

X^{èmes} Journées de Méthodologie Statistique de l'INSEE
Paris, 23-25 mars 2009

Impact du mode d'enquête sur les comportements de mobilité

Application à l'EMD de Lyon – 2006

Caroline Bayart, Patrick Bonnel
Laboratoire d'Economie des Transports, ENTPE



Caroline Bayart, Patrick Bonnel
Laboratoire d'Economie des Transports
CNRS – Université de Lyon - ENTPE



① La mixité des modes d'enquête

Difficulté de mener des EMD représentatives de l'ensemble de la population avec les modes d'enquête traditionnels :

⇒ Le taux de non-réponse est en augmentation
(35% en 1995, 47% en 2006)

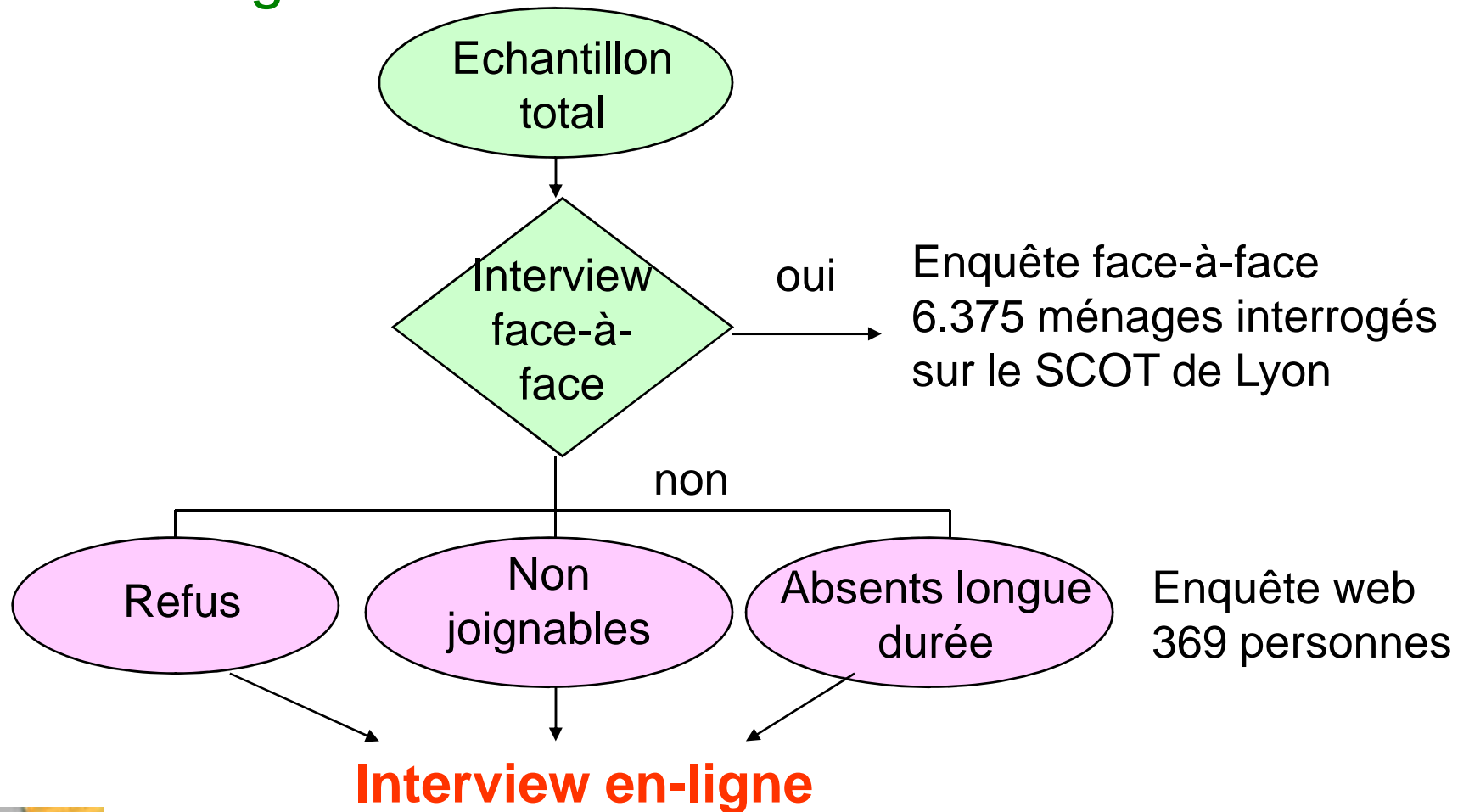
⇒ Certains segments semblent sous représentés

Mais les interviews en face-à-face sont utiles pour collecter des données fiables et permettre l'étude des comportements

Proposer une alternative aux personnes non joignables ou refusant de recevoir un enquêteur



② Methodologie



③ Premières conclusions

Les internautes mobiles se déplacent moins que les individus interrogés en face-à-face (3,78 vs. 4,19 déplacements) :

⇒ Mobilité marche à pied deux fois plus faible

⇒ Moindre participation aux activités d'accompagnement et de loisirs

Hypothèses de recherche :

⇒ Sous-déclaration imputable au média web

⇒ Différences de profils qui impactent la mobilité

**A caractéristiques socioéconomiques égales,
les différences de mobilité subsistent**



④ Le biais de sélection

Mixité des modes d'enquête → les répondants appartiennent à un groupe → les réponses ne sont pas comparables :

⇒ Répartition non aléatoire

⇒ Facteurs externes qui affectent la mobilité

Effet du mode d'enquête sur la mobilité individuelle :

$$Y_i = \beta_i X_i + \alpha_i I_i + u_i$$

⇒ I_i ne peut pas être considéré comme exogène

⇒ Le biais de sélection doit être corrigé

Méthode d'estimation en deux étapes (Heckman, 1979)



5 La méthode d'estimation en deux étapes

Modèle explicatif de la mobilité :

$$\Rightarrow Y_{1i} = \beta_i X_{1i} + u_{1i}, \text{ pour les répondants web}$$

$$\Rightarrow Y_{2i} = \beta_i X_{2i} + u_{2i}, \text{ pour les répondants en face-à-face}$$

Equation de sélection :

$$\Rightarrow I_i^* = \delta_i Z_i + \varepsilon_i, \begin{cases} I_i = 1, \text{ ssi } I_i^* > 0, \\ I_i = 0, \text{ ssi } I_i^* \leq 0 \end{cases}$$

La mobilité observée pour un individu i est :

$$\Rightarrow Y_i = Y_{1i}, \text{ ssi } I_i = 1$$

$$\Rightarrow Y_i = Y_{2i}, \text{ ssi } I_i = 0$$



La méthode d'estimation en deux étapes (2)

La régression de Y_i sur X_i devient :

$$\Rightarrow Y_{1i} = \beta_1 X_{1i} + \rho_1 \sigma_{u1i} \lambda_{1i} + e_{1i}, \text{ pour le web}$$

$$\Rightarrow Y_{2i} = \beta_2 X_{2i} + \rho_2 \sigma_{u2i} \lambda_{2i} + e_{2i}, \text{ pour le face-à-face}$$

Biais de sélection \rightarrow biais de valeur manquante :

$$\begin{aligned} \Rightarrow \lambda_{1i} &= \frac{\phi(\delta_i Z_i)}{\theta(\delta_i Z_i)} \\ \Rightarrow \lambda_{2i} &= \frac{-\phi(\delta_i Z_i)}{1 - \theta(\delta_i Z_i)} \end{aligned} \left. \vphantom{\begin{aligned} \Rightarrow \lambda_{1i} &= \frac{\phi(\delta_i Z_i)}{\theta(\delta_i Z_i)} \\ \Rightarrow \lambda_{2i} &= \frac{-\phi(\delta_i Z_i)}{1 - \theta(\delta_i Z_i)} \end{aligned}} \right\} \begin{array}{l} \phi \text{ et } \Phi = \text{fonctions de densité et} \\ \text{de répartition de la loi normale} \end{array}$$

1^{ère} étape : estimer l'équation de sélection (probit)

2^{ème} étape : estimer β et $\rho\sigma_u$ par régression MCO

Equation de sélection : les répondants web

	Coefficient	Signif.
Constante	-3.712	***
Age	0.049	***
(Age) ²	-4.78e-04	**
Connexion internet : oui	0.495	***
Téléphone portable : oui	0.257	**
Liste téléphone : oui	0.372	***
Liste téléphone : pas de téléphone	0.248	*
Nb de personnes / ménage	-0.101	***
Lieu de travail : non précisé	-0.993	**
Lieu de travail : périphérie	-0.24	**
Lieu de travail : inactif	-0.244	**
Diplôme : supérieur	0.41	***
Densité de la zone de résidence	-2.169e-05	**
Revenu déclaré : oui	0.401	***
Vendredi : oui	0.469	***
Distance domicile / centre	-1.9e-05	*

Significativité :

0 '****'

0.001 '***'

0.01 '**'

0.05 '.'

0.1 ''

Equation de mobilité : les répondants en face-à-face

Echantillon face-à-face	Régression sans correction		Régression avec correction	
	Coeff.	Sign.	Coeff.	Sign.
Constante	2.013	***	2.640	***
Sexe : homme	-0.24	***	-0.24	***
Age	0.065	***	0.057	***
(Age) ²	-7.3e-04	***	-6.4e-04	***
Possession permis : oui	0.465	***	0.444	***
Nb d'enfants / ménage	0.711	***	0.710	***
(Nb d'enfants / ménage) ²	-0.004	*	-0.003	*
Nb de voitures / personne	0.571	***	0.534	***
Nb de personnes / ménage	-0.188	***	-0.181	***
Revenu déclaré : oui	0.325	***	0.268	***
Activité : non actif	0.288	***	0.304	***
Distance domicile / centre	6.7e-06		9.5e-06	*
Mills	NA		-0.18	**

Significativité :

0 '****'

0.001 '***'

0.01 '**'

0.05 '.'

0.1 ''



Equation de mobilité : les répondants web

Echantillon web	Régression sans correction		Régression avec correction	
	Coeff.	Sign.	Coeff.	Sign.
Constante	0.894		3.038	.
Sexe : homme	-0.846	***	-0.870	**
Age	0.077		0.039	
(Age) ²	-8.57e-04		-4.38e-04	
Possession permis : oui	0.263		0.454	
Nb d'enfants / ménage	0.399		0.465	
(Nb d'enfants / ménage) ²	-0.121		-0.151	
Nb de voitures / personne	0.683	.	0.690	.
Nb de personnes / ménage	0.266	.	0.299	.
Revenu déclaré : oui	0.360		0.166	
Activité : non actif	0.741	*	0.823	*
Distance domicile / centre	-5.7e-05	*	-5.2e-05	.
Mills	NA		-0.731	**

Significativité :

0 '****'

0.001 '***'

0.01 '**'

0.05 '.'

0.1 ''



Modèle final : interactions significatives

Modèle final	Coefficients	Significativité
Constante	2.650	***
Sexe : homme	-0.240	***
Age	0.057	***
Age ²	-6.36e-04	***
Possession permis : oui	0.445	***
Nb d'enfants / ménage	0.694	***
(Nb d'enfants / ménage) ²	-0.033	.
Nb de voitures / personne	0.534	***
Nb de personnes / ménage	-0.173	***
Revenu déclaré : oui	0.265	***
Activité : non actif	0.315	***
Distance domicile / centre	9.54e-06	*
Mills	-0.191	***
Mode	-0.504	0.176
(Sexe : homme) * mode	-0.547	0.059
(Nb de personnes / ménage) * mode	0.187	0.102
(Distance domicile / centre) * mode	-5.9e-05	0.053

Significativité :

0 '***'

0.001 '***'

0.01 '**'

0.05 '.'

0.1 ''



⑥ Synthèse des résultats

Le fait de répondre en-ligne influence négativement la mobilité (-0.5 déplacements).

Trois variables interagissent également avec le mode d'enquête (web) :

⇒ Le sexe : impact < 0

⇒ Le nombre de personnes du ménage : impact > 0

⇒ La distance du domicile / centre : impact < 0

Permet de quantifier pour chaque répondant l'impact du mode sur le comportement de mobilité



7 Pistes d'interprétation

Les répondants web se déplacent moins que les répondants en face-à-face :

- ⇒ Le risque est de confondre le phénomène étudié avec le processus de sélection des répondants
- ⇒ La faible diffusion du web et les exigences de qualité ne permettent pas d'occulter ce biais
- ⇒ La procédure d'estimation en deux étapes propose de corriger le biais de sélection

Des analyses complémentaires par mode et par motif sont nécessaires

