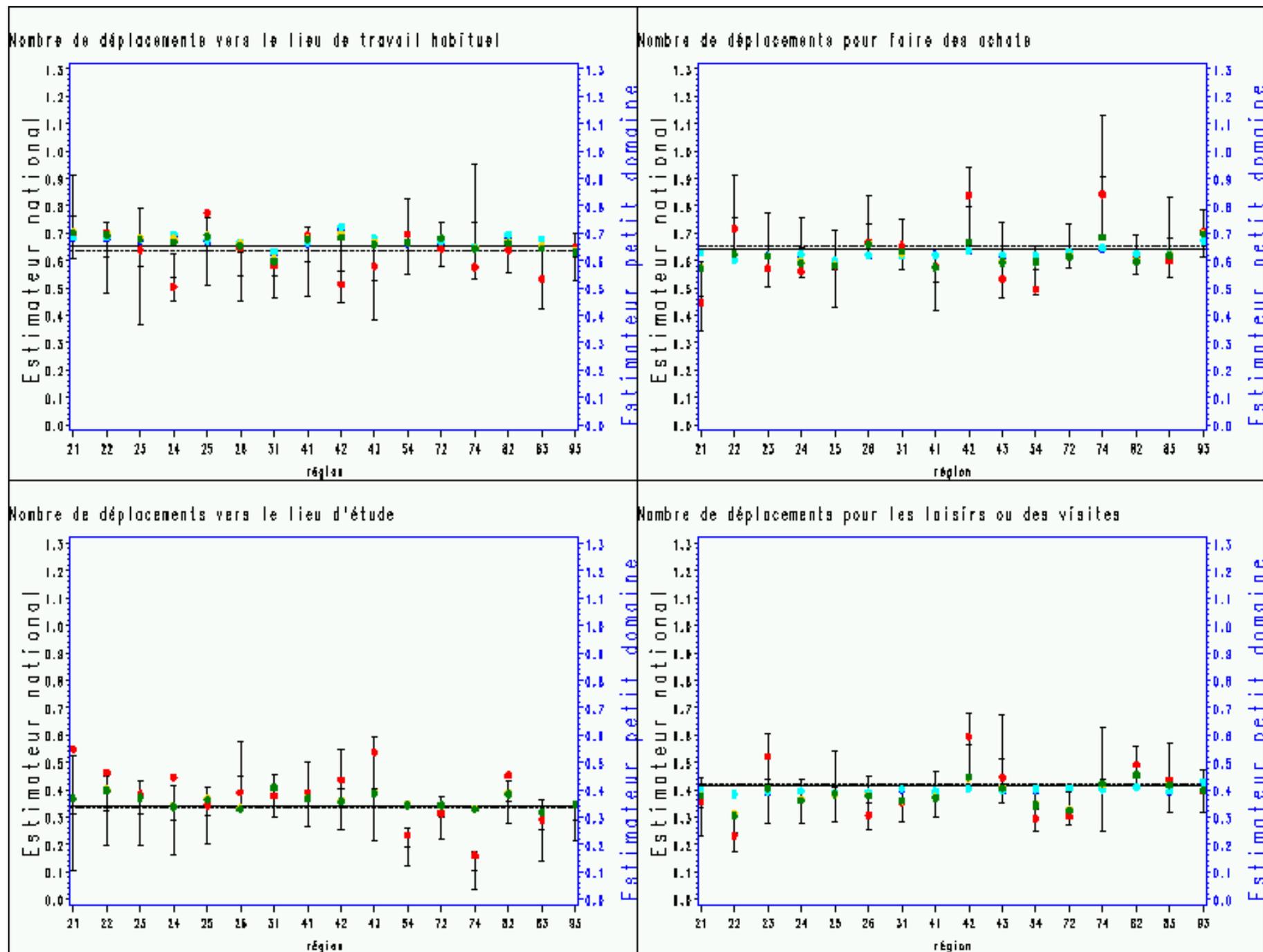
Les deux lignes horizontales noires représentent la moyenne nationale, respectivement en France métropolitaine (trait plein) ou en France de province (trait pointillé), calculée avec les pondérations nationales.

Liste des codes région

21	Champagne-Ardenne	42	Alsace
22	Picardie	43	Franche-Comté
23	Haute-Normandie	54	Poitou-Charentes
24	Centre	72	Aquitaine
25	Basse-Normandie	74	Limousin
26	Bourgogne	82	Rhône-Alpes
31	Nord	83	Auvergne
41	Lorraine	93	Provence Côte d Azur







Bibliographie

- [1] P. Ardilly, « Panorama des principales méthodes d'estimation sur petits domaines », INSEE-UMS, 2006.
- [2] K. Attal-Toubert et O. Sautory, « Estimations régionales à l'aide de techniques d'analyse multidimensionnelle », INSEE 1998, document de travail n°9807.
- [3] C. Couet, « Estimations locales dans le cadre de l'enquête HID », INSEE 2002, document de travail n°F0207.
- [4] M. Fesseau, « Les personnes en difficulté face à l'écrit ». Estimations régionales sur petits domaines dans l'enquête IVQ 2004, INSEE 2006, intranet du pôle ISM.
- [5] D. Ravaille, « Mise en œuvre d'estimations sur « petits domaines » pour l'estimation de taux de chômage et de taux d'activité par région, à partir de l'enquête Emploi », INSEE 2005, rapport de stage ENSAE.
- [6] Techniques d'enquête, volume 29, n°1, juin 2003 : numéro spécial de la revue de Statistique Canada consacré à l'estimation sur petits domaines.
- [7] M.A. Hidirolou et Z. Patak, « Estimation par domaine par la régression linéaire », Techniques d'enquête, volume 30 n°1, juin 2004, pp.73-85.
- [8] C. Lavergne, C. Trottier, « sur l'estimation dans les modèles linéaires généralisés à effets aléatoires », Revue de statistique appliquée, 2000, XLVIII, pp.49-67.

Annexe 1 : effectif de l'échantillon par région

nombre de ménages répondant aux deux visites

Région de résidence	Fréquence	Pourcentage
Ile de France	5454	29.27
Champagne-Ardenne	320	1.72
Picardie	410	2.20
Haute-Normandie	228	1.22
Centre	449	2.41
Basse-Normandie	347	1.86
Bourgogne	387	2.08
Nord	608	3.26
Lorraine	442	2.37
Alsace	279	1.50
Franche-Comté	289	1.55
Pays de la Loire	2676	14.36
Bretagne	1139	6.11
Poitou-Charentes	430	2.31
Aquitaine	708	3.80
Midi-Pyrénées	985	5.29
Limousin	206	1.11
Rhône-Alpes	892	4.79
Auvergne	284	1.52
Languedoc-Roussillon	1294	6.95
Provence Côte d Azur	758	4.07
Corse	47	0.25
<hr/>		
Total métropole	18632	

Annexe 2 : population estimée au 1er janvier 2008 par région

code région	région	population des ménages	taille moyenne des ménages	nombre de ménages	population des ménages âgée de 6 ans ou plus
11	Île-de-France	11421074	2.33567	4889858	10459891
21	Champagne-Ardenne	1305156	2.26222	576935	1210094
22	Picardie	1857879	2.44133	761010	1710109
23	Haute-Normandie	1779492	2.34657	758338	1641628
24	Centre	2467608	2.26604	1088950	2286456
25	Basse-Normandie	1425247	2.27174	627383	1321811
26	Bourgogne	1587744	2.19502	723340	1480543
31	Nord-Pas-de-Calais	3950636	2.46422	1603196	3626474
41	Lorraine	2278752	2.31136	985893	2117736
42	Alsace	1793207	2.36155	759336	1662912
43	Franche-Comté	1136192	2.26739	501101	1049753
52	Pays de la Loire	3419478	2.29385	1490714	3147885
53	Bretagne	3057668	2.21812	1378494	2831784
54	Poitou-Charentes	1703029	2.19290	776612	1589282
72	Aquitaine	3094516	2.20106	1405922	2891037
73	Midi-Pyrénées	2764332	2.18666	1264177	2579320
74	Limousin	715802	2.10624	339849	672761
82	Rhône-Alpes	5965271	2.31269	2579365	5499889
83	Auvergne	1299612	2.15928	601873	1215888
91	Languedoc-Roussillon	2530826	2.21140	1144444	2358746
93	Provence-Alpes-Côte d'Azur	4779058	2.23709	2136288	4454982
94	Corse	296503	2.33500	126982	279019

Annexe 3 : modèle simplifié du calage régional, nomenclature des variables de calage

Variables de calage de niveau ménage :

- zone de résidence
 - 1=communes rurales en espace rural,
 - 2=communes rurales en espace urbain,
 - 3=communes péri-urbaines d'unités urbaines de moins de 100000 habitants,
 - 4=villes-centres d'agglomérations de moins de 100000 habitants,
 - 5=communes péri-urbaines d'agglomérations de 100000 habitants ou plus,
 - 6=villes-centres d'unités urbaines de 100000 habitants ou plus
- type de ménage
 - personne seule ou ménage sans famille,
 - famille monoparentale,
 - couple sans enfants, couple avec enfants
- nombre de voitures du ménage
 - aucune,
 - 1 voiture,
 - 2 voitures ou plus
- catégorie socioprofessionnelle de la personne de référence
 - professions indépendantes incluant les agriculteurs,
 - cadres et professions intermédiaires,
 - ouvriers et employés,
 - inactifs

Variables de calage de niveau individu :

- Sexe
- Âge
 - 6-24 ans,
 - 25-34 ans, 35-49 ans,
 - 50-64 ans,
 - 65 ans et plus
- catégorie socioprofessionnelle de l'individu, les retraités étant reclassés dans leur profession d'origine
 - agriculteurs et professions indépendantes en activité ou retraités,
 - cadres et professions intermédiaires actifs,
 - ouvriers et employés actifs,
 - anciens cadres et professions intermédiaires,
 - anciens ouvriers et employés,
 - inactifs non retraités
- Nombre de voitures du ménage auquel appartient l'individu
 - aucune,
 - 1 voiture,
 - 2 voitures ou plus

Annexe 4 : modèle détaillé, nomenclature des variables explicatives

Variables explicatives de niveau ménage :

- Classe de commune
 - 01 rurale dans un espace urbain
 - 02 rurale dans un espace à dominante rurale
 - 11 agglomération de moins de 20000 habitants - ville centre
 - 12 agglomération de moins de 20000 habitants - banlieue
 - 41 agglomération de 20000 à 50000 habitants - ville centre
 - 42 agglomération de 20000 à 50000 habitants - banlieue
 - 51 agglomération de 50000 à 100000 habitants - ville centre
 - 52 agglomération de 50000 à 100000 habitants - banlieue
 - 61 agglomération de 100000 à 200000 habitants - ville centre
 - 62 agglomération de 100000 à 200000 habitants - banlieue
 - 71 agglomération de 200000 à 2 M d'habitants - ville centre
 - 72 agglomération de 200000 à 2 M d'habitants - banlieue
 - 81 Paris
 - 82 Banlieue parisienne

- type de ménage
 - personne seule,
 - famille monoparentale,
 - couple sans enfants,
 - couple avec enfants,
 - ménage sans famille

- nombre de voitures du ménage
 - aucune,
 - 1 voiture,
 - 2 voitures ou plus

- catégorie socioprofessionnelle de la personne de référence
 - agriculteurs, artisans, commerçants, chefs d'entreprise
 - cadres
 - professions intermédiaires
 - ouvriers
 - employés
 - retraités
 - autres inactifs

Variables explicatives de niveau individu :

- hommes par tranche d'âge :
 - 6-13 ans,
 - 14-18 ans
 - 19-25 ans,
 - 26-29 ans,
 - 30-39 ans,
 - 40-49 ans,
 - 50-60 ans,
 - 61-69 ans,
 - 70-79 ans
 - 80 ans et plus

- femmes selon les mêmes tranches d'âge
- catégorie socioprofessionnelle de l'individu, les retraités étant reclassés dans leur profession d'origine :
 - agriculteurs, actifs et retraités
 - artisans, commerçants, chefs d'entreprise en activité ou retraités,
 - cadres et professions intermédiaires actifs,
 - ouvriers et employés actifs,
 - anciens cadres et professions intermédiaires,
 - anciens ouvriers et employés,
 - élèves, étudiants
 - inactifs autres
- Nombre de voitures du ménage auquel appartient l'individu
 - aucune,
 - 1 voiture,
 - 2 voitures ou plus

**Annexe 5 : estimateur direct par calage sur des totaux régionaux
étendue des pondérations après calage régional**

Région	Pondérations initiales		Rapports de poids : poids de calage/poids initial		Pondérations finales de calage	
	minimum	maximum	minimum	maximum	minimum	maximum
Ile de France	223	2272	0,60	1,90	138	3981
Champagne -Ardennes	826	4218	0,31	2,70	276	9257
Picardie	434	7459	0,25	3,20	212	11463
Haute-Normandie	829	7413	0,45	3,55	380	14217
Centre	827	6574	0,35	3,55	289	13086
Basse-Normandie	816	4337	0,35	2,50	299	9024
Bourgogne	820	7421	0,40	2,20	337	8199
Nord-Pas-de-Calais	816	4567	0,50	1,97	424	7388
Lorraine	814	4529	0,46	2,20	376	9715
Alsace	840	4567	0,33	3,10	555	11078
Franche-Comté	818	6649	0,25	2,60	207	10589
Pays de la Loire	107	2842	0,44	1,40	50	3494
Bretagne	462	4422	0,36	1,69	220	6324
Poitou-Charentes	823	6720	0,24	3,99	259	9325
Aquitaine	811	7330	0,30	1,85	251	12242
Midi-Pyrénées	372	3880	0,51	2,80	198	5270
Limousin	822	4190	0,26	2,74	218	9409
Rhône-Alpes	816	7550	0,60	2,00	520	13566
Auvergne	771	4071	0,30	2,27	248	8402
Languedoc-Roussillon	305	3980	0,50	2,60	155	3948
Provence-Alpes-Côte d'Azur	826	7747	0,45	2,00	373	15495

Annexe 6 : les estimateurs régionaux initiaux, par calage, modélisés

Liste des tableaux

Proportion de la population mobile un jour de semaine (%).....	77
Population mobile un jour de semaine	78
Nombre de déplacements par personne un jour de semaine.....	79
Nombre total de déplacements un jour de semaine par région.....	80
Nombre de déplacements en transport en commun par personne un jour de semaine	81
Nombre de déplacements en transport en commun par région un jour de semaine	82
Nombre de déplacements en voiture (conducteur) par personne un jour de semaine	83
Nombre de déplacements en voiture (conducteur) par région un jour de semaine	84
Nombre de déplacements en voiture (passager) par personne un jour de semaine	85
Nombre de déplacements en voiture (passager) par région un jour de semaine	86
Nombre de déplacements à pied ou en vélo par personne un jour de semaine	87
Nombre de déplacements à pied ou en vélo par région un jour de semaine.....	88
Nombre de déplacements pour le travail habituel par personne un jour de semaine.....	89
Nombre de déplacements pour le travail habituel par région un jour de semaine.....	90
Nombre de déplacements pour les études par personne un jour de semaine.....	91
Nombre de déplacements pour les études par région un jour de semaine.....	92
Nombre de déplacements pour faire des achats par personne un jour de semaine.....	93
Nombre de déplacements pour faire des achats par région un jour de semaine.....	94
Nombre de déplacements pour les loisirs ou des visites par personne un jour de semaine.....	95
Nombre de déplacements pour les loisirs ou des visites par région un jour de semaine.....	96
Durée totale des déplacements par personne un jour de semaine (minutes)	97
Durée moyenne d'un déplacement en transport en commun (minutes).....	98
Durée moyenne d'un déplacement en voiture comme conducteur (minutes).....	99
Distance totale parcourue à vol d'oiseau par personne un jour de semaine (km)	100
Distance moyenne d'un déplacement en voiture (conducteur)	101
Proportion de personnes se rendant sur leur lieu de travail ou d'étude en transport en commun	102
Proportion de personnes se rendant sur leur lieu de travail ou d'étude en voiture (conducteur).....	103
Proportion de personnes se rendant sur leur lieu de travail ou d'étude en voiture (passager).....	104
Proportion de personnes se rendant sur leur lieu de travail ou d'étude à pied ou en vélo.....	105
Proportion des déplacements vers le lieu de travail ou d'étude en transport en commun.....	106
Proportion des déplacements vers le lieu de travail ou d'étude en voiture (conducteur)	107
Proportion des déplacements vers le lieu de travail ou d'étude en voiture (passager).....	108
Proportion des déplacements vers le lieu de travail ou à pied ou en vélo	109

Proportion de la population mobile un jour de semaine (%)

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte aléatoire	
	Ile de France	0.8863	0.8951	0.9039	0.8984	0.8854	0.8985	0.8972	0.8972
Champagne-Ardenne	0.6988	0.7534	0.8081	0.7892	0.8410	0.8089	0.8340	0.8018	
Picardie	0.7895	0.8281	0.8666	0.8399	0.8417	0.8469	0.8351	0.8412	
Haute-Normandie	0.7916	0.8330	0.8745	0.8260	0.8516	0.8464	0.8445	0.8398	
Centre	0.8049	0.8365	0.8681	0.8456	0.8400	0.8339	0.8397	0.8351	
Basse-Normandie	0.8224	0.8588	0.8952	0.8622	0.8286	0.8376	0.8191	0.8312	
Bourgogne	0.7560	0.7936	0.8312	0.7907	0.8284	0.8169	0.8220	0.8113	
Nord	0.8050	0.8313	0.8576	0.8397	0.8639	0.8393	0.8595	0.8373	
Lorraine	0.7598	0.8049	0.8500	0.7923	0.8483	0.8203	0.8402	0.8139	
Alsace	0.8899	0.9194	0.9489	0.9287	0.8604	0.8901	0.8597	0.8904	
Franche-Comté	0.7912	0.8382	0.8852	0.8321	0.8394	0.8392	0.8299	0.8372	
Pays de la Loire	0.8646	0.8831	0.9017	0.8804	0.8493	0.8742	0.8442	0.8747	
Bretagne	0.8653	0.8850	0.9047	0.8810	0.8388	0.8721	0.8371	0.8722	
Poitou-Charentes	0.7811	0.8207	0.8604	0.8242	0.8276	0.8237	0.8157	0.8136	
Aquitaine	0.7681	0.7952	0.8224	0.8056	0.8438	0.8138	0.8363	0.8112	
Midi-Pyrénées	0.8215	0.8365	0.8515	0.8263	0.8418	0.8317	0.8388	0.8285	
Limousin	0.7342	0.8051	0.8760	0.7831	0.8205	0.8100	0.8137	0.8071	
Rhône-Alpes	0.8314	0.8542	0.8770	0.8531	0.8603	0.8516	0.8542	0.8515	
Auvergne	0.7648	0.8162	0.8677	0.7834	0.8288	0.8098	0.8262	0.8050	
Languedoc-Roussillon	0.8265	0.8484	0.8704	0.8419	0.8411	0.8518	0.8354	0.8483	
Provence Côte d Azur	0.8185	0.8455	0.8726	0.8732	0.8612	0.8592	0.8582	0.8612	
France métropolitaine		0.85		0.8531	0.8589				
France hors Ile de France		0.8396		0.8427	0.8528				

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Population mobile un jour de semaine

région	Estimateur national			Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire
Ile de France	9321706	9414342	9506979	9397094	9261375	9398208	9384535	9384347
Champagne-Ardenne	696154	890356	1084558	955056	1017727	978787	1009265	970261
Picardie	1183225	1454551	1725876	1436339	1439396	1448244	1428082	1438500
Haute-Normandie	879738	1060458	1241178	1355909	1398028	1389492	1386317	1378630
Centre	1296511	1610385	1924259	1933358	1920700	1906776	1919921	1909358
Basse-Normandie	1013319	1335419	1657518	1139646	1095308	1107098	1082717	1098638
Bourgogne	1040199	1302786	1565373	1170689	1226417	1209397	1216972	1201239
Nord	3082517	3526842	3971168	3044986	3132739	3043733	3116789	3036272
Lorraine	1419349	1676228	1933107	1677836	1796395	1737156	1779349	1723688
Alsace	1176702	1558824	1940946	1544406	1430842	1480090	1429589	1480689
Franche-Comté	709529	991959	1274389	873451	881125	880965	871237	878893
Pays de la Loire	2693932	2751764	2809595	2771420	2673507	2751747	2657323	2753458
Bretagne	2453073	2508889	2564706	2494889	2375380	2469626	2370514	2469995
Poitou-Charentes	1109323	1413883	1718443	1309961	1315329	1309056	1296448	1292979
Aquitaine	2185957	2497214	2808471	2329113	2439483	2352799	2417741	2345178
Midi-Pyrénées	2097387	2135712	2174038	2131250	2171244	2145154	2163444	2137017
Limousin	348632	577874	807116	526844	552024	544936	547405	542982
Rhône-Alpes	3952510	4330840	4709170	4692036	4731580	4683738	4697800	4682887
Auvergne	797091	1026968	1256846	952499	1007769	984670	1004588	978829
Languedoc-Roussillon	1940028	1991526	2043023	1985718	1984004	2009111	1970454	2000906
Provence Côte d Azur	3069366	3486005	3902644	3890046	3836699	3827858	3823088	3836766

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements par personne un jour de semaine

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	2.93	3.00	3.07	3.01	3.24	3.07	3.03	3.00
Champagne-Ardenne	2.64	2.94	3.23	3.01	3.16	3.09	3.14	3.05	
Picardie	2.94	3.26	3.58	3.55	3.19	3.34	3.13	3.26	
Haute-Normandie	2.77	3.11	3.46	3.22	3.18	3.15	3.16	3.07	
Centre	2.73	2.97	3.20	2.86	3.14	2.96	3.12	2.93	
Basse-Normandie	2.43	2.65	2.87	2.78	3.10	2.83	3.05	2.78	
Bourgogne	2.99	3.26	3.52	3.17	3.09	3.26	3.05	3.19	
Nord	3.02	3.28	3.53	3.45	3.21	3.36	3.28	3.33	
Lorraine	2.65	2.97	3.28	3.09	3.16	3.08	3.15	3.01	
Alsace	3.35	3.57	3.80	3.77	3.23	3.46	3.24	3.43	
Franche-Comté	2.77	3.07	3.36	3.17	3.14	3.12	3.10	3.09	
Pays de la Loire	3.39	3.53	3.66	3.50	3.18	3.40	3.17	3.38	
Bretagne	3.36	3.51	3.67	3.52	3.15	3.42	3.13	3.38	
Poitou-Charentes	2.52	2.72	2.93	2.63	3.08	2.81	3.05	2.71	
Aquitaine	2.72	2.89	3.06	2.82	3.14	2.99	3.14	2.95	
Midi-Pyrénées	3.06	3.17	3.28	3.09	3.15	3.12	3.15	3.08	
Limousin	2.75	3.28	3.82	3.04	3.06	3.16	3.03	3.12	
Rhône-Alpes	3.12	3.28	3.45	3.36	3.21	3.26	3.21	3.23	
Auvergne	2.97	3.26	3.56	3.09	3.09	3.08	3.08	3.03	
Languedoc-Roussillon	3.05	3.20	3.35	3.23	3.17	3.24	3.15	3.20	
Provence Côte d Azur	3.11	3.31	3.51	3.33	3.21	3.28	3.27	3.29	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre total de déplacements un jour de semaine par région

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	30822504	31553823	32285143	31457947	33922863	32129340	31641601	31421858
Champagne-Ardenne	2725854	3471728	4217602	3643561	3820268	3737959	3801688	3693421	
Picardie	4546048	5726437	6906826	6063691	5451500	5712759	5346436	5572230	
Haute-Normandie	3125166	3963504	4801843	5282866	5217816	5177878	5186969	5040024	
Centre	4481843	5715851	6949859	6533005	7170701	6760993	7136767	6693677	
Basse-Normandie	3063275	4127142	5191010	3673729	4095464	3747278	4033313	3670450	
Bourgogne	4137058	5343773	6550488	4687596	4580775	4819206	4508475	4722549	
Nord	11735455	13900932	16066409	12501428	11637746	12180494	11908522	12092947	
Lorraine	5021578	6174923	7328267	6535872	6694019	6521112	6668234	6371192	
Alsace	4451758	6057206	7662654	6269191	5370500	5752607	5394723	5700452	
Franche-Comté	2680049	3628297	4576544	3329247	3296816	3273972	3255351	3243767	
Pays de la Loire	10568244	10990403	11412562	11003373	9998227	10718064	9993696	10627714	
Bretagne	9523126	9959371	10395616	9967844	8913253	9696857	8873285	9579840	
Poitou-Charentes	3688485	4691676	5694867	4180055	4894612	4467796	4845731	4314789	
Aquitaine	7874817	9072834	10270851	8164785	9064980	8631497	9064708	8519176	
Midi-Pyrénées	7824339	8095132	8365924	7967021	8133594	8035007	8113406	7948651	
Limousin	1383711	2356447	3329184	2042700	2055550	2125143	2040076	2097464	
Rhône-Alpes	15027433	16647647	18267862	18467040	17652981	17916930	17639428	17744856	
Auvergne	3175148	4105220	5035291	3760285	3760048	3741889	3743127	3678723	
Languedoc-Roussillon	7152636	7504303	7855969	7616761	7475645	7649466	7428198	7537655	
Provence Côte d Azur	11844281	13632343	15420405	14840229	14286510	14602906	14562672	14660354	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements en transport en commun par personne un jour de semaine

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	0.58	0.61	0.64	0.61	0.45	0.60	0.59	0.59
Champagne-Ardenne	0.03	0.20	0.36	0.35	0.26	0.28	0.22	0.22	
Picardie	0.06	0.12	0.19	0.18	0.24	0.22	0.20	0.20	
Haute-Normandie	0.08	0.23	0.37	0.29	0.28	0.27	0.23	0.24	
Centre	0.05	0.09	0.13	0.12	0.22	0.16	0.19	0.17	
Basse-Normandie	0.05	0.14	0.23	0.20	0.21	0.18	0.17	0.17	
Bourgogne	0.10	0.21	0.31	0.25	0.21	0.21	0.18	0.18	
Nord	0.20	0.27	0.34	0.27	0.35	0.27	0.27	0.26	
Lorraine	0.07	0.14	0.21	0.13	0.26	0.18	0.21	0.20	
Alsace	0.10	0.16	0.22	0.22	0.28	0.22	0.23	0.23	
Franche-Comté	0.09	0.25	0.41	0.28	0.23	0.28	0.19	0.20	
Pays de la Loire	0.18	0.21	0.24	0.22	0.25	0.22	0.19	0.20	
Bretagne	0.12	0.18	0.23	0.17	0.21	0.17	0.18	0.17	
Poitou-Charentes	0.05	0.09	0.13	0.11	0.20	0.16	0.15	0.15	
Aquitaine	0.11	0.16	0.21	0.17	0.24	0.21	0.18	0.18	
Midi-Pyrénées	0.11	0.13	0.16	0.14	0.23	0.19	0.19	0.19	
Limousin	-.00	0.03	0.06	0.03	0.21	0.15	0.16	0.16	
Rhône-Alpes	0.16	0.21	0.26	0.25	0.28	0.27	0.22	0.24	
Auvergne	0.03	0.08	0.14	0.08	0.20	0.12	0.17	0.16	
Languedoc-Roussillon	0.13	0.16	0.19	0.15	0.24	0.20	0.19	0.19	
Provence Côte d Azur	0.15	0.22	0.28	0.27	0.32	0.27	0.25	0.25	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements en transport en commun par région un jour de semaine

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	6059533	6388592	6717650	6420934	4657716	6267285	6161074	6173131
Champagne-Ardenne	31247	230962	430677	418766	313797	341735	263932	271174	
Picardie	98554	218775	338996	311923	404594	372401	350350	347274	
Haute-Normandie	94542	291452	488363	474983	462801	449600	382379	387779	
Centre	93928	178090	262253	263418	509832	363490	428288	397269	
Basse-Normandie	69139	214010	358882	263505	275355	242388	226233	223404	
Bourgogne	153565	343915	534265	366100	304949	311020	264591	264131	
Nord	820615	1162031	1503447	961410	1255619	981906	964523	955271	
Lorraine	136564	285605	434646	268129	553924	386572	441155	414509	
Alsace	151099	271079	391060	359134	464677	367384	388405	381841	
Franche-Comté	95292	297546	499799	298738	244740	293719	199238	211832	
Pays de la Loire	555880	646744	737608	690709	781871	708027	612849	631473	
Bretagne	350920	502463	654006	477934	602244	484888	514094	491759	
Poitou-Charentes	86274	153651	221028	178355	325416	249435	244926	236602	
Aquitaine	338316	509926	681536	503186	687101	610083	515561	529982	
Midi-Pyrénées	280344	339685	399027	365016	600364	494330	487971	479963	
Limousin	-2071	21208	44488	21768	141705	99648	109770	105462	
Rhône-Alpes	785867	1053653	1321439	1376115	1551113	1458918	1228444	1293890	
Auvergne	35025	106533	178041	100140	248163	149914	211723	196874	
Languedoc-Roussillon	302555	370630	438704	356762	575828	472498	451213	444948	
Provence Côte d Azur	604293	892931	1181568	1207098	1444421	1203875	1101619	1110440	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements en voiture (conducteur) par personne un jour de semaine

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	1.00	1.05	1.09	1.04	1.40	1.09	1.04	1.04
Champagne-Ardenne	1.35	1.63	1.91	1.54	1.64	1.60	1.69	1.61	
Picardie	1.64	1.96	2.27	2.08	1.70	1.87	1.73	1.84	
Haute-Normandie	1.47	1.81	2.15	1.70	1.63	1.68	1.71	1.70	
Centre	1.68	1.87	2.06	1.78	1.74	1.74	1.78	1.77	
Basse-Normandie	1.16	1.41	1.65	1.52	1.73	1.51	1.76	1.56	
Bourgogne	1.41	1.72	2.02	1.68	1.73	1.78	1.74	1.77	
Nord	1.65	1.84	2.04	1.83	1.48	1.75	1.70	1.77	
Lorraine	1.36	1.61	1.87	1.66	1.66	1.63	1.74	1.63	
Alsace	1.37	1.57	1.77	1.48	1.70	1.58	1.75	1.63	
Franche-Comté	1.39	1.68	1.97	1.70	1.72	1.70	1.75	1.71	
Pays de la Loire	1.89	2.01	2.12	1.94	1.72	1.88	1.81	1.87	
Bretagne	1.87	2.00	2.13	2.01	1.75	1.96	1.79	1.92	
Poitou-Charentes	1.72	1.93	2.13	1.71	1.75	1.71	1.82	1.73	
Aquitaine	1.62	1.80	1.99	1.75	1.73	1.77	1.83	1.81	
Midi-Pyrénées	1.80	1.88	1.97	1.77	1.74	1.76	1.81	1.77	
Limousin	1.79	2.27	2.74	1.94	1.72	1.80	1.77	1.81	
Rhône-Alpes	1.63	1.79	1.94	1.76	1.68	1.71	1.75	1.72	
Auvergne	1.26	1.52	1.78	1.41	1.75	1.54	1.79	1.60	
Languedoc-Roussillon	1.59	1.70	1.81	1.68	1.65	1.70	1.70	1.69	
Provence Côte d Azur	1.73	1.92	2.12	1.86	1.55	1.80	1.67	1.79	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements en voiture (conducteur) par région un jour de semaine

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	10488839	10997778	11506717	10826223	14594887	11382691	10920081	10837188
Champagne-Ardenne	1449956	1927585	2405213	1865841	1981545	1930275	2040848	1952985	
Picardie	2638991	3437714	4236437	3559644	2904995	3205063	2953596	3145698	
Haute-Normandie	1759027	2304773	2850520	2795808	2668178	2763525	2802087	2784504	
Centre	2830098	3595191	4360284	4066766	3971914	3972345	4078732	4041109	
Basse-Normandie	1618348	2189221	2760094	2010202	2280930	1991393	2326576	2057020	
Bourgogne	2133141	2816111	3499080	2481707	2556319	2640280	2577029	2619177	
Nord	6472522	7818209	9163896	6622262	5369738	6363218	6164181	6404935	
Lorraine	2648156	3358661	4069165	3510920	3515362	3457786	3690749	3459803	
Alsace	1923086	2659670	3396254	2458198	2826627	2629329	2912897	2710013	
Franche-Comté	1361803	1992205	2622607	1786101	1806577	1787333	1841085	1798805	
Pays de la Loire	5887487	6248357	6609226	6092106	5429795	5916845	5698284	5881665	
Bretagne	5301685	5671960	6042235	5686467	4965714	5560842	5066735	5447435	
Poitou-Charentes	2567980	3320674	4073368	2713664	2779230	2721027	2895118	2749792	
Aquitaine	4778098	5663161	6548225	5054768	5001148	5113646	5293825	5222426	
Midi-Pyrénées	4591488	4805705	5019921	4560013	4499638	4538359	4666898	4558486	
Limousin	958059	1626092	2294125	1302098	1154390	1213131	1189110	1214915	
Rhône-Alpes	7969698	9058506	10147314	9702007	9264393	9388504	9647079	9467261	
Auvergne	1415549	1911624	2407698	1717587	2125522	1873902	2176478	1941034	
Languedoc-Roussillon	3733631	3989932	4246232	3961644	3891241	4011186	4019638	3989639	
Provence Côte d Azur	6658038	7925383	9192728	8280109	6885343	8032072	7422291	7996232	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements en voiture (passager) par personne un jour de semaine

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	0.22	0.24	0.26	0.25	0.33	0.28	0.26	0.26
Champagne-Ardenne	0.28	0.38	0.47	0.33	0.40	0.36	0.41	0.37	
Picardie	0.24	0.37	0.50	0.42	0.43	0.41	0.45	0.42	
Haute-Normandie	0.12	0.29	0.47	0.42	0.41	0.39	0.41	0.39	
Centre	0.23	0.31	0.39	0.21	0.41	0.35	0.42	0.36	
Basse-Normandie	0.34	0.43	0.52	0.37	0.42	0.42	0.43	0.43	
Bourgogne	0.34	0.48	0.63	0.43	0.41	0.43	0.40	0.41	
Nord	0.33	0.42	0.52	0.46	0.39	0.44	0.41	0.44	
Lorraine	0.32	0.43	0.54	0.47	0.40	0.41	0.40	0.41	
Alsace	0.33	0.47	0.60	0.55	0.40	0.42	0.39	0.41	
Franche-Comté	0.18	0.30	0.42	0.36	0.41	0.37	0.42	0.38	
Pays de la Loire	0.42	0.49	0.57	0.51	0.41	0.47	0.42	0.47	
Bretagne	0.43	0.53	0.62	0.53	0.42	0.50	0.42	0.50	
Poitou-Charentes	0.17	0.26	0.35	0.28	0.40	0.35	0.42	0.36	
Aquitaine	0.29	0.37	0.45	0.36	0.39	0.36	0.41	0.37	
Midi-Pyrénées	0.32	0.37	0.42	0.44	0.40	0.40	0.40	0.40	
Limousin	0.24	0.47	0.70	0.51	0.37	0.42	0.40	0.43	
Rhône-Alpes	0.29	0.37	0.45	0.41	0.40	0.40	0.40	0.39	
Auvergne	0.45	0.66	0.87	0.74	0.40	0.50	0.39	0.49	
Languedoc-Roussillon	0.43	0.51	0.58	0.55	0.40	0.46	0.41	0.46	
Provence Côte d Azur	0.18	0.24	0.30	0.27	0.36	0.30	0.35	0.29	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements en voiture (passager) par région un jour de semaine

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	2313287	2535119	2756951	2611365	3403958	2948685	2719433	2669317
Champagne-Ardenne	296720	444656	592591	399554	486031	440648	496576	447359	
Picardie	404235	651280	898326	721261	737018	708678	769876	719624	
Haute-Normandie	129704	372623	615542	687035	668161	644773	665913	632735	
Centre	398350	588566	778782	469674	930013	794431	954860	824247	
Basse-Normandie	460372	674347	888321	489440	551032	561533	565948	572140	
Bourgogne	480903	792845	1104786	633271	600209	631153	590375	600174	
Nord	1293755	1801869	2309983	1650867	1409137	1592687	1500144	1605688	
Lorraine	597519	892028	1186537	990769	851760	867885	849925	869501	
Alsace	448900	795493	1142085	909000	666364	705651	654558	687190	
Franche-Comté	199895	354248	508601	376831	431554	393629	440918	401736	
Pays de la Loire	1317609	1540847	1764085	1604648	1284406	1482346	1322182	1471578	
Bretagne	1222700	1496083	1769466	1499572	1176125	1404903	1200725	1423934	
Poitou-Charentes	257982	449351	640721	445341	633152	559480	670890	570289	
Aquitaine	879578	1168538	1457498	1044882	1118586	1053623	1175464	1066297	
Midi-Pyrénées	829434	947422	1065411	1133809	1019891	1040265	1022965	1034722	
Limousin	110079	336157	562236	340048	252078	284496	266837	291104	
Rhône-Alpes	1411683	1854057	2296431	2239135	2196140	2210766	2211614	2133966	
Auvergne	463061	830297	1197532	894733	480276	613012	475333	591605	
Languedoc-Roussillon	1007367	1185417	1363467	1303405	943195	1073373	967399	1081601	
Provence Côte d Azur	689319	990951	1292582	1183972	1586552	1358697	1550456	1288896	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements à pied ou en vélo par personne un jour de semaine

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	0.97	1.03	1.09	1.03	1.00	1.03	1.06	1.04
Champagne-Ardenne	0.49	0.69	0.89	0.74	0.79	0.80	0.76	0.78	
Picardie	0.49	0.74	0.99	0.84	0.76	0.79	0.69	0.73	
Haute-Normandie	0.53	0.76	0.98	0.76	0.80	0.76	0.75	0.68	
Centre	0.51	0.65	0.78	0.72	0.71	0.66	0.67	0.62	
Basse-Normandie	0.42	0.62	0.81	0.65	0.68	0.64	0.63	0.59	
Bourgogne	0.59	0.83	1.07	0.81	0.69	0.80	0.67	0.80	
Nord	0.52	0.71	0.89	0.89	0.93	0.87	0.82	0.83	
Lorraine	0.53	0.77	1.01	0.83	0.78	0.81	0.73	0.78	
Alsace	1.13	1.36	1.60	1.52	0.79	1.19	0.80	1.15	
Franche-Comté	0.50	0.71	0.93	0.69	0.71	0.69	0.68	0.68	
Pays de la Loire	0.65	0.73	0.81	0.75	0.73	0.75	0.68	0.74	
Bretagne	0.63	0.74	0.85	0.74	0.70	0.72	0.68	0.70	
Poitou-Charentes	0.33	0.41	0.50	0.50	0.66	0.54	0.59	0.47	
Aquitaine	0.35	0.45	0.54	0.43	0.72	0.56	0.66	0.49	
Midi-Pyrénées	0.64	0.71	0.79	0.65	0.72	0.68	0.68	0.66	
Limousin	0.27	0.52	0.77	0.56	0.69	0.74	0.65	0.70	
Rhône-Alpes	0.78	0.90	1.02	0.92	0.78	0.85	0.76	0.82	
Auvergne	0.68	0.87	1.05	0.78	0.68	0.76	0.66	0.74	
Languedoc-Roussillon	0.69	0.77	0.85	0.79	0.82	0.81	0.79	0.79	
Provence Côte d Azur	0.60	0.71	0.83	0.77	0.92	0.78	0.94	0.81	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements à pied ou en vélo par région un jour de semaine

région	Estimateur national			Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire
	Ile de France	10228754	10832119	11435483	10825586	10468324	10741894	11048846
Champagne-Ardenne	530579	816718	1102858	896578	958819	970585	923262	946944
Picardie	760222	1294884	1829546	1432091	1299713	1342687	1173904	1241168
Haute-Normandie	609944	962498	1315052	1244163	1312703	1255388	1228283	1122484
Centre	841137	1241984	1642831	1642405	1617024	1502968	1525274	1407340
Basse-Normandie	514694	959324	1403953	859303	902996	848599	833024	785334
Bourgogne	839155	1370134	1901112	1197484	1028564	1186727	988832	1178241
Nord	2156783	3001800	3846816	3211650	3363289	3150550	2985324	3001895
Lorraine	1049233	1605818	2162403	1753909	1645243	1723329	1553482	1651532
Alsace	1610984	2309692	3008400	2533325	1308965	1979593	1336265	1916923
Franche-Comté	504660	841981	1179303	727770	746969	723885	715546	710879
Pays de la Loire	2026452	2279746	2533041	2372922	2286171	2371080	2144119	2339366
Bretagne	1783372	2099401	2415430	2094426	1994961	2049343	1915911	1988580
Poitou-Charentes	506387	713840	921294	800864	1056587	861976	942360	748551
Aquitaine	1061812	1397658	1733504	1230741	2078566	1611143	1893632	1414773
Midi-Pyrénées	1630416	1817576	2004736	1680141	1848175	1762856	1757891	1702634
Limousin	136580	372990	609400	378785	464182	495172	434405	468644
Rhône-Alpes	3861702	4561692	5261682	5049460	4283399	4661421	4194061	4535008
Auvergne	803581	1090155	1376729	952744	831253	927894	802976	904210
Languedoc-Roussillon	1609313	1797767	1986220	1856583	1931799	1912339	1872084	1857563
Provence Côte d Azur	2398443	2939830	3481218	3416681	4087824	3484341	4193936	3611428

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements pour le travail habituel par personne un jour de semaine

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	0.69	0.72	0.74	0.71	0.71	0.71	0.72	0.72
Champagne-Ardenne	0.61	0.76	0.91	0.69	0.66	0.68	0.68	0.70	
Picardie	0.48	0.61	0.74	0.70	0.66	0.68	0.68	0.69	
Haute-Normandie	0.36	0.58	0.79	0.64	0.66	0.67	0.67	0.68	
Centre	0.45	0.54	0.62	0.51	0.67	0.65	0.69	0.67	
Basse-Normandie	0.51	0.63	0.76	0.77	0.66	0.68	0.67	0.69	
Bourgogne	0.45	0.54	0.63	0.65	0.65	0.64	0.66	0.66	
Nord	0.47	0.55	0.62	0.58	0.61	0.58	0.63	0.60	
Lorraine	0.47	0.60	0.72	0.69	0.65	0.67	0.66	0.68	
Alsace	0.45	0.56	0.67	0.51	0.70	0.67	0.72	0.69	
Franche-Comté	0.38	0.53	0.67	0.58	0.67	0.66	0.67	0.66	
Pays de la Loire	0.78	0.85	0.91	0.82	0.69	0.76	0.71	0.77	
Bretagne	0.59	0.67	0.75	0.68	0.67	0.68	0.69	0.69	
Poitou-Charentes	0.55	0.69	0.82	0.70	0.65	0.65	0.66	0.66	
Aquitaine	0.58	0.66	0.74	0.64	0.66	0.67	0.67	0.68	
Midi-Pyrénées	0.76	0.81	0.86	0.73	0.67	0.69	0.69	0.71	
Limousin	0.53	0.74	0.95	0.58	0.64	0.64	0.64	0.64	
Rhône-Alpes	0.56	0.63	0.71	0.64	0.69	0.66	0.69	0.66	
Auvergne	0.42	0.54	0.66	0.54	0.66	0.63	0.67	0.64	
Languedoc-Roussillon	0.53	0.58	0.63	0.54	0.61	0.61	0.62	0.62	
Provence Côte d Azur	0.53	0.61	0.70	0.65	0.63	0.62	0.63	0.63	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements pour le travail habituel par région un jour de semaine

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	7266504	7527477	7788451	7439259	7418922	7436705	7569953	7550899
Champagne-Ardenne	657023	900311	1143598	834708	796609	826175	820486	852433	
Picardie	778978	1072825	1366672	1200597	1135525	1156589	1163980	1186322	
Haute-Normandie	419194	734565	1049935	1049724	1086123	1096139	1100577	1112643	
Centre	780245	1033895	1287545	1155066	1538591	1482743	1581904	1528945	
Basse-Normandie	677246	984435	1291624	1021808	868221	892841	884026	910863	
Bourgogne	654607	889113	1123618	955503	967966	947375	980272	970607	
Nord	1860058	2314231	2768404	2108297	2228333	2105880	2275856	2164133	
Lorraine	928052	1240406	1552759	1463984	1375514	1408870	1396543	1433231	
Alsace	650141	950061	1249982	855098	1167529	1114722	1196668	1139667	
Franche-Comté	376387	623326	870264	608618	704297	693285	707668	692479	
Pays de la Loire	2436971	2643927	2850883	2587162	2178840	2385332	2231878	2433918	
Bretagne	1683780	1907580	2131379	1919526	1888704	1918911	1945394	1955073	
Poitou-Charentes	805861	1183961	1562061	1105501	1036404	1040096	1049531	1055512	
Aquitaine	1725226	2072285	2419343	1855081	1903119	1934630	1931377	1965898	
Midi-Pyrénées	1940454	2068995	2197535	1892890	1739549	1771743	1781102	1820110	
Limousin	272962	532719	792476	387661	431549	428191	433200	431190	
Rhône-Alpes	2740908	3202319	3663730	3517855	3781854	3622566	3788286	3641368	
Auvergne	466549	680003	893456	651075	800949	761675	819944	782604	
Languedoc-Roussillon	1245341	1359000	1472659	1264743	1440563	1430500	1468365	1456662	
Provence Côte d Azur	2082586	2524674	2966762	2887894	2793776	2768550	2826305	2811886	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements pour les études par personne un jour de semaine

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	0.35	0.37	0.40	0.40	0.39	0.40	0.38	0.38
Champagne-Ardenne	0.10	0.31	0.52	0.55	0.38	0.39	0.37	0.37	
Picardie	0.20	0.32	0.45	0.46	0.40	0.41	0.40	0.40	
Haute-Normandie	0.19	0.31	0.43	0.38	0.40	0.39	0.37	0.37	
Centre	0.16	0.29	0.41	0.44	0.35	0.36	0.34	0.34	
Basse-Normandie	0.20	0.30	0.41	0.34	0.36	0.35	0.36	0.36	
Bourgogne	0.32	0.45	0.57	0.39	0.34	0.35	0.33	0.33	
Nord	0.30	0.38	0.45	0.38	0.44	0.42	0.41	0.41	
Lorraine	0.26	0.38	0.50	0.39	0.38	0.38	0.37	0.37	
Alsace	0.26	0.40	0.55	0.44	0.38	0.39	0.36	0.36	
Franche-Comté	0.21	0.40	0.59	0.54	0.37	0.38	0.39	0.39	
Pays de la Loire	0.32	0.35	0.39	0.38	0.38	0.38	0.37	0.37	
Bretagne	0.34	0.40	0.47	0.38	0.36	0.37	0.36	0.36	
Poitou-Charentes	0.12	0.19	0.26	0.24	0.34	0.33	0.34	0.34	
Aquitaine	0.22	0.30	0.38	0.32	0.35	0.34	0.34	0.34	
Midi-Pyrénées	0.25	0.29	0.32	0.36	0.36	0.36	0.34	0.34	
Limousin	0.03	0.10	0.17	0.16	0.31	0.30	0.33	0.33	
Rhône-Alpes	0.28	0.35	0.43	0.45	0.39	0.40	0.38	0.38	
Auvergne	0.14	0.25	0.36	0.29	0.33	0.32	0.32	0.32	
Languedoc-Roussillon	0.31	0.36	0.42	0.38	0.38	0.38	0.38	0.38	
Provence Côte d Azur	0.22	0.29	0.36	0.35	0.38	0.37	0.35	0.35	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements pour les études par région un jour de semaine

région	Estimateur national			Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire
	Ile de France	3683742	3943335	4202928	4155824	4128237	4135557	3989464
Champagne-Ardenne	111235	370584	629932	662441	463006	476894	444459	444459
Picardie	331203	567154	803105	790358	687535	694834	678907	678907
Haute-Normandie	235226	397551	559876	628943	652794	643171	609736	609736
Centre	277767	555885	834004	1015633	808256	816851	771320	771320
Basse-Normandie	268644	472500	676356	454292	480749	469076	480628	480628
Bourgogne	468019	735656	1003293	578627	509579	525269	486926	486926
Nord	1207014	1598685	1990357	1370989	1585495	1538928	1485288	1485288
Lorraine	527993	799138	1070284	818094	815270	800707	777701	777701
Alsace	342221	680697	1019172	724645	632907	650643	594323	594323
Franche-Comté	222832	476134	729437	563125	389806	404118	404521	404521
Pays de la Loire	991281	1102570	1213858	1208092	1180676	1182519	1165239	1165239
Bretagne	950786	1143829	1336871	1084948	1032694	1048063	1018774	1018774
Poitou-Charentes	187988	326450	464912	373861	536296	530192	545773	545773
Aquitaine	659684	936745	1213806	911016	1005595	989271	994219	994219
Midi-Pyrénées	631719	730187	828655	920566	921457	919431	875716	875716
Limousin	14505	73661	132817	107066	208909	204027	221380	221380
Rhône-Alpes	1357028	1798568	2240109	2492840	2144203	2219555	2110098	2110098
Auvergne	143147	316419	489690	354857	402419	392006	386908	386908
Languedoc-Roussillon	720890	850299	979709	889732	892817	897229	887826	887826
Provence Côte d Azur	845985	1178563	1511141	1558255	1695277	1641143	1541379	1541379

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements pour faire des achats par personne un jour de semaine

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	estimeur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	0.58	0.61	0.64	0.59	0.62	0.61	0.62	0.62
Champagne-Ardenne	0.34	0.47	0.60	0.45	0.62	0.58	0.63	0.57	
Picardie	0.60	0.76	0.91	0.72	0.61	0.62	0.60	0.62	
Haute-Normandie	0.51	0.64	0.77	0.57	0.62	0.62	0.62	0.62	
Centre	0.54	0.65	0.76	0.56	0.62	0.60	0.62	0.59	
Basse-Normandie	0.43	0.57	0.71	0.58	0.60	0.59	0.60	0.58	
Bourgogne	0.62	0.73	0.84	0.67	0.62	0.65	0.62	0.66	
Nord	0.57	0.66	0.75	0.66	0.62	0.63	0.62	0.63	
Lorraine	0.42	0.52	0.62	0.58	0.62	0.59	0.62	0.58	
Alsace	0.65	0.80	0.94	0.84	0.63	0.65	0.64	0.67	
Franche-Comté	0.46	0.60	0.74	0.53	0.61	0.59	0.62	0.59	
Pays de la Loire	0.59	0.65	0.71	0.64	0.61	0.62	0.61	0.63	
Bretagne	0.60	0.69	0.78	0.70	0.62	0.65	0.62	0.65	
Poitou-Charentes	0.47	0.56	0.65	0.50	0.62	0.59	0.62	0.59	
Aquitaine	0.57	0.65	0.73	0.63	0.63	0.62	0.63	0.61	
Midi-Pyrénées	0.55	0.59	0.64	0.57	0.62	0.60	0.62	0.60	
Limousin	0.68	0.90	1.13	0.84	0.63	0.67	0.64	0.69	
Rhône-Alpes	0.55	0.62	0.69	0.61	0.62	0.60	0.62	0.60	
Auvergne	0.54	0.68	0.83	0.60	0.62	0.62	0.62	0.62	
Languedoc-Roussillon	0.61	0.69	0.76	0.67	0.64	0.66	0.65	0.66	
Provence Côte d Azur	0.61	0.70	0.78	0.71	0.66	0.68	0.67	0.70	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements pour faire des achats par région un jour de semaine

région	Estimateur national			Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire
	Ile de France	6080861	6403372	6725884	6222093	6475423	6383065	6451639
Champagne-Ardenne	400206	555221	710237	539118	744544	697543	757958	693129
Picardie	976421	1331616	1686811	1226855	1034716	1066140	1029486	1066599
Haute-Normandie	593250	814948	1036645	940087	1016080	1025217	1016015	1013978
Centre	907163	1246366	1585568	1281952	1421077	1368751	1417930	1354262
Basse-Normandie	612667	884112	1155558	766727	794923	780132	793487	769069
Bourgogne	906480	1199668	1492857	988795	919923	966556	916877	975007
Nord	2270917	2798761	3326605	2379851	2234788	2293133	2260862	2291779
Lorraine	796061	1081493	1366925	1222641	1310179	1239654	1317302	1219814
Alsace	1003040	1350588	1698135	1395921	1040828	1080703	1057269	1107782
Franche-Comté	435273	713840	992406	559764	640588	624124	647290	623167
Pays de la Loire	1840162	2021859	2203555	2002027	1921271	1964343	1915756	1971913
Bretagne	1707866	1956211	2204556	1982301	1746987	1838369	1743698	1847888
Poitou-Charentes	715940	972572	1229204	788703	978673	944045	980519	944612
Aquitaine	1689905	2046986	2404067	1812152	1813542	1782215	1815005	1774496
Midi-Pyrénées	1399712	1514743	1629773	1472394	1604030	1555419	1599504	1550976
Limousin	358052	648586	939121	565925	425740	450527	431763	461205
Rhône-Alpes	2714554	3149082	3583611	3372376	3403609	3298223	3416347	3274529
Auvergne	635200	861265	1087329	729871	752170	753077	753870	754162
Languedoc-Roussillon	1442652	1617483	1792314	1568634	1520834	1549862	1537974	1560144
Provence Côte d Azur	2429407	2880865	3332322	3147976	2944124	3051253	2997034	3111512

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements pour les loisirs ou des visites par personne un jour de semaine

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	0.37	0.40	0.42	0.40	0.42	0.41	0.41	0.41
Champagne-Ardenne	0.23	0.34	0.44	0.35	0.40	0.39	0.40	0.38	
Picardie	0.17	0.24	0.31	0.23	0.40	0.33	0.38	0.30	
Haute-Normandie	0.27	0.44	0.61	0.52	0.40	0.42	0.39	0.40	
Centre	0.28	0.36	0.44	0.36	0.40	0.36	0.39	0.36	
Basse-Normandie	0.28	0.41	0.54	0.39	0.39	0.39	0.39	0.38	
Bourgogne	0.25	0.35	0.45	0.31	0.40	0.39	0.39	0.38	
Nord	0.28	0.35	0.42	0.35	0.41	0.37	0.40	0.36	
Lorraine	0.30	0.38	0.47	0.39	0.40	0.39	0.39	0.37	
Alsace	0.45	0.57	0.68	0.59	0.41	0.46	0.40	0.44	
Franche-Comté	0.35	0.51	0.67	0.44	0.40	0.41	0.40	0.40	
Pays de la Loire	0.45	0.51	0.57	0.50	0.41	0.48	0.40	0.47	
Bretagne	0.49	0.56	0.63	0.56	0.41	0.51	0.41	0.51	
Poitou-Charentes	0.25	0.33	0.41	0.29	0.40	0.36	0.40	0.34	
Aquitaine	0.27	0.33	0.39	0.30	0.41	0.33	0.41	0.32	
Midi-Pyrénées	0.38	0.43	0.47	0.40	0.42	0.40	0.41	0.40	
Limousin	0.25	0.44	0.62	0.41	0.40	0.43	0.40	0.42	
Rhône-Alpes	0.42	0.49	0.56	0.49	0.42	0.46	0.41	0.45	
Auvergne	0.32	0.44	0.57	0.44	0.40	0.43	0.39	0.42	
Languedoc-Roussillon	0.40	0.47	0.54	0.50	0.42	0.46	0.42	0.45	
Provence Côte d Azur	0.31	0.39	0.47	0.40	0.43	0.40	0.43	0.40	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Nombre de déplacements pour les loisirs ou des visites par région un jour de semaine

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	3912125	4189531	4466937	4149593	4366496	4268284	4311392	4252264
Champagne-Ardenne	230554	397286	564018	429189	486115	466325	481962	456848	
Picardie	286202	427877	569552	396393	678083	558095	651668	518685	
Haute-Normandie	315113	560528	805942	857254	663683	691309	639456	662485	
Centre	486261	691056	895851	822647	923896	831996	901748	823831	
Basse-Normandie	351907	639769	927631	509100	519895	515590	512971	507487	
Bourgogne	362922	574468	786014	454992	591613	576565	577470	559825	
Nord	1135766	1493687	1851608	1277124	1496324	1330446	1452780	1300879	
Lorraine	587526	797737	1007948	820091	853516	816649	830532	783716	
Alsace	668298	958141	1247983	988081	680516	762306	668281	739893	
Franche-Comté	426696	604829	782962	466365	420513	426715	416091	424857	
Pays de la Loire	1409523	1586643	1763762	1585007	1288687	1502833	1272951	1484208	
Bretagne	1388302	1588376	1788449	1578459	1161886	1442792	1153269	1436099	
Poitou-Charentes	406365	565789	725214	467939	636801	566752	634494	546074	
Aquitaine	817296	1037022	1256749	879337	1187124	962763	1172994	936781	
Midi-Pyrénées	960518	1085929	1211340	1036472	1071768	1038566	1049855	1019593	
Limousin	116909	312823	508737	277403	272413	286241	270636	283823	
Rhône-Alpes	2054647	2471878	2889109	2701041	2287092	2553667	2250973	2497163	
Auvergne	321990	556276	790561	530804	484201	517894	476731	506082	
Languedoc-Roussillon	949830	1104808	1259787	1170692	1001369	1076011	989531	1060206	
Provence Côte d Azur	1270161	1621529	1972897	1762160	1919881	1789907	1903186	1777792	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Durée totale des déplacements par personne un jour de semaine (minutes)

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
Ile de France	70	72	74	72	63	71	71	71	
Champagne-Ardenne	42	47	52	47	55	50	53	50	
Picardie	50	57	63	60	55	58	54	58	
Haute-Normandie	46	53	59	53	56	53	54	53	
Centre	43	47	51	46	55	49	54	50	
Basse-Normandie	36	40	44	43	54	45	52	46	
Bourgogne	45	49	54	48	54	51	53	50	
Nord	49	52	56	51	58	51	54	51	
Lorraine	46	51	56	51	56	52	54	51	
Alsace	56	60	64	62	57	59	55	59	
Franche-Comté	47	55	62	53	55	55	53	55	
Pays de la Loire	52	55	57	54	56	54	54	54	
Bretagne	51	55	58	55	54	55	54	55	
Poitou-Charentes	45	49	54	49	54	49	52	48	
Aquitaine	46	49	53	46	55	51	53	51	
Midi-Pyrénées	51	53	55	53	55	54	54	54	
Limousin	42	53	63	51	53	53	51	53	
Rhône-Alpes	51	55	58	55	57	55	55	56	
Auvergne	42	47	52	44	54	49	53	50	
Languedoc-Roussillon	50	53	56	53	55	54	53	54	
Provence Côte d Azur	57	63	70	67	58	63	54	62	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Durée moyenne d'un déplacement en transport en commun (minutes)

région	Modèle simplifié				Modèle détaillé			
	Estimateur national		Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	borne inf. de l'intervalle de confiance	estimateur ponctuel						borne sup. de l'intervalle de confiance
Ile de France	43	45	46	45	40	44	45	45
Champagne-Ardenne	16	23	30	23	36	28	35	33
Picardie	21	31	42	32	37	34	38	37
Haute-Normandie	19	27	34	26	37	30	35	34
Centre	26	32	38	39	37	36	37	34
Basse-Normandie	31	36	41	35	37	36	37	37
Bourgogne	25	31	36	32	37	31	37	33
Nord	28	32	36	30	38	32	35	34
Lorraine	21	26	30	26	38	29	36	31
Alsace	23	28	33	30	37	32	34	33
Franche-Comté	24	36	48	33	36	39	36	41
Pays de la Loire	30	32	34	32	36	33	34	35
Bretagne	28	35	41	36	36	35	36	35
Poitou-Charentes	18	25	32	24	36	30	36	33
Aquitaine	32	37	43	33	38	35	36	37
Midi-Pyrénées	29	33	37	32	37	35	35	35
Limousin	10	32	54	40	35	32	35	33
Rhône-Alpes	28	37	46	39	38	37	35	39
Auvergne	26	36	46	37	37	35	37	33
Languedoc-Roussillon	27	30	32	29	36	33	35	34
Provence Côte d Azur	31	36	40	35	38	36	32	34

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Durée moyenne d'un déplacement en voiture comme conducteur (minutes)

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	21	22	22	22	18	21	21	21
Champagne-Ardenne	15	16	18	16	17	16	17	16	
Picardie	16	17	19	17	17	17	17	18	
Haute-Normandie	16	19	22	17	17	18	17	18	
Centre	15	16	17	16	17	17	17	17	
Basse-Normandie	13	14	16	15	17	16	17	16	
Bourgogne	13	14	15	14	17	15	17	15	
Nord	15	16	17	15	17	16	16	16	
Lorraine	15	17	18	17	17	17	17	16	
Alsace	17	18	20	18	17	18	17	17	
Franche-Comté	16	17	19	17	17	17	17	17	
Pays de la Loire	14	15	15	15	17	15	17	15	
Bretagne	14	15	16	15	17	15	17	16	
Poitou-Charentes	16	18	20	19	17	17	16	17	
Aquitaine	14	15	16	15	17	16	16	16	
Midi-Pyrénées	15	16	17	17	17	17	17	17	
Limousin	14	16	19	16	17	16	17	17	
Rhône-Alpes	16	17	18	16	17	17	17	17	
Auvergne	14	15	17	15	17	17	17	16	
Languedoc-Roussillon	16	16	17	16	17	16	17	17	
Provence Côte d Azur	16	19	22	21	17	19	17	19	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Distance totale parcourue à vol d'oiseau ⁴ par personne un jour de semaine (km)

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	13.8	14.3	14.8	14.4	13.8	14.3	13.7	14.0
Champagne-Ardenne	12.0	15.3	18.7	13.0	16.6	15.0	16.7	15.5	
Picardie	17.4	21.3	25.1	22.4	17.6	19.5	18.1	20.3	
Haute-Normandie	14.6	19.3	23.9	17.0	16.1	16.9	16.5	17.9	
Centre	12.8	15.3	17.9	16.1	17.5	16.8	17.7	17.6	
Basse-Normandie	10.4	12.5	14.6	12.0	18.0	14.5	18.5	15.2	
Bourgogne	11.1	13.3	15.4	13.4	17.7	15.3	18.1	15.6	
Nord	13.4	15.2	16.9	14.1	14.2	14.2	14.2	14.2	
Lorraine	15.3	18.1	20.8	17.0	16.4	16.8	16.6	16.8	
Alsace	13.9	16.8	19.7	16.2	16.5	16.4	16.4	16.5	
Franche-Comté	15.5	18.7	21.8	17.6	17.7	17.9	17.8	18.3	
Pays de la Loire	17.3	18.5	19.8	18.1	17.1	17.8	17.5	17.6	
Bretagne	16.9	18.9	20.8	19.3	17.6	18.8	18.0	18.8	
Poitou-Charentes	19.5	22.7	25.8	21.9	17.9	18.8	18.3	19.1	
Aquitaine	13.9	16.2	18.4	14.0	16.8	15.5	16.9	16.2	
Midi-Pyrénées	15.2	16.5	17.8	16.6	17.1	16.7	17.5	17.0	
Limousin	14.1	18.9	23.7	17.3	17.3	17.0	17.3	17.2	
Rhône-Alpes	14.6	16.4	18.2	16.5	16.3	16.7	16.2	17.0	
Auvergne	12.3	15.5	18.8	14.3	17.7	16.5	18.0	16.9	
Languedoc-Roussillon	12.9	14.3	15.8	14.0	16.0	14.7	16.0	14.7	
Provence Côte d Azur	12.7	14.5	16.2	14.4	13.5	13.8	13.0	13.2	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

⁴ Il s'agit des distances à vol d'oiseau entre les centres des communes (des arrondissements à Paris, Lyon, Marseille) de départ et d'arrivée de chaque déplacement. Par construction, les déplacements internes à une même commune sont donc affectés d'une distance nulle.

Distance moyenne d'un déplacement en voiture (conducteur)

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	6.3	6.7	7.1	6.9	6.1	6.7	6.4	6.5
Champagne-Ardenne	5.4	7.1	8.7	5.9	7.1	6.9	7.1	7.0	
Picardie	6.6	8.0	9.4	8.1	7.3	7.4	7.4	7.7	
Haute-Normandie	6.5	9.0	11.5	7.8	6.8	7.7	6.8	7.9	
Centre	5.3	6.2	7.2	6.8	7.2	7.0	7.1	7.2	
Basse-Normandie	5.9	6.9	7.9	6.1	7.4	7.0	7.6	7.4	
Bourgogne	4.7	5.5	6.3	5.6	7.4	6.1	7.5	6.5	
Nord	5.3	6.1	6.9	5.7	6.2	5.8	5.7	5.7	
Lorraine	6.7	7.8	9.0	7.4	6.8	7.1	6.8	7.1	
Alsace	7.2	8.7	10.1	8.6	6.8	7.9	6.8	7.7	
Franche-Comté	7.1	8.5	9.8	8.0	7.3	7.8	7.4	7.9	
Pays de la Loire	6.2	6.7	7.2	6.8	7.0	6.8	7.1	6.8	
Bretagne	6.2	6.9	7.5	7.0	7.2	7.0	7.3	7.1	
Poitou-Charentes	8.6	10.1	11.6	10.5	7.4	8.6	7.4	8.6	
Aquitaine	5.3	6.0	6.8	5.3	6.9	6.1	6.7	6.3	
Midi-Pyrénées	6.0	6.5	7.1	6.8	7.0	6.9	7.0	6.9	
Limousin	4.8	7.0	9.3	7.2	7.3	7.1	7.4	7.4	
Rhône-Alpes	6.2	6.8	7.4	6.4	6.7	6.8	6.6	6.9	
Auvergne	6.2	7.2	8.2	7.2	7.3	7.6	7.3	7.6	
Languedoc-Roussillon	5.6	6.2	6.7	6.0	6.8	6.2	6.8	6.4	
Provence Côte d Azur	5.1	5.8	6.5	5.7	5.8	5.5	5.6	5.3	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Proportion de personnes se rendant sur leur lieu de travail ou d'étude en transport en commun

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	0.2921	0.3107	0.3293	0.3097	0.2259	0.3065	0.3023	0.3022
Champagne-Ardenne	0.0193	0.1236	0.2279	0.2138	0.1573	0.1664	0.1349	0.1349	
Picardie	0.0152	0.0603	0.1054	0.1009	0.1507	0.1182	0.1412	0.1410	
Haute-Normandie	0.0401	0.1012	0.1623	0.1387	0.1662	0.1354	0.1407	0.1404	
Centre	0.0300	0.0787	0.1274	0.0899	0.1429	0.1120	0.1213	0.1216	
Basse-Normandie	0.0144	0.1184	0.2225	0.1609	0.1402	0.1392	0.1250	0.1246	
Bourgogne	0.0438	0.1510	0.2582	0.1954	0.1386	0.1437	0.1260	0.1259	
Nord	0.0940	0.1522	0.2104	0.1443	0.1966	0.1507	0.1498	0.1505	
Lorraine	0.0561	0.1133	0.1706	0.0845	0.1613	0.1238	0.1340	0.1342	
Alsace	0.0227	0.0685	0.1142	0.1006	0.1586	0.1106	0.1326	0.1330	
Franche-Comté	0.0366	0.1606	0.2847	0.1755	0.1468	0.1626	0.1289	0.1288	
Pays de la Loire	0.0996	0.1223	0.1449	0.1194	0.1470	0.1300	0.1221	0.1210	
Bretagne	0.0683	0.1184	0.1685	0.1258	0.1367	0.1122	0.1210	0.1210	
Poitou-Charentes	0.0051	0.0240	0.0428	0.0259	0.1364	0.0874	0.1123	0.1123	
Aquitaine	0.0561	0.0954	0.1346	0.1154	0.1487	0.1352	0.1150	0.1148	
Midi-Pyrénées	0.0505	0.0690	0.0874	0.0717	0.1424	0.1159	0.1178	0.1179	
Limousin	- .0084	0.0079	0.0241	0.0056	0.1362	0.0924	0.1132	0.1136	
Rhône-Alpes	0.0854	0.1258	0.1663	0.1452	0.1627	0.1488	0.1295	0.1298	
Auvergne	- .0193	0.1015	0.2223	0.0994	0.1364	0.1142	0.1203	0.1209	
Languedoc-Roussillon	0.0768	0.1022	0.1276	0.1006	0.1544	0.1284	0.1288	0.1290	
Provence Côte d Azur	0.0855	0.1325	0.1795	0.1573	0.1822	0.1583	0.1294	0.1293	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Proportion de personnes se rendant sur leur lieu de travail ou d'étude en voiture (conducteur)

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	0.2889	0.3057	0.3225	0.2992	0.3972	0.3261	0.3075	0.3075
Champagne-Ardenne	0.4529	0.5822	0.7116	0.4792	0.4835	0.4896	0.5130	0.5130	
Picardie	0.4869	0.5925	0.6981	0.5914	0.4973	0.5224	0.5183	0.5176	
Haute-Normandie	0.4203	0.5421	0.6638	0.5465	0.4745	0.5190	0.5155	0.5145	
Centre	0.4990	0.5916	0.6842	0.5582	0.5202	0.5476	0.5503	0.5515	
Basse-Normandie	0.4615	0.5555	0.6496	0.5339	0.5190	0.5426	0.5411	0.5394	
Bourgogne	0.3503	0.4457	0.5411	0.4474	0.5258	0.4993	0.5447	0.5440	
Nord	0.4399	0.5068	0.5736	0.5137	0.4182	0.4739	0.4869	0.4892	
Lorraine	0.4501	0.5382	0.6263	0.5297	0.4906	0.5162	0.5258	0.5266	
Alsace	0.3939	0.4917	0.5896	0.4595	0.4957	0.4783	0.5275	0.5291	
Franche-Comté	0.3666	0.4929	0.6192	0.4652	0.5138	0.5111	0.5314	0.5310	
Pays de la Loire	0.5004	0.5302	0.5600	0.5126	0.5056	0.5222	0.5395	0.5344	
Bretagne	0.5103	0.5596	0.6090	0.5473	0.5200	0.5467	0.5405	0.5409	
Poitou-Charentes	0.6210	0.7078	0.7947	0.6623	0.5334	0.5790	0.5623	0.5621	
Aquitaine	0.5066	0.5749	0.6433	0.5681	0.5134	0.5460	0.5535	0.5523	
Midi-Pyrénées	0.5229	0.5550	0.5871	0.5300	0.5146	0.5187	0.5510	0.5511	
Limousin	0.7264	0.8229	0.9195	0.7142	0.5333	0.5528	0.5610	0.5630	
Rhône-Alpes	0.4968	0.5558	0.6148	0.5235	0.4844	0.5066	0.5166	0.5179	
Auvergne	0.3629	0.4776	0.5924	0.4836	0.5331	0.5184	0.5562	0.5589	
Languedoc-Roussillon	0.4561	0.4933	0.5306	0.4623	0.4793	0.4843	0.5095	0.5100	
Provence Côte d Azur	0.4228	0.4906	0.5584	0.4427	0.4383	0.4594	0.4967	0.4966	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Proportion de personnes se rendant sur leur lieu de travail ou d'étude en voiture (passager)

région	Estimateur national			Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire
	Ile de France	0.0347	0.0425	0.0502	0.0444	0.0609	0.0528	0.0455
Champagne-Ardenne	0.0256	0.0642	0.1028	0.0517	0.0816	0.0751	0.0828	0.0763
Picardie	0.0312	0.1053	0.1794	0.1240	0.0891	0.0941	0.0976	0.0981
Haute-Normandie	0.0059	0.0360	0.0660	0.0400	0.0815	0.0700	0.0780	0.0694
Centre	0.0135	0.0569	0.1003	0.0509	0.0830	0.0754	0.0844	0.0790
Basse-Normandie	0.1323	0.2050	0.2777	0.1267	0.0873	0.1223	0.0927	0.1174
Bourgogne	0.0002	0.0543	0.1083	0.0405	0.0848	0.0754	0.0803	0.0732
Nord	0.0526	0.0943	0.1360	0.1174	0.0777	0.0986	0.0830	0.0963
Lorraine	0.0331	0.0855	0.1380	0.0647	0.0821	0.0814	0.0803	0.0809
Alsace	0.0218	0.0919	0.1620	0.1383	0.0781	0.0848	0.0732	0.0750
Franche-Comté	- .0136	0.0481	0.1098	0.0911	0.0838	0.0773	0.0896	0.0845
Pays de la Loire	0.0687	0.0907	0.1128	0.1009	0.0812	0.0908	0.0884	0.0957
Bretagne	0.0402	0.0750	0.1099	0.0773	0.0861	0.0859	0.0886	0.0893
Poitou-Charentes	0.0659	0.1189	0.1718	0.1236	0.0824	0.0938	0.0944	0.0988
Aquitaine	0.0411	0.0728	0.1044	0.0631	0.0788	0.0699	0.0844	0.0722
Midi-Pyrénées	0.0481	0.0714	0.0947	0.0775	0.0811	0.0816	0.0780	0.0768
Limousin	- .0074	0.0611	0.1295	0.1121	0.0764	0.0793	0.0885	0.0887
Rhône-Alpes	0.0271	0.0549	0.0828	0.0633	0.0794	0.0718	0.0777	0.0681
Auvergne	0.1034	0.1980	0.2927	0.2518	0.0821	0.1070	0.0778	0.0939
Languedoc-Roussillon	0.1004	0.1320	0.1635	0.1328	0.0845	0.0987	0.0924	0.1004
Provence Côte d Azur	0.0197	0.0467	0.0737	0.0498	0.0711	0.0642	0.0624	0.0558

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Proportion de personnes se rendant sur leur lieu de travail ou d'étude à pied ou en vélo

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	0.1896	0.2066	0.2235	0.2124	0.1848	0.1917	0.2101	0.2101
Champagne-Ardenne	0.0483	0.1594	0.2706	0.1795	0.1544	0.1633	0.1496	0.1496	
Picardie	0.0702	0.1225	0.1748	0.1100	0.1443	0.1463	0.1284	0.1283	
Haute-Normandie	0.0931	0.2063	0.3195	0.1723	0.1538	0.1603	0.1422	0.1419	
Centre	0.0488	0.1490	0.2491	0.2040	0.1339	0.1432	0.1250	0.1252	
Basse-Normandie	0.0052	0.0677	0.1302	0.1147	0.1331	0.1174	0.1251	0.1247	
Bourgogne	0.0803	0.1761	0.2719	0.1719	0.1327	0.1427	0.1316	0.1315	
Nord	0.0755	0.1227	0.1699	0.1138	0.1789	0.1559	0.1477	0.1484	
Lorraine	0.0752	0.1453	0.2155	0.1670	0.1451	0.1477	0.1391	0.1393	
Alsace	0.1582	0.2461	0.3340	0.2097	0.1440	0.1758	0.1482	0.1486	
Franche-Comté	0.0405	0.1272	0.2140	0.1414	0.1340	0.1337	0.1362	0.1361	
Pays de la Loire	0.0754	0.0995	0.1235	0.1011	0.1393	0.1228	0.1295	0.1283	
Bretagne	0.0773	0.1164	0.1555	0.1138	0.1341	0.1230	0.1292	0.1293	
Poitou-Charentes	0.0271	0.0893	0.1516	0.1074	0.1271	0.1246	0.1161	0.1161	
Aquitaine	0.0461	0.0868	0.1276	0.0917	0.1358	0.1196	0.1249	0.1246	
Midi-Pyrénées	0.1208	0.1457	0.1707	0.1441	0.1370	0.1401	0.1292	0.1292	
Limousin	-.0071	0.0342	0.0754	0.0932	0.1319	0.1297	0.1254	0.1258	
Rhône-Alpes	0.1331	0.1809	0.2287	0.1887	0.1468	0.1649	0.1513	0.1517	
Auvergne	0.0679	0.1438	0.2198	0.1247	0.1294	0.1290	0.1268	0.1274	
Languedoc-Roussillon	0.0936	0.1236	0.1536	0.1431	0.1572	0.1463	0.1561	0.1563	
Provence Côte d Azur	0.0978	0.1557	0.2136	0.2031	0.1787	0.1753	0.1859	0.1858	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Proportion des déplacements vers le lieu de travail ou d'étude en transport en commun

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	0.3112	0.3303	0.3494	0.3330	0.2354	0.3196	0.3195	0.3200
Champagne-Ardenne	0.0112	0.1390	0.2669	0.2370	0.1645	0.1901	0.1442	0.1412	
Picardie	0.0286	0.0795	0.1304	0.0949	0.1572	0.1340	0.1472	0.1465	
Haute-Normandie	0.0417	0.1663	0.2909	0.1995	0.1744	0.1800	0.1523	0.1526	
Centre	0.0404	0.0867	0.1330	0.1081	0.1496	0.1189	0.1329	0.1340	
Basse-Normandie	0.0240	0.1153	0.2065	0.1481	0.1462	0.1357	0.1295	0.1296	
Bourgogne	0.0872	0.1694	0.2516	0.2056	0.1447	0.1487	0.1327	0.1318	
Nord	0.1157	0.1723	0.2289	0.1636	0.2047	0.1613	0.1623	0.1657	
Lorraine	0.0595	0.1238	0.1881	0.1029	0.1688	0.1392	0.1414	0.1404	
Alsace	0.0286	0.0670	0.1055	0.1003	0.1676	0.1125	0.1447	0.1456	
Franche-Comté	0.0727	0.1672	0.2616	0.1700	0.1545	0.1713	0.1320	0.1315	
Pays de la Loire	0.1061	0.1276	0.1490	0.1373	0.1555	0.1461	0.1310	0.1264	
Bretagne	0.0735	0.1146	0.1556	0.1191	0.1433	0.1148	0.1294	0.1286	
Poitou-Charentes	0.0090	0.0401	0.0712	0.0425	0.1433	0.1013	0.1173	0.1180	
Aquitaine	0.0847	0.1262	0.1676	0.1439	0.1559	0.1485	0.1231	0.1234	
Midi-Pyrénées	0.0524	0.0690	0.0856	0.0803	0.1504	0.1210	0.1300	0.1292	
Limousin	- .0001	0.0180	0.0361	0.0139	0.1441	0.1007	0.1194	0.1204	
Rhône-Alpes	0.0855	0.1237	0.1620	0.1424	0.1707	0.1536	0.1387	0.1392	
Auvergne	0.0034	0.0667	0.1300	0.0598	0.1421	0.0956	0.1270	0.1302	
Languedoc-Roussillon	0.0736	0.0959	0.1183	0.1000	0.1600	0.1257	0.1332	0.1337	
Provence Côte d Azur	0.0918	0.1451	0.1985	0.1685	0.1911	0.1695	0.1445	0.1460	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Proportion des déplacements vers le lieu de travail ou d'étude en voiture (conducteur)

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	0.2908	0.3088	0.3267	0.2991	0.3991	0.3213	0.3149	0.3133
Champagne-Ardenne	0.4511	0.5960	0.7408	0.4633	0.4854	0.5234	0.5157	0.5280	
Picardie	0.4814	0.5980	0.7146	0.5896	0.4947	0.5315	0.5160	0.5309	
Haute-Normandie	0.3742	0.5052	0.6363	0.5010	0.4726	0.4886	0.5126	0.5144	
Centre	0.5135	0.6125	0.7115	0.5642	0.5198	0.5504	0.5489	0.5661	
Basse-Normandie	0.4679	0.5716	0.6753	0.5234	0.5208	0.5468	0.5443	0.5532	
Bourgogne	0.3706	0.4584	0.5462	0.4695	0.5268	0.5011	0.5469	0.5336	
Nord	0.4371	0.5011	0.5652	0.5051	0.4152	0.4550	0.4848	0.4844	
Lorraine	0.4287	0.5221	0.6154	0.5102	0.4864	0.5181	0.5222	0.5293	
Alsace	0.2926	0.3925	0.4925	0.3429	0.4923	0.4194	0.5203	0.4749	
Franche-Comté	0.3069	0.4435	0.5800	0.4141	0.5131	0.4888	0.5290	0.5065	
Pays de la Loire	0.5366	0.5718	0.6070	0.5489	0.5070	0.5613	0.5404	0.5591	
Bretagne	0.5133	0.5645	0.6156	0.5674	0.5216	0.5666	0.5435	0.5580	
Poitou-Charentes	0.6433	0.7328	0.8222	0.6634	0.5363	0.5910	0.5679	0.5887	
Aquitaine	0.5096	0.5834	0.6571	0.5764	0.5158	0.5494	0.5559	0.5682	
Midi-Pyrénées	0.5373	0.5672	0.5970	0.5344	0.5197	0.5287	0.5553	0.5535	
Limousin	0.7583	0.8450	0.9318	0.7173	0.5397	0.5782	0.5633	0.5748	
Rhône-Alpes	0.4645	0.5263	0.5880	0.4830	0.4850	0.4801	0.5120	0.4927	
Auvergne	0.3834	0.4875	0.5917	0.4859	0.5361	0.5012	0.5619	0.5399	
Languedoc-Roussillon	0.4651	0.5051	0.5451	0.4630	0.4826	0.4920	0.5084	0.5022	
Provence Côte d Azur	0.4475	0.5155	0.5834	0.4563	0.4412	0.4590	0.4914	0.4810	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Proportion des déplacements vers le lieu de travail ou d'étude en voiture (passager)

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	Estimateur national estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	0.0532	0.0614	0.0696	0.0644	0.0938	0.0832	0.0676	0.0677
Champagne-Ardenne	0.0155	0.0416	0.0677	0.0359	0.1193	0.1053	0.1179	0.1155	
Picardie	0.0676	0.1721	0.2767	0.2072	0.1299	0.1372	0.1380	0.1373	
Haute-Normandie	0.0160	0.0476	0.0792	0.0436	0.1206	0.1107	0.1165	0.1167	
Centre	0.0325	0.0862	0.1399	0.0521	0.1214	0.1143	0.1202	0.1212	
Basse-Normandie	0.1321	0.1944	0.2568	0.1357	0.1262	0.1354	0.1282	0.1283	
Bourgogne	0.0695	0.1421	0.2146	0.0956	0.1229	0.1265	0.1176	0.1168	
Nord	0.0889	0.1335	0.1782	0.1490	0.1163	0.1298	0.1214	0.1239	
Lorraine	0.0813	0.1366	0.1920	0.1357	0.1210	0.1206	0.1185	0.1176	
Alsace	0.0767	0.1839	0.2911	0.2440	0.1165	0.1318	0.1066	0.1072	
Franche-Comté	0.0240	0.0853	0.1466	0.1044	0.1228	0.1152	0.1253	0.1249	
Pays de la Loire	0.0964	0.1207	0.1450	0.1275	0.1195	0.1207	0.1215	0.1172	
Bretagne	0.0773	0.1195	0.1617	0.1234	0.1252	0.1286	0.1234	0.1226	
Poitou-Charentes	0.0575	0.1052	0.1529	0.1213	0.1195	0.1198	0.1280	0.1287	
Aquitaine	0.0975	0.1391	0.1806	0.1295	0.1154	0.1186	0.1193	0.1196	
Midi-Pyrénées	0.0774	0.0985	0.1196	0.1070	0.1185	0.1174	0.1115	0.1108	
Limousin	- .0035	0.0862	0.1760	0.1691	0.1116	0.1129	0.1171	0.1180	
Rhône-Alpes	0.0533	0.0938	0.1342	0.1078	0.1180	0.1162	0.1161	0.1165	
Auvergne	0.1357	0.2446	0.3535	0.3213	0.1186	0.1334	0.1119	0.1147	
Languedoc-Roussillon	0.1515	0.1898	0.2282	0.2123	0.1230	0.1417	0.1255	0.1259	
Provence Côte d Azur	0.0460	0.0845	0.1229	0.0927	0.1061	0.1049	0.0945	0.0955	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Proportion des déplacements vers le lieu de travail ou à pied ou en vélo

région	Estimateur national				Modèle simplifié			Modèle détaillé	
	borne inf. de l'intervalle de confiance	estimateur ponctuel	borne sup. de l'intervalle de confiance	Estimateur direct par calage régional	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	Estimateur synthétique de type régression	Estimateur par modèle mixte effet régional aléatoire	
	Ile de France	0.2467	0.2659	0.2850	0.2679	0.2379	0.2466	0.2633	0.2626
Champagne-Ardenne	0.0494	0.1856	0.3219	0.2232	0.2053	0.2143	0.1988	0.2043	
Picardie	0.0820	0.1354	0.1887	0.1040	0.1945	0.1864	0.1798	0.1764	
Haute-Normandie	0.1387	0.2544	0.3700	0.2094	0.2055	0.2072	0.1910	0.1898	
Centre	0.0616	0.1800	0.2983	0.2468	0.1845	0.1843	0.1725	0.1723	
Basse-Normandie	0.0087	0.1055	0.2024	0.1812	0.1824	0.1606	0.1764	0.1659	
Bourgogne	0.1328	0.2231	0.3134	0.2277	0.1822	0.2022	0.1803	0.1945	
Nord	0.1363	0.1930	0.2497	0.1823	0.2346	0.2151	0.1952	0.1957	
Lorraine	0.1270	0.2175	0.3080	0.2511	0.1982	0.2080	0.1916	0.2016	
Alsace	0.2253	0.3460	0.4667	0.3085	0.1972	0.2544	0.2024	0.2297	
Franche-Comté	0.1176	0.2653	0.4130	0.2858	0.1846	0.2155	0.1934	0.2072	
Pays de la Loire	0.1061	0.1382	0.1703	0.1498	0.1915	0.1729	0.1811	0.1752	
Bretagne	0.1244	0.1781	0.2319	0.1677	0.1864	0.1790	0.1804	0.1785	
Poitou-Charentes	0.0428	0.1085	0.1742	0.1580	0.1764	0.1660	0.1654	0.1618	
Aquitaine	0.0725	0.1092	0.1458	0.1034	0.1874	0.1560	0.1748	0.1515	
Midi-Pyrénées	0.1881	0.2148	0.2415	0.2155	0.1864	0.1958	0.1759	0.1831	
Limousin	-.0045	0.0507	0.1059	0.0997	0.1796	0.1703	0.1788	0.1781	
Rhône-Alpes	0.1925	0.2482	0.3038	0.2622	0.1994	0.2285	0.2066	0.2229	
Auvergne	0.1057	0.1749	0.2441	0.1151	0.1794	0.1696	0.1750	0.1708	
Languedoc-Roussillon	0.1420	0.1782	0.2144	0.1999	0.2115	0.1972	0.2134	0.2040	
Provence Côte d Azur	0.1260	0.1834	0.2408	0.2236	0.2333	0.2157	0.2384	0.2280	

Champ : population âgée de 6 ans ou plus, déplacements pendant un jour de semaine du lundi au vendredi dans un rayon de 80 km autour du domicile.

Annexe 7 - Estimation de la variance de l'estimateur direct

L'échantillon de l'enquête Transport et déplacements résulte de la fusion par partage des poids de deux échantillons : l'échantillon national⁵ d'une part, l'extension locale sur le périmètre du département de Loire-Atlantique et de la Division territoriale d'aménagement de l'estuaire de la Loire⁶, d'autre part.

Chacun de ces échantillons a été sélectionné selon son propre plan de sondage. L'échantillon national n'étant pas stratifié par région, son intersection avec la région Pays de la Loire ou avec le département de Loire-Atlantique (comme avec n'importe quelle autre région de métropole) constitue un « domaine » au sens de la théorie des sondages, c'est-à-dire un sous-échantillon dont la taille, aléatoire, n'est pas contrôlée par les paramètres du tirage. Au contraire, l'extension locale a été sélectionnée directement dans une base de sondage ne contenant que les logements situés en Loire-Atlantique, et dans les communes du Maine et Loir concernées.

Pour simplifier, on désignera par la suite par Loire-Atlantique le champ géographique de l'extension d'échantillon. Le lecteur se rappellera que celui-ci inclut également la DTA complète.

Cette note décrit les diverses étapes d'estimation de la variance d'un paramètre estimé par l'enquête Transport, après redressement par calage de l'aléa d'échantillonnage et du biais de non-réponse.

LE PLAN DE SONDRAGE ET L'ARTICULATION DES QUESTIONNAIRES

Les figures 1 et 2 résument le schéma de collecte de l'enquête Transport. Celle-ci s'est faite en deux visites, qui représentent deux degrés du plan de sondage.

En première visite, le questionnaire d'enquête observe différents types d'unités, correspondant aussi à des degrés différents du plan de sondage : le ménage, l'individu, le véhicule. Le ménage s'identifie à la résidence principale. Il est sélectionné par le biais d'un échantillonnage de logements. Individus et véhicules des ménages sélectionnés sont décrits de façon exhaustive. Certains d'entre eux ont par ailleurs été sélectionnés aléatoirement à l'intérieur de chaque ménage pour répondre à une partie du questionnaire. Selon la variable étudiée, on se situe donc, conditionnellement à la base de sondage utilisée -- restriction du recensement de 1999 ou de la base des logements neufs aux communes de l'échantillon-maître -- dans un sondage à un degré, dans un sondage en grappes, ou dans un sondage à deux degrés.

En fin de première visite, un individu Kish unique est sélectionné dans le ménage avec des probabilités différenciées par unité, pour renseigner un carnet journalier de déplacements pendant une semaine et répondre au questionnaire de la deuxième visite à l'aide de ce carnet.

Lors de la deuxième visite, un jour de la semaine, du lundi au vendredi, est sélectionné à probabilités égales pour décrire tous les déplacements locaux (de moins de 80 km autour du domicile) réalisés par l'individu Kish pendant cette journée.

Les voyages de longue distance (couvrant 80 km ou plus autour du domicile) constituent une partie du questionnaire de deuxième visite. Ils sont relevés sur une période de référence de 13 semaines précédant l'enquête. Ils avaient vocation à être décrits de façon exhaustive pour chaque individu Kish, mais la présence dans l'échantillon de quelques grands voyageurs a contraint là encore à un échantillonnage des voyages lorsque leur fréquence dépassait le seuil de 10.

Une phase supplémentaire de non-réponse s'est interposée entre les deux questionnaires : certains ménages ayant répondu en première visite ont refusé de poursuivre l'enquête à son terme, ou n'en ont pas eu le temps dans les délais impartis par le calendrier de la collecte. Il s'ensuit que l'échantillon de répondants en visite 2 est réduit par rapport à l'échantillon de répondants en première visite.

⁵ On inclut ici dans l'échantillon national les extensions régionales par tirage complémentaire dans l'EMEX.

⁶ Division territoriale d'aménagement (DTA) : zone d'étude spécifique à la direction régionale de l'Équipement des Pays de la Loire. La DTA est principalement incluse dans la Loire-Atlantique, mais comprend également 20 communes du Maine et Loire.

Figure 1. Le plan de sondage national

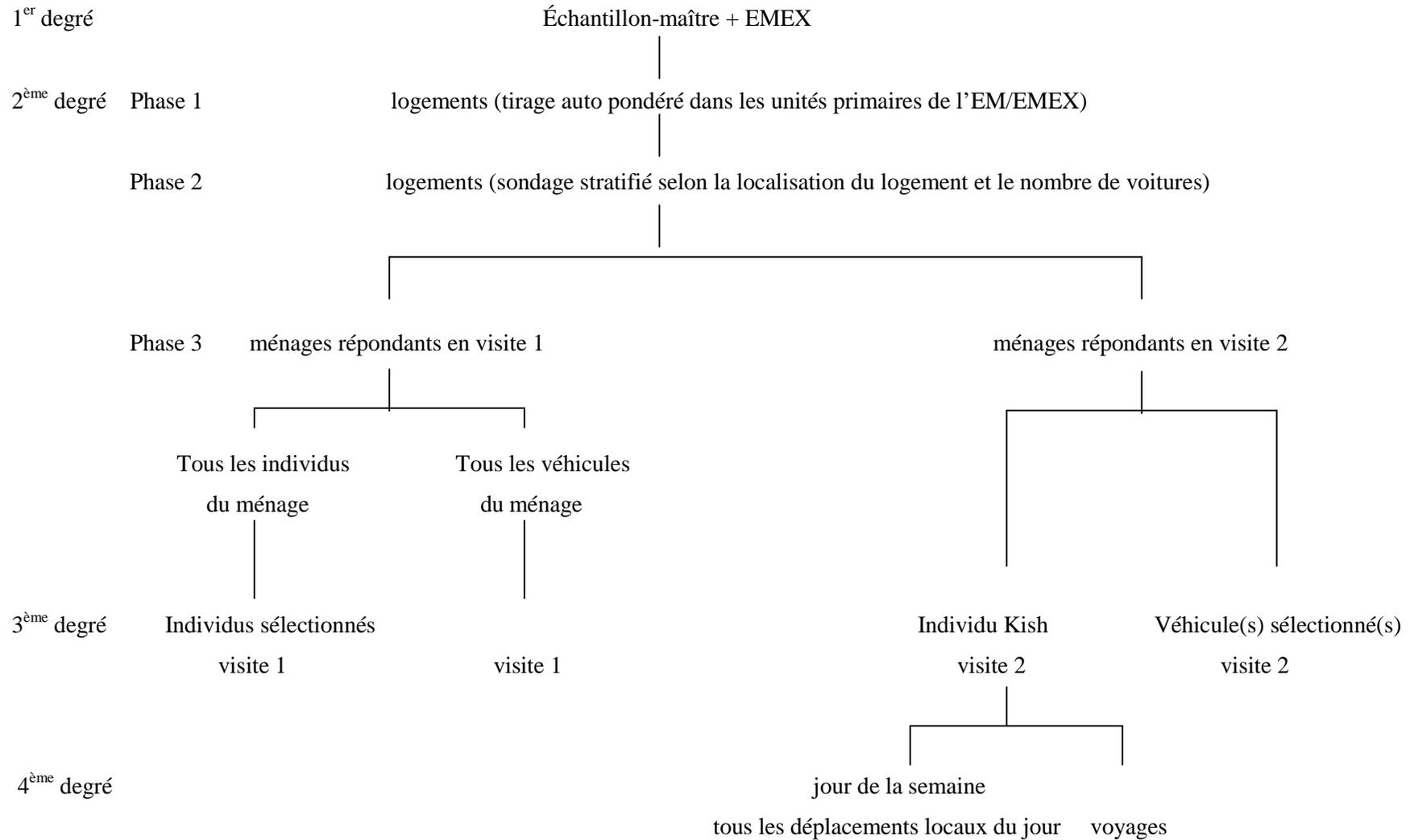
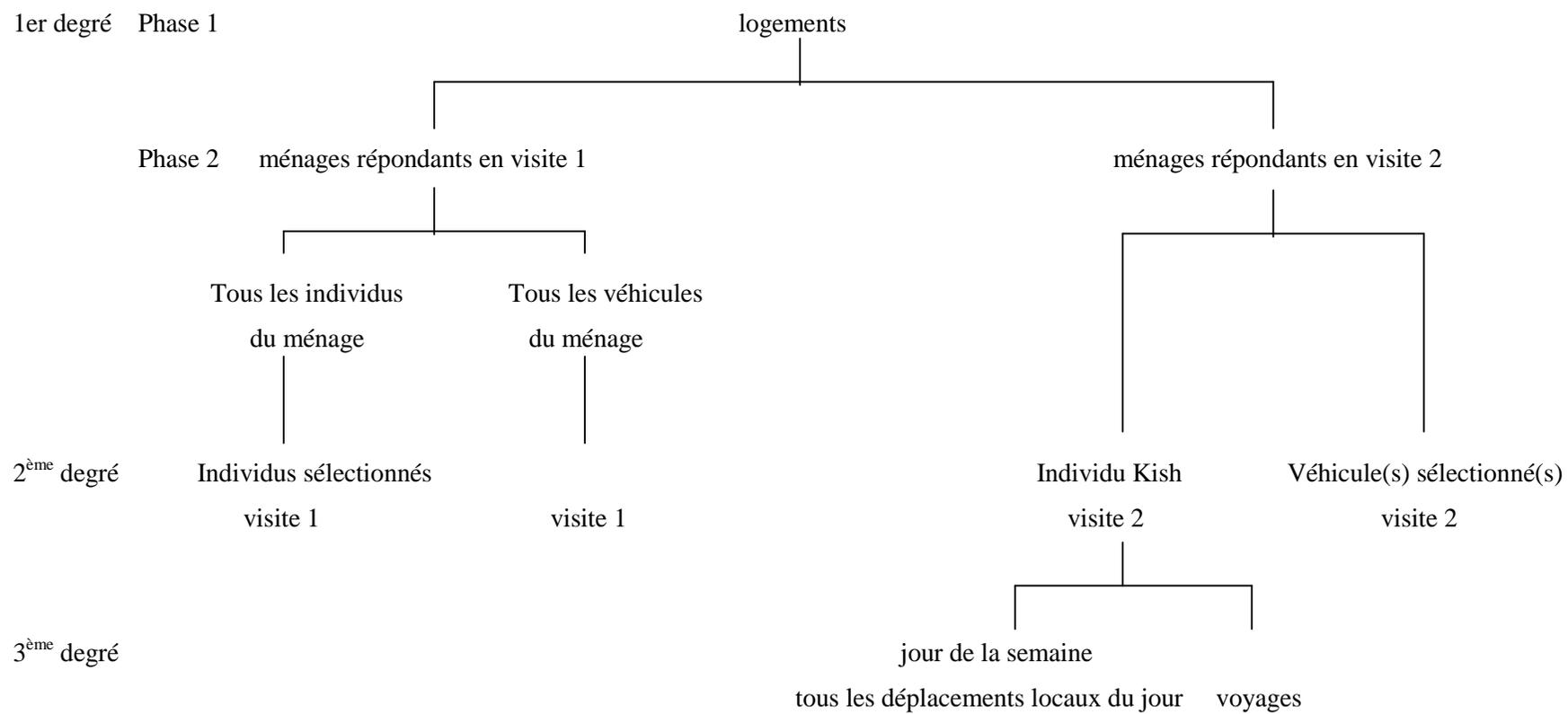


Figure 2. Le plan de sondage de l'extension locale de Loire-Atlantique



Le plan de sondage des logements

L'échantillon national

L'échantillon national de logements est sélectionné par un plan qu'on peut assimiler à un sondage à deux degrés, le premier étant constitué par la réunion de l'échantillon maître (EM) et de l'échantillon maître pour les extensions régionales (EMEX) dans les cinq régions ayant réalisé des extensions, par l'échantillon-maître seul dans le reste de la France.

L'échantillon maître est un échantillon aréolaire, stratifié par taille d'unité urbaine et groupe de régions. Selon la strate de tirage, l'unité sélectionnée (unité primaire) est un canton rural, une unité urbaine ou un îlot. Les unités primaires ont été tirées avec des probabilités proportionnelles à leur taille en nombre de résidences principales au recensement de 1999.

L'EMEX est un échantillon complémentaire d'unités urbaines et de communes rurales, destiné à assurer une représentativité régionale. Il a été sélectionné de façon coordonnée par rapport à l'échantillon maître national, mais selon un plan de sondage identique. Dans une région avec extension régionale, l'échantillon de logements est sélectionné dans les communes appartenant à l'EM et à l'EMEX, les unités primaires étant repondérées par partage des poids entre les deux échantillons.

On a traité l'ensemble des unités primaires mobilisées pour le tirage de l'enquête Transport comme un échantillon tiré par un plan de sondage unique.

Au second degré, les logements ont été sélectionnés dans deux bases de sondage : les logements recensés en 1999 (RP99) et les logements construits depuis 1999 (BSLN). Dans les deux cas, on tire dans l'intersection de la base de logements avec l'échantillon maître (et l'EMEX selon les régions).

Le tirage dans le RP99 s'est fait en deux phases. En première phase, l'échantillon est alloué entre les strates de tirage de l'EM proportionnellement à la taille des strates dans la base de sondage, puis réparti de façon fixe entre les unités primaires d'une même strate. Le sondage est auto pondéré.

En deuxième phase, les logements sont stratifiés selon leur localisation (ville-centre d'une unité urbaine, banlieue, espace rural) et le nombre de voitures du ménage de 1999. Ils ont été sélectionnés par sondage aléatoire simple dans les strates, avec des taux de sondage permettant de surreprésenter les ménages ayant plusieurs voitures et les logements situés en milieu rural.

Les logements neufs sont sélectionnés par sondage aléatoire simple dans les unités primaires, avec des allocations assurant un sondage auto pondéré. En deuxième phase, les logements sont stratifiés selon leur localisation, afin de surreprésenter le milieu rural.

Au total, l'échantillon national de logements, avec ses extensions régionales tirées dans l'EMEX, se présente comme un sondage à deux degrés, avec un premier degré aréolaire et un deuxième degré en deux phases : sondage auto pondéré de logements dans les unités primaires en première phase, sondage stratifié avec échantillonnage aléatoire simple dans les strates en phase 2. Enfin la non-réponse ajoute une troisième phase, modélisée comme un sondage poissonnien.

L'extension locale

L'extension locale de Loire-Atlantique a été tirée dans la base de logements du recensement de 1999 et dans la base Sitadel des logements construits depuis, ces deux tirages étant indépendants. C'est un sondage direct à un degré stratifié, avec sondage aléatoire simple dans les strates pour les logements sélectionnés dans le recensement de 1999, tirage systématique dans un fichier ordonné à l'intérieur des strates pour les logements neufs. On estimera la variance comme celle d'un sondage aléatoire simple dans les strates dans les deux cas.

La non réponse ajoute une deuxième phase poissonnienne au sondage.

Individus et véhicules en visite 1

En visite 1, tous les individus du ménage et tous ses véhicules font l'objet d'une description exhaustive. Il s'agit des variables présentes dans les tables de diffusion TCM_INDIVIDU et QR_véhicule. Ces deux niveaux seront assimilés aux unités secondaires d'un sondage en grappes, la grappe étant le logement.

Le plan de sondage des individus kish de visite 1

Dans les ménages comprenant plus de deux enfants scolarisés de 6 ans ou plus, un enfant seulement dans ce champ était généralement sélectionné par sondage aléatoire simple. Tous les autres membres du ménage étaient interrogés. Les variables de ce niveau sont diffusées dans les tables Q_INDIVIDU et Q_IND_LIEU_TEG. Dans l'échantillon national, il s'agit du troisième degré de sondage, du second degré de tirage dans l'extension, comportant une strate exhaustive à l'intérieur du ménage.

Le plan de sondage des véhicules

Dans les ménages disposant de plusieurs voitures (respectivement de plusieurs vélos), deux voitures (respectivement deux vélos) au maximum étaient sélectionnés : par sondage aléatoire simple pour les voitures, par sondage à probabilités inégales pour les vélos. Il s'agit des informations contenues dans les tables QF_VOITVUL et QF_VELO. On les traitera comme deux échantillons distincts sélectionnés par un sondage à trois degrés dans l'échantillon national, par un sondage à deux degrés dans l'extension.

L'échantillon des individus kish de visite 2

Un individu unique a été sélectionné à l'intérieur de chaque ménage pour répondre au questionnaire de la deuxième visite. Le tirage a été fait avec des probabilités inégales, dépendant de la fréquence des déplacements des personnes, d'après leurs réponses au premier questionnaire. Tous les ménages ayant répondu en visite 1 n'ont pas accepté de répondre en visite 2. On traitera cet échantillon Kish comme le résultat d'un sondage à trois degrés (deux dans l'extension), le second degré (respectivement premier) ayant une phase de non-réponse incluant à la fois la non-réponse en visite 1 et en visite 2. Les données sont incluses dans la table de diffusion K_MOBILITE.

Voyages et déplacements des individus kish de visite 2

Les individus Kish sélectionnés en fin de première visite ont rempli un carnet de déplacements pendant la semaine de référence entre les deux entrevues avec l'enquêteur. Un jour de la semaine a été sélectionné pour décrire l'intégralité des déplacements quotidiens de la personne. Ce sont les informations contenues dans la table K_DEPLOC.

Les voyages non quotidiens, réalisés pendant les 4 semaines précédant l'enquête, sont également échantillonnés lorsque leur fréquence dépasse 10. Les effets de mémoire, en omettant les événements anciens au profit d'un passé plus récent, introduisent par ailleurs une sélection aléatoire par l'individu interrogé, équivalente à un phénomène de non-réponse. Il s'agit des informations contenues dans les tables K_VOYAGE et K_VOY_DEPDET.

LE REDRESSEMENT PAR CALAGE EN RÉGION PAYS DE LA LOIRE

Les échantillons de logements, national et local, ont été fusionnés par partage des poids sur le champ géographique de l'extension.

L'échantillon fusionné de ménages avec tous leurs individus, a été redressé par calage simultané sur les totaux dans la population, régionale et locale, de variables auxiliaires connues par ailleurs. La non-réponse totale, concernant seulement le niveau ménage, a été redressée directement par calage. On notera \mathbf{X}_M la matrice (n,q) des variables de calage utilisées au niveau ménage-individus.

L'échantillon des individus Kish de visite 1 a été calé sur des totaux de population, séparément du calage des ménages. On notera \mathbf{X}_I la matrice des variables de calage utilisées à ce niveau.

L'échantillon des individus Kish de visite 2 a été calé sur des totaux de population, séparément du calage des ménages. On notera \mathbf{X}_K la matrice des variables de calage utilisées à ce niveau.

L'échantillon de véhicules (voitures et vélos) sélectionné en visite 1 (tables QF) a été calé sur le parc de véhicules estimé à partir du recensement des véhicules du ménage dans le questionnaire principal, pondéré par le poids de calage du ménage répondant en première visite. On notera \mathbf{X}_V la matrice des variables de calage utilisées à ce niveau.

LES ÉTAPES DE L'ESTIMATION DE LA VARIANCE

Linéarisation de la variable estimée

Le programme joint propose d'estimer la variance de trois types de paramètres :

- Le total d'une variable numérique ou l'effectif d'une modalité dans la population de référence
- La moyenne d'une variable numérique ou la proportion d'une modalité dans la population de référence
- Le rapport de deux variables numériques dans la population de référence.

La population de référence peut être la population totale des Pays de la Loire ou un sous-ensemble de cette population, défini par un critère quelconque.

L'estimation de l'effectif de la population possédant la modalité j d'une variable catégorielle s'analyse comme celle du total de la variable numérique binaire, indicatrice d'appartenance à la catégorie j . De même, l'estimation d'une proportion s'identifie à celle de la moyenne de l'indicatrice d'appartenance à la modalité concernée.

La moyenne dans une sous-population (ou domaine), et un ratio de deux variables d'intérêt, ne sont pas des formes linéaires des observations de l'échantillon. Le paramètre à estimer est donc d'abord transformé par la technique de linéarisation d'une fonction complexe, selon la méthode proposée par J.C. Deville, basée sur le développement limité de Taylor à l'ordre 1.

Soit : $\vartheta = f(Y, X)$ le paramètre à estimer. ϑ est une fonction des totaux dans la population de référence U des variables X et Y : $X = \sum_{\lambda \in U} X_\lambda$ et $Y = \sum_{\lambda \in U} Y_\lambda$.

L'estimateur du paramètre ϑ est approximé par l'expression linéaire :

$$\hat{\vartheta} = f(\hat{Y}, \hat{X}) \approx \vartheta + (\hat{Y} - Y) \frac{\partial f}{\partial Y} + (\hat{X} - X) \frac{\partial f}{\partial X}$$

Dans le cas d'un ratio $\vartheta = \frac{Y_1}{Y_2}$, il vient :

$$\hat{\vartheta} = \frac{\hat{Y}_1}{\hat{Y}_2} \approx \vartheta + (\hat{Y}_1 - Y_1) \frac{1}{Y_2} + (\hat{Y}_2 - Y_2) \left(-\frac{Y_1}{[Y_2]^2} \right) = \vartheta + \frac{1}{Y_2} (\hat{Y}_1 - \vartheta \hat{Y}_2)$$

Son erreur quadratique moyenne est donc approchée par celle de la forme linéaire :

$$EQM(\hat{\vartheta}) = E(\hat{\vartheta} - \vartheta)^2 \approx V \left[\frac{1}{Y_2} (\hat{Y}_1 - \vartheta \hat{Y}_2) \right] = \frac{1}{(Y_2)^2} V(\hat{Y}_1 - \vartheta \hat{Y}_2) = \frac{1}{(Y_2)^2} V \left(\sum_{\lambda \in s} \frac{y_{1\lambda} - \vartheta y_{2\lambda}}{\pi_\lambda} \right)$$

qu'on estime par :

$$E\hat{Q}M(\hat{\vartheta}) = \frac{1}{(\hat{Y}_2)^2} \hat{V} \left(\sum_{\lambda \in s} \frac{y_{1\lambda} - \hat{\vartheta} y_{2\lambda}}{\pi_\lambda} \right)$$

Dans la pratique, on forme la variable auxiliaire : $z_\lambda = \frac{y_{1\lambda} - \hat{\vartheta} y_{2\lambda}}{\hat{Y}_2}$ pour chaque unité λ de l'échantillon. Dans le cas de l'estimation d'une moyenne dans une sous-population U_d , la variable $y_{2\lambda}$ est l'indicatrice d'appartenance de l'unité λ au domaine U_d , et \hat{Y}_2 est l'estimateur de la taille du domaine.

La variance du paramètre ϑ est alors estimée comme celle de l'estimateur du total :

$$\hat{V}(\hat{\vartheta}) = \hat{V}(\hat{Z}) = \hat{V}\left(\sum_{\lambda \in s} \frac{z_\lambda}{\pi_\lambda}\right) .$$

Prise en compte du calage

Soient :

S_{nat} l'échantillon national sélectionné sous le plan \mathbf{p}_1 , de taille n_1

S_d l'extension locale en Loire-Atlantique sélectionnée sous le plan \mathbf{p}_2 , de taille n_2

S l'échantillon fusionné : $S = S_{\text{nat}} \cup S_d$

$sr_{\text{nat}}, sr_d, sr = sr_{\text{nat}} \cup sr_d$ les échantillons répondants respectifs, de tailles $r_1, r_2, r = r_1 + r_2$

π_l la probabilité d'inclusion de l'unité l dans le sous-échantillon S_{nat} ou S_d auquel il appartient

p_l la probabilité conditionnelle de réponse de l'unité l

$d_l, d_l^*, \tilde{d}_l^*, w_l$ le poids de l'unité l , respectivement au tirage, après fusion des échantillons S_{nat} et S_d , après correction de la non-réponse, après calage

\mathbf{X} (r, q) la matrice des variables de calage dans l'échantillon répondant

\mathbf{Z} (r) le vecteur de la variable d'intérêt linéarisée dans l'échantillon répondant

L'échantillon fusionné a été redressé par calage sur les totaux connus des variables auxiliaires \mathcal{X} . L'estimateur de Z redressé par calage est asymptotiquement équivalent à l'estimateur redressé par régression sur les variables \mathcal{X} :

$$\hat{Z}_{\text{cal}} = \sum_{l \in sr} w_l z_l \approx \hat{Z}_{\text{reg}} = \hat{\mathbf{b}}' \mathbf{X} + \sum_{l \in sr} \tilde{d}_l^* u_l \quad (1)$$

$$\text{où : } \begin{cases} u_l = z_l - \hat{\mathbf{b}}' x_l \text{ est le résidu de la régression de } \mathbf{Z} \text{ sur } \mathbf{X} \text{ pour l'unité } l \\ \hat{\mathbf{b}} = [\mathbf{X}' \mathbf{D} \mathbf{X}]^{-1} \mathbf{X}' \mathbf{D} \mathbf{Z} \text{ est le vecteur des coefficients de régression de } Z \text{ sur } X \\ \mathbf{D} = \text{diag}(\tilde{d}_l^*) \end{cases}$$

Sa variance peut être estimée par celle des résidus de la régression de Z sur \mathcal{X} , pondérée par les poids dans l'échantillon fusionné, spécifiés en entrée de Calmar, et corrigés de la non-réponse :

$$V(\hat{\vartheta}_{\text{cal}}) \approx V(\hat{Z}_{\text{reg}}) = V\left(\sum_{l \in sr} \tilde{d}_l^* u_l\right)$$

$$\hat{V}(\hat{\vartheta}_{cal}) = \hat{V}(\hat{Z}_{reg}) = \hat{V}\left(\sum_{l \in sr} \tilde{d}_l^* u_l\right) = \hat{V}(\hat{U}_{\tilde{d}^*}) \quad (2)$$

En pratique, on a estimé par régression logistique les probabilités de réponse des ménages, respectivement en visite 1 conditionnellement à l'échantillon sélectionné (\hat{p}_{1l}), et en visite 2 conditionnellement à l'échantillon répondant en visite 1 ($\hat{p}_{2/1,l}$). Le programme effectue les régressions avec les pondérations :

$$\tilde{d}_l^* = d_l^* \times \frac{1}{\hat{p}_{1l}} \text{ pour les variables relevant du questionnaire de visite 1}$$

$$\tilde{d}_l^* = d_l^* \times \frac{1}{\hat{p}_{1l} \times \hat{p}_{2/1,l}} \text{ pour les variables relevant du questionnaire de visite 2.}$$

Prise en compte du partage des poids

Le π -estimateur du total des résidus de régression, calculé avec les poids fusionnés, se décompose de la façon suivante :

$$\hat{U}_{\tilde{d}^*} = \sum_{l \in sr} \tilde{d}_l^* u_l = \sum_{j=1}^2 \sum_{l \in sr_j} (\tilde{d}_l \lambda_j) u_l = \sum_{j=1}^2 \sum_{l \in sr_j} \tilde{d}_l (\lambda_j u_l) = \sum_{l \in sr} \tilde{d}_l t_l \quad (3)$$

où : λ_j est le coefficient de partage des poids appliqué au sous-échantillon S_j . Il vaut 1 hors du champ géographique de l'extension locale

$$t_l = \sum_{j=1}^2 1_{l \in S_j} \lambda_j u_l \text{ est le résidu transformé par le partage des poids}$$

$$\tilde{d}_l = d_l \times \frac{1}{\hat{p}_l} = \frac{1}{\pi_l \hat{p}_l} \text{ est le poids initial de tirage corrigé de l'inverse de la probabilité de réponse.}$$

Le programme calcule la variable t en multipliant, sur chaque unité, le résidu de régression par le coefficient de partage des poids du sous-échantillon auquel appartient l'unité. La variance du paramètre $\hat{\vartheta}$ est donc finalement approchée par celle du total estimé de la variable transformée t , pondérée par le poids initial de sondage corrigé de la non-réponse des ménages.

$$V(\hat{\vartheta}_{cal}) \approx V(\hat{T}_\pi) = V\left(\sum_{l \in sr} \frac{t_l}{\pi_l p_l}\right)$$

Prise en compte de la non-réponse

Une partie des ménages sélectionnés n'a pas répondu. Le comportement de réponse est assimilé à une phase supplémentaire, aléatoire, de l'enquête. On le représente comme le tirage, dans l'échantillon initial, de l'échantillon de répondants avec des probabilités de réponse inconnues p_l . L'échantillon de répondants est donc modélisé comme le résultat d'une procédure de sélection en deux phases : la première constituée de l'échantillonnage initial et la seconde de la sélection des répondants parmi les unités échantillonnées.

La variance, sous un plan en deux phases, se décompose en deux éléments :

$$V(\hat{\vartheta}_{cal}) \approx \underbrace{V \left[\underbrace{E(\hat{T}_{\tilde{d}} / \mathbf{s})}_{\substack{\text{Espérance conditionnelle} \\ \text{à l'échantillon initial}}} \right]}_{\substack{\text{Variance due à l'échantillonnage} \\ \text{initial}}} + \underbrace{E \left[\underbrace{V(\hat{T}_{\tilde{d}} / \mathbf{s})}_{\substack{\text{Variance conditionnelle} \\ \text{à l'échantillon initial}}} \right]}_{\substack{\text{Espérance dans la population de référence} \\ \text{Variance due à la non-réponse} \\ \text{des ménages}}} = V_1 + V_2 \quad (4)$$

Le premier élément V_1 représente la variance due à l'échantillonnage initial. Son expression dépend du plan de sondage adopté.

Le second élément exprime la variance due à la sélection des répondants à l'intérieur de l'échantillon initial. Son expression dépend du modèle retenu de comportement de réponse. On a ici assimilé la phase de réponse à un sondage poissonnien, à probabilités inégales et avec indépendance des comportements de réponse entre deux ménages distincts.

$$V_2 = E \left(V \left(\sum_{l \in sr} \frac{t_l}{\pi_l p_l} \right) / (s_{nat} \cup s_d) \right) = E \left(V \left(\sum_{l \in sr} \frac{\check{t}_l}{p_l} \right) \right) = E \left(\sum_{l \in sr} \sum_{\lambda \in sr} \delta_{l\lambda} \frac{\check{t}_l}{p_l} \frac{\check{t}_\lambda}{p_\lambda} \right)$$

avec : $\check{t}_l = \frac{t_l}{\pi_l}$

$$\delta_{l\lambda} = p_{l\lambda} - p_l p_\lambda = \begin{cases} p_l(1-p_l) & \Leftrightarrow l = \lambda \\ 0 & \Leftrightarrow l \neq \lambda \end{cases}$$

Sous l'hypothèse d'un sondage de Poisson, la variance due à la non-réponse est estimée par :

$$\hat{V}_2 = \sum_{l \in sr} \sum_{\lambda \in sr} \frac{\delta_{l\lambda}}{p_{l\lambda}} \frac{\check{t}_l}{p_l} \frac{\check{t}_\lambda}{p_\lambda} = \sum_{l \in sr} \frac{(1-p_l)}{(p_l)^2} \left(\frac{t_l}{\pi_l} \right)^2$$

π_l est la probabilité initiale d'inclusion de l'unité l dans l'échantillon, respectivement national ou local.

Variance due à l'échantillonnage

On détaille ici l'élément V_1 de l'équation (4). Le π -estimateur du total des résidus de régression, transformés par le partage des poids, se décompose sur les deux échantillons S_{nat} et S_d :

$$\hat{T}_{\pi^*} = \sum_{l \in sr} \frac{t_l}{\pi_l p_l} = \sum_{l \in sr_{nat}} \frac{t_l}{\pi_{1l} p_l} + \sum_{m \in sr_d} \frac{t_m}{\pi_{2m} p_m} = \hat{T}_{p_1} + \hat{T}_{p_2} \quad (5)$$

où : $t_l = \sum_{j=1}^2 1_{l \in s_j} \lambda_j u_l$,

λ_j est le coefficient de partage des poids appliqué au sous-échantillon S_j .

π_{1l} est la probabilité d'inclusion de l'unité l dans l'échantillon national

π_{2m} est la probabilité d'inclusion de l'unité m dans l'extension

\hat{T}_π est la somme de deux π -estimateurs de la variable transformée t_l dans les deux échantillons, national et local. Ces deux échantillons étant quasi-indépendants, la variance de la somme est proche de la somme des variances de chacun des termes. On estimera la variance due à l'échantillonnage comme la somme de la variance dans l'échantillon national sous le plan de sondage \mathbf{p}_1 et de la variance dans l'extension sous le plan \mathbf{p}_2 .

$$V\left[E\left(\hat{T}_\pi^* / \mathbf{s}\right)\right] = V\left[E\left(\hat{T}_{\pi_1} / \mathbf{s}_{\text{nat}}\right)\right] + V\left[E\left(\hat{T}_{\pi_2} / \mathbf{s}_d\right)\right] \quad (6)$$

Variance due à l'échantillonnage national

La difficulté d'accès à tous les paramètres nécessaires à l'usage du logiciel POULPE a conduit à simplifier le modèle d'échantillonnage par une hypothèse de tirage avec remise des unités primaires de l'échantillon-maître au premier degré. Sous cette hypothèse, on dispose d'un estimateur sans biais de la variance ne reposant que sur les probabilités d'inclusion simples et ne nécessitant que l'estimation de la variance due au premier degré de sondage.

$$\hat{T}_{\pi_1} = \sum_{l \in s_{\text{nat}}} \frac{t_l}{\pi_l p_l} = \sum_{i \in EM} \frac{1}{\pi_i} \sum_{l \in s_{i'}} \frac{t_l}{\pi_{l/i} p_l}$$

$$V_{1\text{nat}} = V\left(E\left(\sum_{i \in EM} \sum_{l \in U_i} \frac{t_l}{\pi_i \pi_{l/i} p_l}\right) / (s_{\text{nat}})\right) = V\left(\sum_{i \in EM} \sum_{l \in U_i} \frac{t_l}{\pi_i \pi_{l/i}}\right) = \sum_h \frac{1}{m_h} \sum_{i \in U} A_i \left(\frac{T_i}{A_i} - T_h\right)^2 + \frac{1}{m_h} \sum_{i \in U} \frac{V_i}{A_i}$$

i est l'indice d'une unité primaire de l'échantillon-maître (ou EMEX), h l'indice de strate de l'EM/EMEX

m_h est le nombre d'unités primaires dans l'échantillon-maître et l'EMEX, dans la strate h

A_i est la probabilité de sélection dans l'échantillon-maître de l'UP i au $j^{\text{ième}}$ tirage

V_i est la variance entre ménages à l'intérieur de l'UP i

$T_i = \sum_{l \in U_i} t_l$ est le total dans l'UP i et T_h le total dans la population de référence de la strate h .

La variance est estimée sans biais par :

$$\hat{V}_{1\text{nat}} = \sum_h \frac{1}{m_h(m_h - 1)} \sum_{i \in EM} \left(\frac{\hat{T}_i}{A_i} - \hat{T}_h\right)^2 \quad (7)$$

où : $\hat{T}_i = \sum_{l \in s_i} \frac{t_l}{\pi_{l/i} p_l}$ $\hat{T}_h = \sum_h \sum_{i \in s_h} \frac{\hat{T}_i}{\pi_i}$ $A_i \approx \frac{\pi_i}{m}$

T_i et T_h sont estimés dans l'échantillon de ménages répondants, par conséquent avec des pondérations initiales corrigées de la non-réponse.

Le programme utilise la procédure SURVEYMEANS du logiciel SAS, dont les estimations sont basées sur l'hypothèse de tirage avec remise des unités primaires en cas de sondage multi-degrés et avec des probabilités de tirage inégales. La formule (7) est appliquée à l'estimation de la variance aux niveaux ménages et aux niveaux individus ou véhicules sélectionnés au troisième degré en visite 1 ou 2.

Dans ce dernier cas, la pondération spécifiée dans la procédure SURVEYMEANS est le poids total, corrigé de la non-réponse, de l'unité Kish (unités des tables Q_INDIVIDU, QF_véhicule ou K_MOBILITE). \hat{T}_i et \hat{T}_h estiment des totaux d'individus (de véhicules) et non plus de ménages.

Variance due à l'échantillonnage de l'extension locale

On explicite ici le deuxième terme de l'équation (6). Le tirage des logements de l'extension étant à un seul degré, avec sondage aléatoire simple dans les strates, la variance des variables observées au niveau des ménages ou des individus du TCM peut être estimée directement à partir de l'estimateur général Horvitz-Thomson de la variance. Au second degré, les individus Kish étant uniques dans chaque ménage, on ne peut pas estimer la variance due au second degré de tirage. On a donc, pour les informations des tables Q_INDIVIDU, QF_véhicule, K_MOBILITE, fait l'hypothèse d'un tirage des ménages avec remise.

Niveau ménages et individus du TCM

La variance a pour expression générale la formule Horvitz-Thomson :

$$V_{1s_d} = \left(\sum_h \sum_{l \in U} \sum_{\lambda \in U} \Delta_{l\lambda} \frac{t_l}{\pi_l} \frac{t_\lambda}{\pi_\lambda} \right)$$

et est estimée sans biais par :

$$\hat{V}_{1s_d} = \left(\sum_h \sum_{l \in sr_d} \sum_{\lambda \in sr_d} \frac{\Delta_{l\lambda}}{\pi_{l\lambda} p_{l\lambda}} \frac{t_l}{\pi_l} \frac{t_\lambda}{\pi_\lambda} \right)$$

h est ici la strate de tirage de l'extension.

a) si $l = \lambda$

$$\begin{aligned} \pi_{l\lambda} &= \pi_l = f_h \\ \Delta_{l\lambda} &= \pi_l(1 - \pi_l) \\ \frac{\Delta_{l\lambda}}{\pi_{l\lambda}\pi_l\pi_\lambda} &= \frac{\pi_l(1 - \pi_l)}{(\pi_l)^3} = \frac{(1 - \pi_l)}{(\pi_l)^2} \\ \hat{V}_h &= \sum_{l \in sr_h} \frac{(1 - \pi_l)(t_l)^2}{(\pi_l)^2 p_l} = \sum_{l \in sr_h} \frac{(1 - f_h)(t_l)^2}{(f_h)^2 p_l} \end{aligned}$$

b) si $l \neq \lambda$

$$\begin{aligned} \Delta_{l\lambda} &= \pi_{l\lambda} - \pi_l\pi_\lambda \\ \pi_{l\lambda} &= \frac{n_h(n_h - 1)}{N_h(N_h - 1)} \\ \frac{\Delta_{l\lambda}}{\pi_{l\lambda}\pi_l\pi_\lambda} &= \frac{1}{\pi_l\pi_\lambda} - \frac{1}{\pi_{l\lambda}} = \frac{f_h - 1}{(f_h)^2(n_h - 1)} \end{aligned}$$

D'où :

$$\hat{V}_{1s_d} = \sum_h \frac{(1 - f_h)}{(f_h)^2} \left[\sum_{l \in sr_h} p_l (\tilde{t}_l)^2 + \frac{1}{(n_h - 1)} \left(\sum_{l \in sr_h} (\tilde{t}_l)^2 - (r_h \bar{\tilde{t}}_h)^2 \right) \right]$$

avec :

$$\tilde{t}_l = \frac{t_l}{p_l}$$

$$\bar{t}_h = \frac{1}{r_h} \sum_{l \in sr_h} t_l$$

r_h est le nombre de ménages répondants dans la strate h

n_h est le nombre de logements sélectionnés dans la strate h

f_h est le taux de sondage dans la strate h

Niveau individus Kish ou véhicules sélectionnés

Sous l'hypothèse d'un tirage des logements avec remise, la variance est estimée sans biais par :

$$\hat{V}_{1s_d} = \sum_h \frac{1}{n_h(n_h-1)} \sum_{l \in sd_h} \left(\frac{\hat{T}_l}{A_l} - \hat{T}_h \right)^2 \quad (8)$$

Avec :

$$\hat{T}_l = \sum_{k \in sd_l} \frac{t_k}{\pi_{k/l}}$$

$$\hat{T}_h = \sum_{l \in sd_h} \frac{\hat{T}_l}{\pi_l}$$

n_h est le nombre de logements sélectionnés dans la strate h .

$$A_l \approx \frac{\pi_l}{n_h}$$

L'expression (8) fait appel à l'échantillon sélectionné complet de logements. Elle doit être estimée à partir de l'échantillon de ménages répondants, soit par :

$$\hat{V}_{1s_d} = \sum_h \frac{1}{n_h(n_h-1)} \sum_{l \in sr_d_h} \frac{1}{p_l} \left(\frac{\hat{T}_l}{A_l} - \hat{T}_h \right)^2 \quad (9)$$

où : $\hat{T}_h = \sum_{l \in sd_h} \frac{\hat{T}_l}{\pi_l p_l}$

LES PARAMÈTRES DU PROGRAMME D'ESTIMATION

Le programme d'estimation est une macro SAS paramétrée : ENTD. Les paramètres en sont les suivants.

- **ECHANT**

Nom de la table contenant les données de l'enquête, dont on souhaite estimer la variance. Cette table doit contenir une observation par unité répondante avec les variables suivantes :

- l'identifiant de l'unité observée et celui du logement auquel elle se rattache (variable IDENT_LOG)
- la pondération finale de l'unité après redressement. Il s'agit du poids de calage fourni par le PISM dans la région Pays de la Loire.
- toutes les variables spécifiées dans les paramètres VARINTNUM, VARINTCAT et RATIOVAR ci-dessous, variables d'intérêt de l'enquête dont on estime la variance.

Les identifiants doivent être ceux figurant dans les tables initiales de l'enquête, tels qu'ils ont été nommés dans les fichiers de l'INRETS. Ce sont donc les variables : IDENT_LOG, IDENT_MEN, IDENT_IND, IDENT_NUMVEH, IDENT_DEP, IDENT_DEPLD, IDENT_VOY.

- **VARAUXIL**

Nom d'une table de données auxiliaires, extraite de la table des marges de calage utilisée pour le redressement, contenant la liste des noms des variables de calage et leur nombre de modalités (variables VAR et N de la table de marges). La table à spécifier dépend du niveau d'observation auquel se rattache la table indiquée dans le paramètre ECHANT. L'ensemble des tables de calage sont livrées avec le programme ENTD. Le tableau synoptique présente la correspondance entre le type d'unité observée et le nom des tables de calage s'y rapportant.

- **ECHAUXIL**

Nom d'une table de données auxiliaires contenant une observation par unité enquêtée et la valeur des variables de calage utilisées dans le redressement, ainsi que les pondérations initiales intermédiaires. L'ensemble des tables de données auxiliaires sont livrées avec le programme ENTD. On trouvera dans le tableau synoptique ci-dessous le nom de la table à indiquer en fonction du niveau d'observation des unités de la table spécifiée dans le paramètre ECHANT.

- **POIDSCAL**

Nom de la variable contenant le poids de calage de l'unité répondante. Il s'agit de la pondération finale résultant du redressement de l'échantillon en région Pays de la Loire. Cette variable doit figurer dans la table spécifiée dans le paramètre ECHANT.

- **VISITE**

1 ou **2**, selon que la table spécifiée dans le paramètre ECHANT se rapporte au questionnaire de la première ou de la seconde visite.

- **NIVEAU**

Il s'agit du niveau d'observation auquel se rapportent les variables d'intérêt dont on estime la variance. Ce paramètre doit être renseigné par les noms initiaux des tables de l'enquête créées par l'INRETS. Il peut prendre les valeurs :

- Q_TCM_MENAGE, Q_TCM_INDIVIDU, Q_MENAGE, Q_INDIVIDU, Q_IND_IEUTEG
- QR_VOITVUL, QR_VELO, QF_VOITVUL, QF_VELO
- K_MOBILITE, K_VOYAGE, K_DEPLOC, K_VOY_DEPDET

- **REG**

OUI si l'on souhaite faire imprimer le résultat de la régression des variables d'intérêt sur les variables de calage. Par défaut, REG=non.

- **VARINTNUM**

Noms des variables d'intérêt **numériques** de l'enquête dont on souhaite estimer la variance. S'il y en a plusieurs, les noms doivent être séparés par des blancs.

- **VARINTCAT**

Noms des variables d'intérêt **catégorielles** de l'enquête dont on souhaite estimer la variance. S'il y en a plusieurs, les noms doivent être séparés par des blancs.

- **RATIOVAR**

Noms de deux variables d'intérêt numériques dont on souhaite estimer le ratio et sa variance. La première variable sera prise comme numérateur et la seconde comme dénominateur. Les noms doivent être séparés par un blanc.

- **DOMAINE**

Nom d'une variable binaire valant 1 lorsque l'unité appartient à la sous-population de référence, 0 sinon. Ce paramètre est facultatif. S'il n'est pas renseigné, le programme estime la variance des paramètres sur l'ensemble de la région Pays de la Loire. Sinon, les paramètres sont estimés dans la sous-population spécifiée dans DOMAINE.

- **SORTIE**

Nom d'une table SAS qui contiendra les résultats produits par la macro ENTD. Cette table peut être permanente ou temporaire. Par défaut, le programme crée en sortie une table WORK.__VARIANCE contenant les résultats de l'estimation.

Les paramètres REG, DOMAINE et SORTIE sont facultatifs. L'un seulement des paramètres VARINTNUM, VARINTCAT et RATIOVAR est obligatoire. Tous les autres paramètres doivent être renseignés.

Les variables d'intérêt spécifiées dans les paramètres VARINTNUM, VARINTCAT et RATIOVAR doivent être renseignées sur toutes les unités de la table ECHANT. Dans le cas d'une variable non renseignée dans le fichier pour les unités non concernées par la question, il faut préalablement codifier par une modalité spécifique le caractère « non concerné ».

Les tables de données auxiliaires appelées par la macro sont livrées avec le programme. Elles doivent être implantées dans le même répertoire : selon les cas, la macro lit le ddname alloué pendant la session de l'utilisateur à ce répertoire pour y rechercher une seconde table de données auxiliaires.

Le tableau suivant présente les noms des tables de données auxiliaires à utiliser selon le niveau d'observation des variables d'étude.

Visite	Niveau d'observation	Table ECHAUXIL	Table VARAUXIL
1	Q_TCM_MENAGE		
1	Q_TCM_INDIVIDU		
1	Q_MENAGE	V1MENAGE	MARGES_V1MENIND
1	QR_VOITVUL		
1	QR_VELO		
1	Q_INDIVIDU	V1QINDIV	MARGES_V1QIND
1	Q_IND_LIEUTEG		
2	Q_TCM_MENAGE		
2	Q_TCM_INDIVIDU	V2MENAGE	MARGES_V2MENIND
2	Q_MENAGE		
2	K_MOBILITE		
2	K_DEPLOC	V2KISH	MARGES_V2KISH
2	K_VOYAGE		
2	K_VOY_DEPDET	V2VOYAGE	

Exemples d'utilisation de la macro ENTD

```
/* allocation des répertoires contenant respectivement les fichiers d'enquête et
les tables de données auxiliaires */
```

```
libname fin "d:\iz91eg\mes documents\transport2007\calage pays de la loire\fichiers pondérés";
libname va "d:\iz91eg\mes documents\transport2007\variance";
```

```
options mprint;
```

```
/* niveau ménage visite 1 - champ Pays de la Loire */
```

```
%ENTD(echant=fin.q_menage,
      echauxil=va.v1menage,
      varauxil=va.marges_v1menind,
      poidsca1=pondv1,
      niveau=q_menage,
      visite=1,
      reg=non,
      varintnum=v1_jnbveh v1_jnbveloadt v1_jnbcyclo v1_jnbccvul,
      varintcat=v1_logestatdispo v1_logestatvelo,
      ratiovar=v1_jnbveh v1_jnbveloadt,
      domaine= ,
      sortie=varmen_pl);
```

```
/* niveau individus sélectionnés en visite 1 - champ : Loire Atlantique */
```

```
data qteg;
  set fin.q_ind_lieu_teg (where=(rg='52'));
  array v v1_ntravdistetudgare v1_btravparkfraisc;
  do over v;
    if v=. then v=0;
  end;
  d44=(dep='44');
run;
```

```
%ENTD(echant=qteg,
      echauxil=va.v1qindiv,
      varauxil=va.marges_v1qind,
      poidsca1=poids_qind1,
      niveau=q_ind_lieu_teg,
      visite=1,
      reg=non,
      varintnum=v1_btravdistetudgare v1_btravparkfraisc,
      domaine=d44,
      sortie=varteg_pl);
```