Quelques éléments méthodologiques sur les enquêtes de mobilité

Ph. L. Toint

Rapport 03/21 9 décembre 2003

Groupe de Recherche sur les Transports (GRT),
Département de Mathématique,
Facultés Universitaires ND de la Paix,
61, rue de Bruxelles, B-5000 Namur, Belgique,
Email : philippe.toint@fundp.ac.be

Quelques éléments méthodologiques sur les enquêtes de mobilité

Philippe Toint (GRT, FUNDP)

8 décembre 2003

Résumé

Ce texte présente succinctement des éléments méthodologiques concernant l'organisation pratique, par les communes ou autres entités territoriales, d'enquêtes sur les comportements de mobilité des personnes. On aborde, en particulier, les contraintes qui déterminent la pertinence de telles enquêtes ainsi que les étapes successives de la méthode qui doit être appliquée dans ce domaine pour garantir des résultats fiables.

1 Introduction

La mobilité des personnes constitue une des problématiques de la société contemporaine dont l'importance va croissant. Son organisation préoccupe de plus en plus de décideurs et de citoyens. Parallèlement, la gestion de la mobilité nécessite davantage que par le passé d'en bien connaître et comprendre les différentes composantes. En simplifiant beaucoup, on peut dire que la mobilité est la rencontre, dans l'espace et dans le temps, d'une offre de déplacement relativement objective (infrastructures routières ou ferrées, offre de transport public, ...) et d'une demande associée, qui est davantage déterminée par des aspects comportementaux plus subjectifs. Il est donc souvent crucial, tant du point de vue du gestionnaire que de celui de l'utilisateur des transports, de rassembler des données sur la mobilité, ce qui implique donc de collecter des informations sur l'offre et sur la demande afin de permettre leur mise en conjonction. Cette note s'intéresse au second de ces aspects complémentaires : la collecte d'informations sur les comportements de la mobilité des personnes, en se focalisant sur les aspects relevant de la mobilité quotidienne comme différente de la mobilité occasionnelle (tourisme, longs voyages).

2 Les enquêtes de comportement : où et quand ?

Une enquête sur les comportements de mobilité consiste à observer un nombre suffisant de choix effectués par des personnes dans leurs déplacements quotidiens, et de relier ces choix à des indicateurs qui décrivent soit l'offre, soit l'environnement individuel (socio-économique, spatial ou temporel) dans lequel les choix sont effectués. Cette nature même impose quelques limites au cadre dans lequel de telles enquêtes ont du sens. Nous en abordons les principales dans cette section.

2.1 Limites statistiques

A moins d'être exhaustive, ce qui ne peut être envisagé que dans le cas de très petites entités, une enquête de mobilité procède en collectant des données sur un échantillon de la population que l'on souhaite étudier. Ceci implique qu'une erreur (appelée erreur d'échantillonnage) rend les résultats obtenus moins précis. Pour limiter cette imprécision, il est nécessaire d'interroger un nombre suffisant de personnes (souvent de l'ordre du millier ou plus). La précision des résultats augmente avec ce nombre, mais n'est jamais parfaite. La qualité des conclusions est aussi fortement influencée par la méthodologie d'échantillonnage utilisée et la rigueur de l'analyse (voir plus bas).

2.2 Limites organisationnelles

La limitation probablement la plus importante est d'ordre organisationnel. Comme nous le décrirons plus bas, la mise sur pied effective d'une procédure d'interrogation d'un nombre significatif de répondants est lourde, complexe et potentiellement coûteuse. Un tel effort doit donc être justifié par l'importance des résultats attendus. En particulier, il est souvent important de couvrir une population "cible" qui ne soit pas trop réduite (afin de rendre les résultats significatifs à une niveau qui ne soit pas trop microscopique). Des agglomérations de 10.000 habitants et plus sont donc raisonnablement adaptées.

La taille de la population à étudier ne consiste pas le seul facteur limitant dans l'organisation des enquêtes de mobilité. Les moyens humains et budgétaires nécessaires sont encore plus souvent déterminants dans la décision d'organiser une telle enquête.

- Du point de vue humain d'abord, l'organisation d'une enquête de mobilité nécessite l'intervention d'enquêteurs qui doivent être formés, et, plus crucialement encore, de gestionnaires de l'enquête qui en comprennent bien les difficultés (cohérence de l'organisation, techniques de rédaction des questionnaires, rigueur de la procédure, aspects statistiques) et qui peuvent ensuite l'analyser avec succès (familiarité avec les analyses typiques, maîtrise des moyens informatiques et logiciels nécessaires). Soit ces personnes appartiennent au personnel communal ou urbain et ont été formées spécifiquement à ce type de travail, soit un consultant externe est engagé.
- Les aspects budgétaires couvrent, outre la formation du personnel lorsque nécessaire et le temps que celui-ci passe au projet, les coûts d'enquête proprement dits (impression des questionnaires et autres documents papier, trajets des enquêteurs, coûts informatiques et téléphoniques, ...)

On le voit, ces coûts peuvent s'avérer élevés pour des pouvoirs locaux dont les moyens sont limités, ce qui renforce l'observation que ce sont les entités plus importantes qui se lancent le plus souvent dans l'aventure.

2.3 Limites de situation

Finalement, une enquête de mobilité dans une commune n'a de sens que si l'on interroge les acteurs réels de la mobilité locale : dans une commune de transit, il faut bien sûr interroger les personnes qui traversent la commune plutôt que les résidents, ou, en tous cas, à côté d'eux.

3 Les étapes d'une enquête de mobilité

Malgré ces difficultés, la réalisation d'une enquête de mobilité représente une source d'information remarquable et doit être encouragée dans la mesure du possible. Nous analysons rapidement, dans les paragraphes qui suivent, les phases principales d'une telle entreprise.

3.1 Les objectifs

La définition des objectifs de l'enquête est un préalable dont on ne peut sous-estimer l'importance. Une telle définition couvre, entre autres,
- la définition de la zone à étudier,
- la définition de la population “cible”, c'est-à-dire de la population dont on veut analyser le comportement (âge, résidence, citoyenneté, ...),
- la définition des indicateurs principaux que l'on souhaite mesurer (caractéristiques socio-économiques des individus et des ménages, types d'utilisation des transports, opinions, ...)
- de la période d'enquête,
- du type d'organisation de celle-ci (aussi appelé *protocole*, que nous abordons plus bas),
- du type de diffusion à assurer aux résultats.

Il est en effet relativement difficile de modifier la marche d'une entreprise aussi complexe “au milieu du gué”...

En ce qui concerne la population, on peut distinguer une enquête qui s'adresse à des ménages ou à des individus, les deux approches pouvant être couplées (comme dans l'Enquête Nationale MOBEL 1999 ou l'Enquête Régionale ERMM).

3.2 Le cadre d'échantillonnage

Une fois la population ciblée bien circonscrite, il faut rechercher un *cadre d'échantillonnage* qui lui soit adapté, c'est-à-dire une liste (dans notre cas, de personnes) dans laquelle un échantillon aussi représentatif que possible sera tiré. En effet, il est en général impossible d'interroger exhaustivement toute la population ciblée, pour des raisons qui combinent le prix de revient et l'ampleur de l'organisation. Les sources traditionnelles pour le cadre d'échantillonnage d'enquête de transports sont :
- les indicateurs téléphoniques,
- les listes électorales,
- les listes administratives,

– les listes maintenues par des compagnies de distribution de services comme le gaz, l'eau ou l'électricité,
– les autres listes assemblées par le secteur privé (souvent spécialisé).

En Belgique, l'administration fédérale maintient, avec l'aide des communes, une liste de la troisième de ces catégories : le *Registre National.* Ce registre d'état civil recense de manière relativement complète tous les résidents dans le pays. Il est donc plus complet que d'autres sources comme les listes électorales (limitées aux Belges) ou d'autres fichiers commerciaux (incomplets, biaisés et souvent chers). Régulièrement mis à jour, il risque moins que les autres fichiers de contenir des informations périmées ou des doublons. Enfin, il contient, pour chaque résident, plusieurs variables comme le nom, l'adresse, le sexe, la date de naissance, la position dans le ménage (chef de ménage, conjoint, enfant), la structure et la composition du ménage ..., informations qui sont très utiles lors de la sélection de l'échantillon. Le seul défaut du Registre National, de notre point de vue, est son accès, qui, pour des raisons évidentes de protection de la vie privée, n'est pas direct, mais bien règlementé.

3.2.1 La sélection de l'échantillon

Une fois le cadre d'échantillonnage choisi, l'échantillon de la population peut alors être sélectionné. La plupart des méthodes acceptables pour cette sélection, l'*échantillonnage*, sont basées sur le principe du tirage aléatoire dans la population ciblée (voir Ardilly [1], Richardson, Meyburg et Ampt [6] ou Morin [5] pour de plus amples justifications et une présentation des autres méthodes possibles).

L'échantillonnage aléatoire simple L'idée fondamentale d'un tel échantillonnage est que le tirage de chaque unité de la population ciblée (dans le cadre d'échantillonnage) est établi indépendamment, et que chaque unité a une probabilité égale d'être sélectionnée dans l'échantillon. La taille de l'échantillon (le nombre d'individus sélectionnés) permet de prévoir le degré de précision obtenu sur les indicateurs calculés à partir de cet échantillon. Formellement, l'erreur standard sur la moyenne des indicateurs dérivés d'un échantillon de taille n sélectionné dans une population de taille N est approximé, pour les taux de sondage (n/N) faibles, par la formule

$$\sigma\sqrt{\frac{N-n}{Nn}} \approx \frac{\sigma}{\sqrt{n}},$$

où σ est la déviation standard (écart type) de la population complète pour l'indicateur considéré. Cette formule montre que, dans ces circonstances, la précision des indicateurs ne dépend plus réellement de la taille N de la population, mais uniquement de la taille n de l'échantillon. Sous cette hypothèse, un échantillon de taille 2500 assure une erreur standard approximative de 2%.

Autres techniques d'échantillonnage D'autres techniques plus complexes sont souvent utilisées, afin de garantir la réprésentativité de sous-groupes de la population étudiée. Nous décrivons brièvement les plus courantes.

L'échantillonnage aléatoire stratifié : Utilisé seul, l'échantillonnage aléatoire simple assure la représentativité globale de la population mais n'implique pas nécessairement la représentativité au niveau de segments particuliers de la population qui sont considérés comme potentiellement importants pour les objectifs de l'enquête. Pour corriger ce défaut, on procède en trois étapes successives.

1. On détermine les segments de la population dont on veut assurer la représentativité. Ces segments sont appelés *strates* de la population. Ils sont d'autant plus adaptés qu'ils représentent un sous-groupe homogène de la population, alors que les différences entre strates peuvent être plus grandes.
2. L'échantillon total est alors subdivisé en sous-échantillons dont la taille est calculée de façon à obtenir la précision (en termes d'erreur standard) souhaitée.
3. L'échantillonnage total est alors réparti entre les strates, et un échantillonnage aléatoire simple (de taille fixée) est ensuite réalisé à l'intérieur de chaque strate.

L'enquête est ensuite réalisée sur la combinaison des échantillons par strate.

L'échantillonnage par quotas : Cette méthode intègre l'échantillonnage à l'exécution de l'enquête elle-même. L'idée est de déterminer *a priori* un certain nombre de strates (comme ci-dessus) et ensuite de tirer aléatoirement *en cours d'enquête* des répondants dans chaque strate jusqu'à ce qu'un nombre suffisant (*quota*) d'entre eux puisse être effectivement contacté.

Une difficulté se pose à cette étape. En effet, les décisions importantes qui concernent la mobilité (choix des heures de départ, du mode de déplacement, voire des activités elles-mêmes) sont prises non pas au niveau de l'individu lui-même, mais au niveau du ménage auquel cet individu appartient. Questionner des individus isolés dans des ménages ne peut donc refléter le contexte de ces décisions que partiellement. Pour pallier cet inconvénient, il est possible, comme dans MOBEL ou l'ERMM, de sélectionner des ménages dans l'échantillon et d'ensuite interroger tous leurs membres. Cette procédure permet de rassembler en particulier de l'information sur la corrélation des décisions de mobilité dans les ménages, mais doit faire l'objet d'un traitement statistique spécial au niveau de l'estimation de l'erreur.

3.3 Le questionnaire

Vient ensuite la rédaction du questionnaire, qui doit impérativement tenir compte des contraintes suivantes :

complétude : Le questionnaire doit rassembler toutes les informations nécessaires au calcul des indicateurs retenus dans les objectifs de l'enquête.

clarté : Les questions doivent être rédigées clairement, dans un language abordable par tous et les termes liés à la mobilité doivent y être définis avec précision. La mise en page doit favoriser la lisibilité et encourager les réponses.

taille : Le questionnaire ne doit pas être trop long, au risque de fatiguer ou de décourager les répondants.

économie : Il ne faut inclure de questions qui ne concernent qu'une marge très faible de la population et n'auraont donc pas de réponse statistiquement fiable.

Les questionnaires sont le plus souvent l'objet d'expérimentations (appelés *pilotes*) auprès d'échantillons de répondants très petits, afin de les améliorer avant de les appliquer dans l'enquête proprement dite.

Il est aussi important que les questionnaires de différentes enquêtes de mobilité réalisées en Région Wallonne s'appuient sur des définitions communes (voir l'ERMM) et un tronc commun d'indicateurs, afin de rendre les résultats comparables entre entités wallonnes différentes et de permettre l'intégration des résultats obtenus localement au niveau sous-régional ou régional.

3.4 Le protocole d'enquête

Le protocole d'enquête doit ensuite être mis au point. Il doit au moins aborder les questions suivantes.

3.4.1 L'orientation temporelle

On peut interroger les répondants sur leurs choix passés (un jour, plusieurs jours, plusieurs mois) : on parle alors de *protocole rétrospectif*, ou de *préférences révélées*. C'est la méthode la plus souvent utilisée, mais il convient d'en sélectionner l'horizon (la période sur laquelle on interroge les répondants) avec soin. En effet, plus l'horizon est long, plus la mémoire risque d'être défaillante et l'information incomplète ou biaisée.

Il est aussi possible d'interroger les répondants sur leurs choix futurs, y compris dans des situations hypothétiques (par exemple, dans l'hypothèse où l'offre de transport serait modifiée). Il s'agit alors d'un *protocole prospectif*, ou de *préférences déclarées*. Ce type de protocole est rarement utilisé seul, il ne se révèle fiable que dans le cas où les questions posées au répondant sur ses choix futurs sont articulées sur son expérience passée.

3.4.2 La technique de contact

Les répondants peuvent être contactés et interrogés de différentes manières.

- Les *enquêtes en face-à-face*, où un enquêteur interroge le répondant en sa présence, sont qualitativement les meilleures, mais sont considérablement plus coûteuses à organiser, surtout si la zone géographique considérée est étendue.
- Les *enquêtes téléphoniques* sont plus abordables, mais limitent nettement la taille du questionnaire, le nombre de personnes interrogées dans un ménage et la population qui peut être atteinte.

– Les *enquêtes postales*, qui demandent au répondant de remplir un questionnaire papier chez lui et de le renvoyer ensuite. Leur inconvénient est qu'elles suscitent généralement un taux de réponse assez faible et introduisent parfois un biais lié aux caractéristiques socio-économiques des répondants. En pratique, les enquêtes postales pures (via toutes-boîtes, par exemple) sont peu utilisées pour ces raisons.

On combinera fréquemment ces types d'enquête en un protocole plus complexe (face-à-face remplacé, le cas échéant, par un entretien téléphonique, questionnaire postal supporté par des appels téléphoniques de motivation et/ou de rappel, ...), mais il reste utile d'expérimenter avec la méthode choisie avant son application en vraie grandeur. Les deux premiers types de protocole sont souvent guidés par des programmes informatiques, ce qui donne lieu au CAPI (computer aided personal inteview) ou au CATI (computer aided telephone inteview). Des expériences ont actuellement lieu pour utiliser le Web comme moyen de communiquer avec les répondants.

3.5 La récolte réelle des données et l'encodage

Un fois toutes ces préparations effectuées, il reste alors à réaliser l'enquête sur le terrain, c'est-à-dire à interroger effectivement les répondants. Cette phase nécessite une bonne organisation pratique et un suivi méticuleux des actions entreprises et du protocole.

Les réponses obtenues sont le plus souvent encodées au fur et à mesure. L'encodage lui-même doit faire l'objet d'une attention toute particulière pour assurer la qualité des données. Des programmes informatiques de vérifications croisées ou de double encodage sont couramment utilisés.

3.6 Validation et correction

Si des réponses apparaissent comme impossibles ou même peu plausibles, il est fréquent que les répondants soient recontactés pour s'assurer de leur réponses. Certaines erreurs peuvent aussi être corrigées directement sans nouveau contact, mais la prudence s'impose alors pour ne pas introduire de biais (voir, par exemple, [4] pour une discussion de méthodes comme le "hot deck"). Cette phase de *correction* est généralement cruciale pour la qualité finale des résultats de l'enquête. Il est aussi recommandé de procéder à une *validation* aussi large que possible, c'est-à-dire à une verification systématique des réponses avec un sous-ensemble des répondants.

3.7 Analyse

L'analyse des résultats peu enfin commencer... Elle doit extraire des données récoltées non seulement des valeurs fiables pour les indicateurs définis dans les objectifs, mais elle présente aussi des corrélations entre ceux-ci. Elle est

– d'abord *descriptive* quand elle présente ces résultats et ces corrélations sans tentative d'interprétation élaborée,

– et ensuite *modélisatrice* lorsqu'elle propose des modèles plus complexes du comportement des individus, modèles dont les paramètres sont calibrés sur les données obtenues par enquête.

A titre d'exemple, l'ouvrage [3] présente une analyse descriptive de l'enquête MOBEL, tandis que [2] décrit un modèle comportemental (pour les choix des destinations) qui utilise ces mêmes données.

Notons que les contraintes de la préservation de la vie privée des répondants doivent rester présentes à toutes les phases de l'analyse, ce qui impose un traitement confidentiel des données et une attention spécifique pour éviter de rendre les comportements individuels reconnaissables au travers des indicateurs publiés.

3.8 Présentation des résultats

L'analyse se conclut par la rédaction de rapports ou la présentation des résultats sous d'autres formes (conférences, sites Web, ...).

4 Conclusions

Nous avons brièvement résumé dans cette note les limites et les techniques associées aux enquêtes sur les comportements de mobilité. Il est bien clair que notre exposé a laissé un bon nombre de sujets importants dans l'ombre (panels, redressements, protocoles complexes, modèles d'interprétation, détails organisationnels, ...) et nous encourageons le lecteur intéressé à completer son information dans les ouvrages cités. Nous espérons cependant que ce bref aperçu permettra une première approche en mettant en évidence quelques unes des questions de fond qui se posent lorsque l'on souhaite aborder ce sujet passionnant.

Si une commune estime qu'une enquête sur les comportements de mobilité s'impose sur son territoire, notre discussion montre en particulier qu'il est d'abord important de mesurer l'impact des contraintes de ce type de procédure (coût, formation, temps passé) et qu'il est ensuite crucial d'insister sur la nature rigoureuse de l'approche (échantillonnage, vocabulaire utilisé, protocole,...) pour que les résulats obtenus ne déçoivent pas leurs commenditaires.

Remerciements

Merci à Eric Cornélis et Jean-Paul Hubert pour leur relecture de ce texte.

Références

[1] P. Ardilly. *Les techniques du sondage.* Editions Technip, Paris, 1994.

[2] C. Cirillo, E. Cornélis, L. Legrain, and Ph. L. Toint. Combining spatial and temporal dimensions in destination choice models. Rapport GRT 03/13, Groupe de Recherche sur les Transports, Département de Mathématique, Facultés Universitaires ND de la Paix, Namur, Belgique, 2003.

[3] J.-P. Hubert and Ph. L. Toint. *La mobilité quotidienne des Belges.* Number 1 in Mobilité et Transports. Presses Universitaires de Namur, Namur, Belgium, 2002.

[4] O. Lothaire and Ph. L. Toint. A toolbox approach to data correction and imputation. In *Capturing Long-Distance Travel*, pages 244–255. Research Science Press, Baldock, UK, 2003.

[5] H. Morin. *Théorie de l'échantillonnage.* Presses Universitaires Laval, Sainte Foy, Canada, 1993.

[6] A. J. Richardson, E. S. Ampt, and A. H. Meyburg. *Survey Methods for Transport Planning.* Eucalyptus Press, Melbourne, Australia, 1995.