



Vers une catégorisation des déplacements domicile - travail des frontaliers luxembourgeois en 2003.

Annexes

Philippe GERBER, Michel RAMM - CEPS/INSTEAD

Annexe 1 - L'imputation : aperçu méthodologique, mise en oeuvre et résultats

Dans un questionnaire, les enquêtés ne répondent pas forcément à toutes les questions. Les données, produites en fonction des réponses apportées aux questions posées, peuvent donc être incomplètes. Certaines réponses manquantes doivent alors faire l'objet d'une estimation pour que le biais dû à la non réponse soit le plus faible possible. Cette estimation peut se baser : soit sur des données provenant d'autres sources, soit sur des règles logiques, soit encore sur la base d'un modèle statistique d'estimation. Ce dernier processus est appelé imputation.

Dans le cadre des « Enquêtes Frontaliers » 2002 et 2003, portant entre autres sur la mobilité des déplacements domicile-travail des travailleurs frontaliers au Grand-Duché de Luxembourg, le modèle choisi pour l'imputation est un modèle multivarié de régression séquentielle, implémenté dans le logiciel IVE (Imputation and Variance Estimation) de l'Université de Michigan (<http://www.isr.umich.edu/src/smp/ive/>). Son algorithme peut être décrit de la façon suivante :

- soit U un ensemble de variables explicatives sans données manquantes,

- soit $Y_1, Y_2, \dots, Y_k$, un ensemble de variables dépendantes ordonnées selon le taux croissant de données manquantes.

La séquence des imputations est déterminée par la factorisation qui suit : $[Y_1/U], [Y_2/U, Y_1], [Y_k/U, Y_1, \dots, Y_{k-1}]$,

où $[Y/U]$ est la distribution conditionnelle jointe de Y sachant U. Autrement dit, après chaque itération i, la variable Y_i qui vient d'être imputée s'ajoute à l'ensemble des variables explicatives.

Pour les besoins de ce document, sont imputées les deux variables « temps de trajet » et « distance de trajet », dès lors qu'elles ne sont pas renseignées sur au moins une portion du trajet domicile - travail. Du fait du type de ces variables à expliquer (elles sont continues), une régression linéaire généralisée a été pratiquée pour les deux vagues d'enquête de 2002 et 2003. Les variables explicatives utilisées dans le modèle d'imputation sont : le pays d'installation, le sexe, les tranches d'âge, la catégorie socioprofessionnelle, le niveau d'éducation et le moyen de transport.

Les bornes inférieures des futures valeurs imputées des deux variables à expliquer sont construites à partir de la somme partielle des données renseignées sur les portions de trajet déclarées. Les bornes supérieures de ces mêmes variables ont été fixées en référence aux distances et vitesses maximales plausibles.

Pour éviter une discordance entre l'imputation des variables « temps » et « distance », nous choisissons celle qui est la plus renseignée pour l'imputer en premier lieu. La vitesse est imputée par après. Ainsi, pour la première vague de l'enquête (2002), la distance (4 378 répondants) et la vitesse sont imputées d'abord pour en déduire le temps ensuite (cf. *tableau A1*).

Ainsi, pour l'imputation de la distance de la vague un, les bornes inférieure et supérieure sont déterminées par la valeur du minimum et du maximum déclarés par les répondants, c'est-à-dire respectivement 2 et 360 km. Pour la vitesse, calculée a posteriori en fonction des déclarations, les valeurs minimales et maximales à imputer sont de 8 et de 130 km/h (vitesse maximale autorisée). La même procédure et la même méthode sont appliquées pour la vague deux, en fonction cette fois du temps de déplacement, variable la plus renseignée. Mais avant toute imputation de cette dernière vague, les données manquantes de la distance et du temps sont remplacées par celles de la vague un, si elles sont renseignées et si la personne habite toujours la même commune. A l'issue de cette procédure appelée « editing »

(cf. *tableau A1*), le temps de trajet devient la variable la plus renseignée. Par conséquent, le temps de trajet et la vitesse sont imputés pour en déduire la distance.

Quelle que soit l'année, les résultats obtenus révèlent un profil de frontaliers bien particulier. En effet, les valeurs moyennes imputées des temps et distances de déplacement restent bien plus faibles que celles des répondants. Par exemple, en 2003, la distance moyenne de déplacement est de 43,4 km pour les répondants, alors qu'elle ne dépasse pas les 18 km par rapport à la moyenne des valeurs imputées des non-répondants. L'écart se creuse si l'on observe le temps moyen, qui passe de 42,7 minutes pour les répondants à 6,9 minutes en moyenne pour les autres.

En fonction de cette analyse, nous constatons que les profils des frontaliers n'ayant pas répondu à ces questions diffèrent de ceux qui y ont répondu. Il existe donc une sélectivité de la non-réponse où le domicile officiel (le questionnaire a été envoyé à l'adresse connue grâce au fichier de l'IGSS) ne correspond pas (ou plus) au domicile effectivement occupé (supposé *a posteriori*) du frontalier. En fait, la distance ou le temps de déplacement laisse supposer que le lieu d'habitation se situe au Luxembourg. Selon ces considérations, comme le lieu de départ habituel du déplacement n'a pas été spécifié au sein du questionnaire, nous écartons de l'analyse les questionnaires qui ne disposent pas de réponses déclarées de la part des frontaliers.

DIA Modou, GERBER Philippe, CEPS/
INSTEAD

A1 Imputation des variables temps, distance et vitesse des vagues une (2002) et deux (2003)

	Temps (mn)			Effectifs	Distance (km)			Effectifs	Vitesse (km/h)			Effectifs
	Minimum	Maximum	Moyenne		Minimum	Maximum	Moyenne		Minimum	Maximum	Moyenne	
Répondants 2002	4 369	1	580	44,0	4 378	2	360	40,2	4 348	8	600	63,3
Non-répondants 2002	227	1	174,5	10,2	218	2	176,5	7,4	248	8,0	126,4	64,1
Total 2002	4 596	0	580	41,8	4 596	2	360	40,2	4 596	8	600	63,4
Répondants 2003	5 942	2	190	42,7	5 894	1	264	43,4	5 945	4	225	62,3
Non-répondants 2003	204	2,0	18,4	6,9	233	0,7	170	17,9	261	18,5	110	61,6
dont édités	60	1	130	42,3	79	8	120	42,6				
Total 2003	6 206	1	190	41,6	6 206	0,7	264	43,4	6 206	4	225	62,3

Source : STATEC 2002-2003, CEPS/INSTEAD 2004 "Enquête Frontaliers" 2002 et 2003.
Champ : frontaliers ayant participé à l'enquête.

Annexe 2 - Correspondances entre les nomenclatures nationales de niveau de formation et classification CITE retenue

Correspondance CITE retenue

France

- Primaire	primaire ou secondaire inférieur
- BEPC	primaire ou secondaire inférieur
- CAP	secondaire supérieur
- BEP	secondaire supérieur
- Baccalauréat professionnel	secondaire supérieur
- Baccalauréat technique	secondaire supérieur
- Baccalauréat général	secondaire supérieur
- BTS/DUT	enseignement supérieur
- Autres Bac+2 (DEUG, ...)	enseignement supérieur
- Bac+3 (licence, ...)	enseignement supérieur
- Bac+4 (maîtrise, ...)	enseignement supérieur
- Bac+5 (DESS, DEA, ...)	enseignement supérieur
- Supérieur à Bac+5	enseignement supérieur

Belgique

- Primaire	primaire ou secondaire inférieur
- Secondaire inférieur général	primaire ou secondaire inférieur
- Secondaire inférieur technique	primaire ou secondaire inférieur
- Secondaire inférieur professionnel	primaire ou secondaire inférieur
- Secondaire supérieur général	secondaire supérieur
- Secondaire supérieur technique	secondaire supérieur
- Secondaire supérieur professionnel	secondaire supérieur
- Post-secondaire non supérieur (7 ^e année d'enseignement secondaire, enseignement secondaire professionnel complémentaire (4 ^e degré), chef d'entreprise des Classes moyennes)	secondaire supérieur
- Enseignement supérieur non universitaire de type court 1er cycle graduat, bac+3)	enseignement supérieur
- Université ou haute école de type long 1er cycle (candidature)	enseignement supérieur
- Université ou haute école de type long 2ème cycle (licence, ingénieur, docteur sans thèse,...)	enseignement supérieur
- Université ou haute école : diplôme complémentaire (DESS, DEC, DEA),	
- maîtrise, agrégation de l'enseignement secondaire supérieur,...)	enseignement supérieur
- Doctorat avec thèse	enseignement supérieur

Allemagne

- Grundschule	primaire ou secondaire inférieur
- Hauptschule	primaire ou secondaire inférieur
- Realschule	primaire ou secondaire inférieur
- Berufliche Lehre (Gesellenprüfung – Kaufmännischer Abschluss)	secondaire supérieur
- Fachoberschule	secondaire supérieur
- Gymnasium	secondaire supérieur
- Berufliche Lehre (Meisterprüfung)	secondaire supérieur
- Fachhochschule	enseignement supérieur
- Universitätsdiplom	enseignement supérieur
- Aufbaustudium, Dissertation, Habilitation	enseignement supérieur