

Article

Symposium 2008 :

Collecte des données : défis, réalisations et nouvelles orientations

Effets du nombre de tentatives d'appel sur les taux de non-réponse et sur le biais dû à la non-réponse – Résultats de certaines études de cas réalisées à Statistics Sweden

par Jan Hörngren, Peter Lundquist et Sara Westling

2009



Effets du nombre de tentatives d'appel sur les taux de non-réponse et sur le biais dû à la non-réponse – Résultats de certaines études de cas réalisées à Statistics Sweden

Jan Hörngren, Peter Lundquist et Sara Westling¹

Résumé

Dans une enquête téléphonique, le choix du nombre de tentatives d'appel représente une décision importante. Un grand nombre de tentatives rend la collecte des données longue et dispendieuse, tandis qu'un petit nombre réduit l'ensemble de réponses sur lequel sont fondées les conclusions et accroît la variance. La décision peut aussi avoir une incidence sur le biais de non-réponse. Dans le présent article, nous examinons les effets du nombre de tentatives d'appel sur le taux de non-réponse et sur le biais dû à la non-réponse dans deux enquêtes réalisées par Statistics Sweden, à savoir la Labour Force Survey (LFS) et la Household Finances (HF).

Au moyen de paradonnées, nous calculons le taux de réponse en fonction du nombre de tentatives d'appel. Pour estimer le biais de non-réponse, nous utilisons les estimations de certaines variables de registre, pour lesquelles des observations sont disponibles pour les répondants ainsi que les non-répondants. Nous calculons aussi les estimations de certains paramètres d'enquête réels en fonction d'un nombre variable de tentatives d'appel. Les résultats indiquent qu'il est possible de réduire le nombre courant de tentatives d'appel sans accroître le biais de non-réponse.

Mots clés : Tentatives d'appel (événements WinDATI), taux de non-réponse, biais de non-réponse.

1. Introduction

Pourquoi le choix du nombre de tentatives d'appel est-il important dans une enquête (téléphonique)?

D'une part, si le nombre de tentatives d'appel est élevé, la collecte des données prend beaucoup de temps et est donc très coûteuse – le temps, c'est de l'argent! D'autre part, un petit nombre de tentatives d'appel réduit l'ensemble de réponses, ce qui accroît la variance, autrement dit produit de plus grandes erreurs d'échantillonnage, de plus grands intervalles de confiance et, conséquence probablement plus grave, un plus grand risque de biais dû à la non-réponse.

Dans le présent article, nous utilisons des paradonnées pour calculer le taux de réponse en fonction du nombre de tentatives d'appel. Pour avoir une idée du biais de non-réponse, nous utilisons des estimations calculées pour diverses variables de registre, pour lesquelles des observations sont disponibles pour les répondants ainsi que les non-répondants. En outre, dans ce cas, nous calculons les estimations pour divers nombres de tentatives d'appel. L'article s'appuie principalement sur deux études réalisées à Statistics Sweden : Lundquist et coll. (2007) et Westling (2008). Une étude à grande échelle dans ce domaine a été réalisée antérieurement par Japac et Hörngren (2005), mais leurs résultats ne sont pas comparables à ceux présentés ici, parce que leur définition d'une tentative d'appel est différente.

¹Jan Hörngren, méthodologiste principal, Statistics Sweden, 702 33 Örebro, Sweden (jan.horngren@scb.se); Peter Lundquist, méthodologiste principal, Statistics Sweden, Sweden, Box 24300 104 51 Stockholm; Sara Westling, méthodologiste principale, Statistics Sweden, 702 33 Örebro, Sweden

2. Possibilités et contraintes

Nous décrivons dans cette section les possibilités et les contraintes relatives à l'analyse et à l'élaboration de stratégies de tentatives d'appel à Statistics Sweden. Nous nous appuyons pour cela sur l'analyse de parodonnées provenant du système WinDATI, le système d'interview assistée par ordinateur (IAO) de Statistics Sweden. Quand un intervieweur utilise WinDATI, chaque action produit ce que l'on appelle un *événement* (événement WinDATI). Les événements WinDATI qui suivent ont été utilisés dans l'analyse :

- interview achevée;
- interview interrompue → nouveau rendez-vous;
- interview interrompue → code final;
- code de rendez-vous;
- numéro de téléphone occupé;
- pas de réponse (pas de contact);
- tentatives d'appel supplémentaires durant la journée terminées;
- code final (code de non-réponse);
- fermeture/terminaison du cas pour l'unité échantillonnée.

Chaque événement est doté d'un code unique et, ici, chaque événement (susmentionné) correspond à une tentative d'appel².

Un autre élément important, pour ce qui est des possibilités et contraintes, est l'utilisation de deux équipes d'intervieweurs, l'une formée d'intervieweurs sur place et l'autre, d'intervieweurs centralisés (intervieweurs par ITAO) :

- les intervieweurs sur place sont répartis dans le pays et travaillent à partir de leur domicile. Ils effectuent des interviews sur place **et** des interviews téléphoniques, et reçoivent périodiquement un certain nombre de cas par enquête, lesquels représentent leur charge de travail personnelle;
- les intervieweurs par ITAO travaillent dans un bureau central à partir duquel ils effectuent uniquement des interviews téléphoniques. Les cas qui leur sont attribués proviennent d'une base de données commune.

Les intervieweurs sur place et les intervieweurs par ITAO ont aussi des contrats différents.

Statistics Sweden souhaiterait que les cas soient attribués à tous les intervieweurs (dans les enquêtes téléphoniques) à partir d'une base de données commune. Ce n'est probablement que dans de telles conditions qu'une stratégie de tentative d'appel pourrait être fructueuse.

3. Études empiriques des effets du nombre de tentatives d'appel

3.1 Données d'enquête et analyses

Les études sont fondées sur deux enquêtes :

- La Labour Force Survey (LFS) est une enquête mensuelle réalisée auprès d'un échantillon de 21 000 personnes chaque mois. La base de sondage est le registre de la population suédoise. La collecte des données est effectuée par des intervieweurs par ITAO et des intervieweurs sur place. Le taux de réponse est de 83 % (valeur moyenne).

² Dans le présent article, nous utilisons le concept de l'événement WinDATI pour éviter des comparaisons invalides avec d'autres concepts de tentative d'appel.

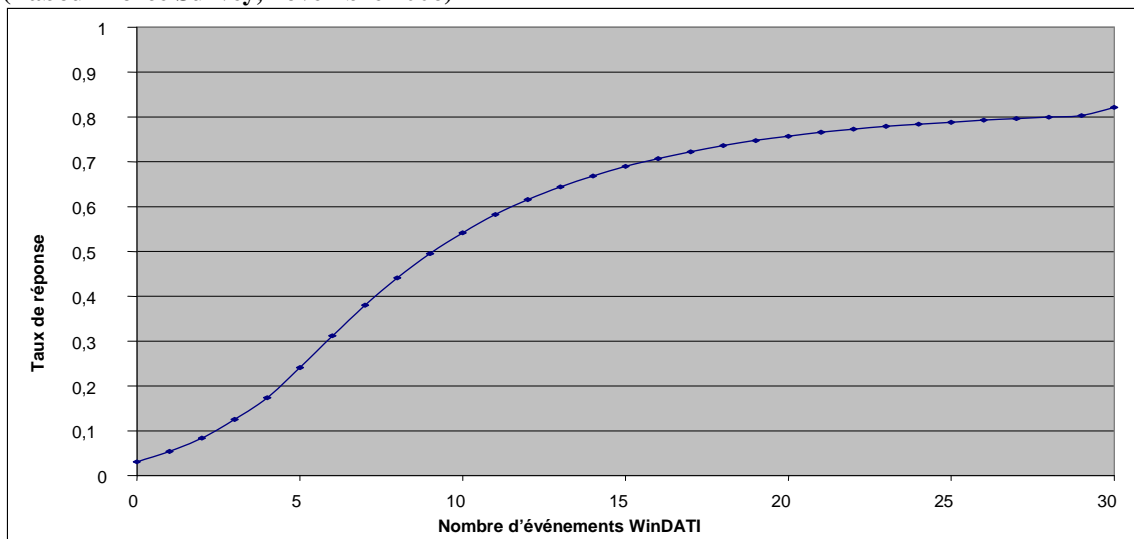
- La Household Finances (HF) est une enquête annuelle réalisée auprès d'un échantillon de 17 000 ménages. Pour cette enquête également, la base de sondage est le registre de la population suédoise (échantillon en réseau). La collecte des données est effectuée par des intervieweurs sur place. Le taux de réponse est de 70 %.

Les études des effets du nombre de tentatives d'appel se sont concentrées sur le taux de non-réponse et le biais de non-réponse. Le taux de non-réponse est un indicateur de base de la qualité d'une enquête. Le biais de non-réponse est un indicateur très important dans le contexte de stratégies de tentatives d'appel. Il peut vraiment nous indiquer ce qui se passe en ce qui concerne les estimations d'intérêt.

3.2 Effet sur le taux de non-réponse

Dans cette section, nous étudions l'effet des efforts sur le terrain (exprimés en nombre d'événements WinDATI) sur le taux de réponse. La figure 3.2-1 donne le taux de réponse cumulé en fonction du nombre d'événements WinDATI pour la LFS. (Dans cette figure et celles qui suivent, la valeur valide pour le nombre maximal d'événements WinDATI (=30) sur l'axe des x est ≥ 30 . Les cas pour lesquels le nombre d'événements est supérieur à 30 sont très rares.)

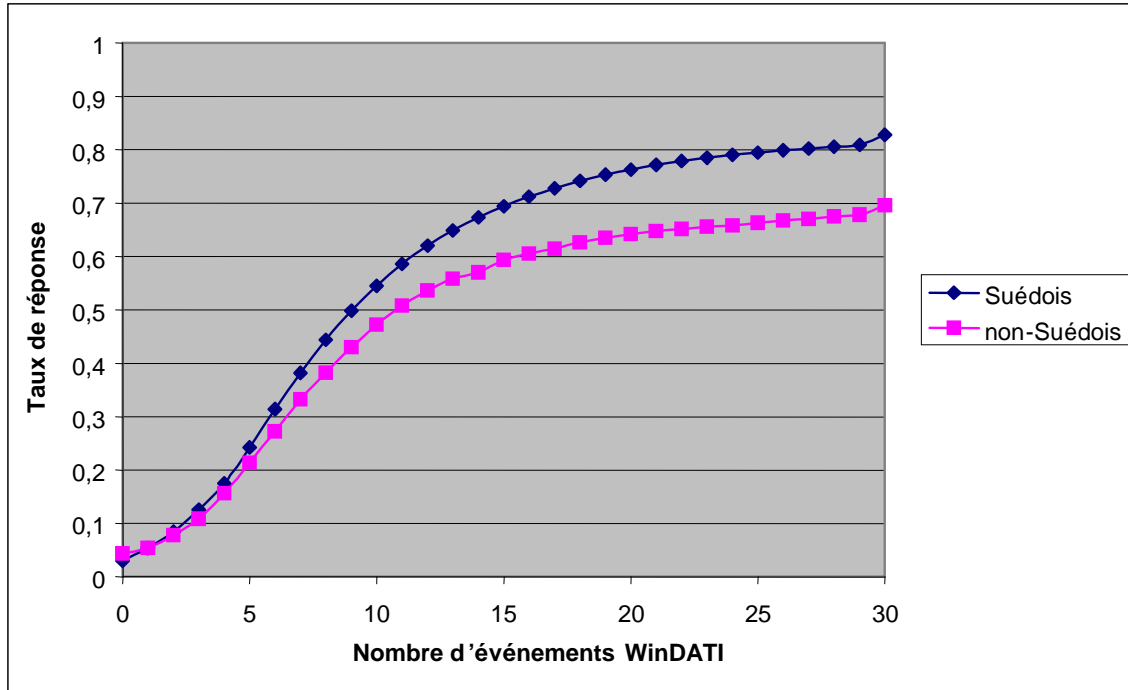
Figure 3.2-1
Taux de réponse (global) cumulé en fonction du nombre d'événements WinDATI
(Labour Force Survey, novembre 2006)



Nous voyons que le taux de réponse semble stable après 20 événements WinDATI et que le « niveau de réponse » atteint environ 70 % après 15 événements. La figure 3.2-1 est valide pour l'échantillon complet. La situation reste-t-elle la même pour différents domaines?

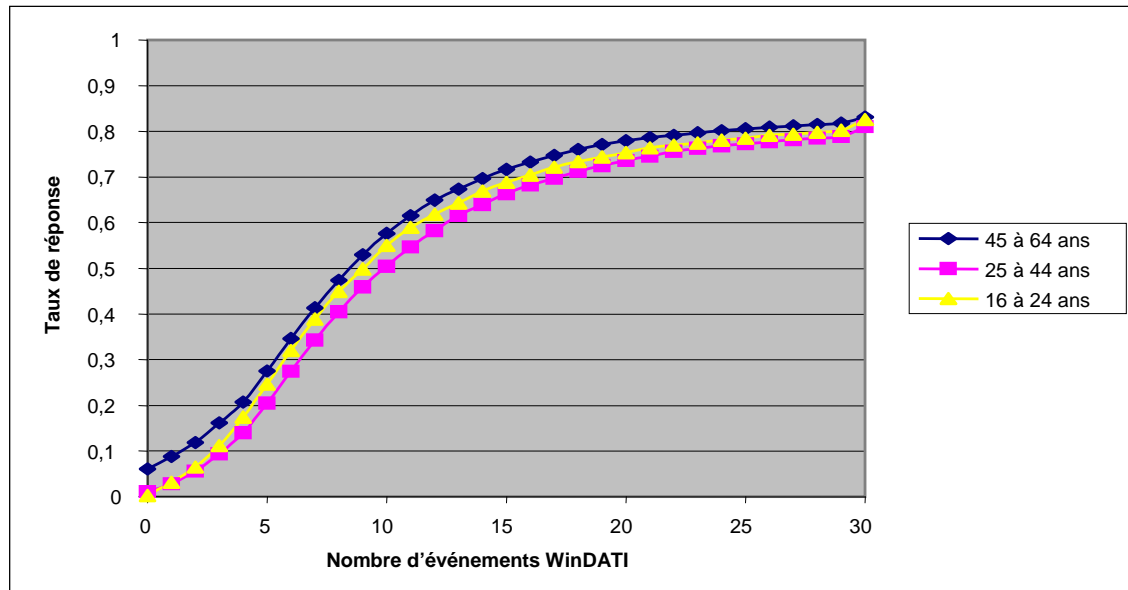
La figure 3.2-2 montre le taux de réponse cumulé ventilé selon la citoyenneté, et la figure 3.2-3 le taux de réponse cumulé ventilé selon le groupe d'âge.

Figure 3.2-2
Taux de réponse cumulé en fonction du nombre d'événements WinDATI selon la citoyenneté
 (Labour Force Survey, novembre 2006)



La forme de la courbe est la même qu'à la figure 3.2-1 (taux global). Le taux de réponse cumulé selon la citoyenneté ne nous dit pas grand-chose, à part que le taux de non-réponse est plus élevé chez les non-Suédois, fait qui est déjà bien connu.

Figure 3.2-3
Taux de réponse cumulé en fonction du nombre d'événements WinDATI selon le groupe d'âge
 (Labour Force Survey, novembre 2006)



La forme de la courbe se répète encore. Comme dans les deux figures qui précèdent, le taux de réponse se stabilise après 20 événements WinDATI et atteint un niveau de 70 % après 15 événements.

3.3 Effet sur le biais de non-réponse

Pour estimer le biais de non-réponse, nous utilisons des variables de registre. Quand nous apparions une variable de registre avec l'échantillon complet, nous obtenons des observations concernant cette variable de registre arbitraire pour les répondants ainsi que les non-répondants. Nous pouvons alors estimer le biais de non-réponse comme il suit.

Calculer les estimations en se fondant sur les répondants, \hat{Y}_r .

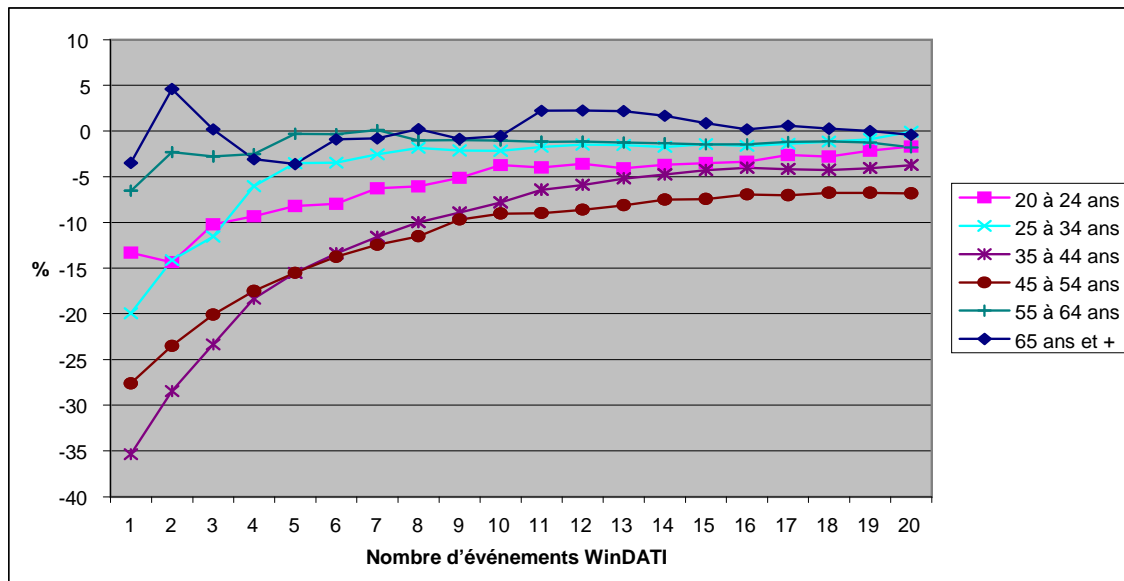
Calculer les estimations en se fondant sur les répondants ainsi que les non-répondants, \hat{Y}_s .

$$\text{Biais relatif estimé (\%)} = \frac{(\hat{Y}_r - \hat{Y}_s)}{\hat{Y}_s} \cdot 100 \text{ d'après la LFS.}$$

En fait, nous estimons le biais de la variable du registre et non de la variable d'enquête réelle. Cependant, nous pouvons trouver des variables de registre qui sont liées à la variable d'enquête d'intérêt. Ainsi, dans la LFS, la variable « personne occupée » du registre (basée sur la déclaration de revenus) est étroitement liée à cette variable dans la LFS. Mais dans ce cas, nous utilisons la variable « personne occupée » du registre comme information auxiliaire dans la procédure d'estimation d'un estimateur par calage. Donc, nous pourrions dire que la variable de registre est déjà employée. L'analyse sera contaminée si nous utilisons une variable de registre qui est déjà employée dans la procédure d'estimation.

Dans l'exemple qui suit, le salaire annuel d'après le registre suédois de l'impôt correspond à \hat{Y} dans la formule susmentionnée. Dans une procédure post-enquête, nous avons apparié HF 2006 (échantillon complet) aux données du registre suédois de l'impôt et aux données (WinDATI) au niveau de l'unité d'échantillonnage.

Figure 3.3-1
Salaire annuel – Estimations du biais relatif (%) en fonction du nombre d'événements WinDATI selon le groupe d'âge (Household Finances 2006)



La figure 3.3-1 illustre le biais relatif pour divers groupes d'âge. Nous voyons qu'il est presque négligeable tout le long de la courbe bleue (≥ 65 et plus). Les personnes de 65 ans et plus ont un salaire nul ou très faible, ce qui explique probablement ce résultat. La courbe est comparable pour les personnes de 55 à 64 ans (courbe verte). En ce

qui concerne les groupes d'âge, nous constatons qu'il existe un biais négatif important au début de la collecte des données et qu'il diminue graduellement en fonction du nombre d'événements WinDATI.

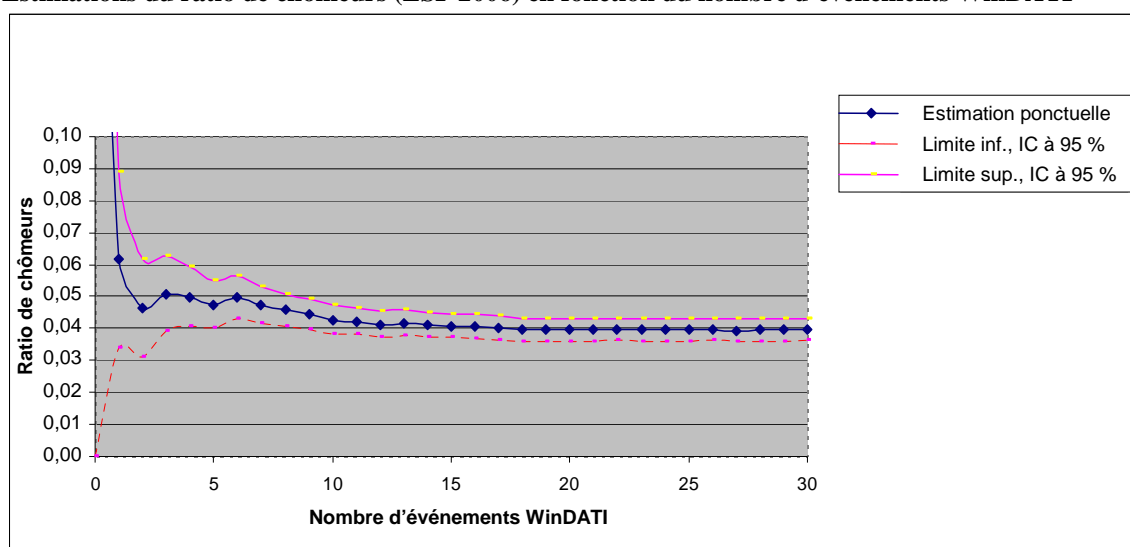
L'écart entre les divers groupes d'âge semble négligeable après 15 événements WinDATI. Une autre étude confirme ces résultats. Isaksson, Lundquist et Thorburn (2008) ont estimé le biais relatif du salaire annuel dans le cas de la LFS. Leurs résultats montrent que le biais relatif se stabilise après 15 événements WinDATI.

3.4 Effet du nombre d'événements WinDATI sur un paramètre réel d'enquête

Comme nous l'avons mentionné à la section 3.3, l'estimation du biais de la variable de registre en tant qu'indicateur du biais des variables d'enquête présente un intérêt pour l'analyse de l'effet du nombre d'événements WinDATI sur un paramètre d'enquête réel. À la figure 3.4-1, nous illustrons les estimations du ratio de chômeurs (LSF) en fonction du nombre d'événements WinDATI.

Figure 3.4-1

Estimations du ratio de chômeurs (LSF 2006) en fonction du nombre d'événements WinDATI



La figure 3.4-1 donne les estimations du ratio de chômeurs ainsi que les limites d'un intervalle de confiance (95 %). Les estimations fondées sur un petit nombre d'événements WinDATI sont à peine valides, car elles sont fondées sur un très petit nombre d'observations. Par contre, pour l'étude d'une variable d'enquête réelle aussi, l'estimation paraît stable après 15 événements WinDATI.

4. Une expérience en vue de réduire le nombre de tentatives d'appel

Typiquement, Statistics Sweden tire l'échantillon d'une enquête auprès des particuliers ou des ménages du registre de la population. Pour chaque personne échantillonnée, les intervieweurs possèdent le nom, l'adresse postale, l'âge et le sexe. En utilisant le numéro d'identification personnel, il est également possible d'effectuer un appariement avec d'autres variables, telles que le niveau d'études, la profession et le revenu. Comme nous l'avons mentionné plus haut (section 2), Statistics Sweden recourt à deux équipes d'intervieweurs : les intervieweurs sur place et les intervieweurs par ITAO. Dans certaines grandes enquêtes, les deux équipes participent à la collecte des données. Des études antérieures ont montré que les intervieweurs sur place qui utilisent l'information auxiliaire réussissent à obtenir un taux de réponse personnel plus élevé (Japac et Lundquist, 2000). Les intervieweurs sur place ont élaboré leurs propres stratégies de prise de contact, tandis que les intervieweurs par ITAO se servent d'un ordonnanceur d'appels automatique qui n'utilise pas l'information auxiliaire ni celle provenant des vagues antérieures des enquêtes par panel.

La régression logistique exécutée par Japac et Lundquist (2000) a confirmé que l'effet des variables de registre diminue après les premières tentatives d'appel. Les variables du processus sont celles qui jouent le rôle le plus important dans l'explication de la probabilité de réponse. Dans Lundquist et coll. (2007), nous nous sommes concentrés sur les facteurs qui pourraient expliquer la probabilité de répondre à la première et à la deuxième tentative d'appel. Nous avons examiné les différences et les similarités des stratégies de prise de contact (en nous basant sur les données observées) adoptées pour deux enquêtes, la LSF et la HF. La LSF est une enquête auprès des particuliers avec renouvellement de panel (huit vagues) et la HF est une enquête auprès des ménages ne comprenant qu'une seule vague.

Une régression logistique nous a permis de constater certaines similarités entre les deux enquêtes. Les appels en soirée (de 17 h à 22 h) et les appels à des personnes de plus de 65 ans sont des facteurs qui accroissent le taux de réponse. Les personnes qui vivent dans les grandes villes font baisser le taux de réponse. Dans le cas de la LSF, les appels effectués du lundi au jeudi produisent un taux de réponse plus élevé, mais cela n'a pas été observé dans le cas de la HF. (Ce résultat pourrait refléter la façon de travailler des intervieweurs. Les mêmes intervieweurs ont participé aux deux enquêtes.) Le mode d'occupation du logement (locataire ou propriétaire) a un effet dans le cas de la HF, mais non de la LSF. Par contre, le revenu a un effet dans la LSF, mais non dans la HF. Il est plus facile de trouver des facteurs significatifs dans le cas de la LSF que dans celui de la HF, peut-être parce que, dans la LSF, les répondants sont les personnes échantillonnées. Au moment de la deuxième tentative d'appel, les similarités générales demeurent significatives, même si les effets sont un peu plus faibles dans le cas de la HF, situation due à une différence de mode de collecte des données entre les deux enquêtes. L'une des conclusions est que nous devons élaborer des stratégies différentes pour les enquêtes.

Une analyse plus approfondie des données concernant la LSF, en appliquant des méthodes de recherche opérationnelle aux probabilités de réponse estimées d'après la régression logistique, a produit un taux de réponse optimal estimé pour le premier appel, sachant la répartition des ressources des intervieweurs durant la journée. La solution consiste à utiliser uniquement le quart de soirée pour la première tentative d'appel. Selon le modèle, cette stratégie permettrait d'accroître le taux de réponse de 10 % comparativement à la stratégie courante.

Durant le premier trimestre de 2008, nous avons lancé les travaux de conception d'une expérience intégrée dans la LSF. Cette expérience a pour but de déterminer s'il est possible de réduire le temps consacré par les intervieweurs aux tentatives d'appel. Une ébauche de plan d'expérience a été produite conformément au guide sur les expériences utilisé à Statistics Sweden. (L'élaboration d'une stratégie lorsqu'on travaille avec des expériences intégrées est décrite dans Karlberg et coll., 2002.)

L'expérience sera effectuée par le groupe central d'intervieweurs par ITAO. L'objectif est de planifier le moment des deux premiers appels dans le groupe expérimental (le groupe témoin travaillera comme à l'ordinaire). Pour le nouveau panel, les deux appels devront être faits en soirée, du lundi au jeudi; pour les vagues 2 à 8 (7/8 de l'échantillon), nous savons déjà quand l'interview précédente a eu lieu. Cette information, conjuguée à celle sur le moment approprié pour l'interview suivante (question supplémentaire) est utilisée automatiquement par l'ordonnanceur d'appel pour affecter l'échantillon à une tranche de temps et à une journée particulières.

5. Sommaire

Nous avons étudié les effets de l'effort de collecte, exprimé en nombre d'événements WinDATI, utilisé dans la LSF et dans la HF sur le taux de non-réponse et sur le biais de non-réponse. Notre analyse a montré que :

- le taux de non-réponse semble se stabiliser après 20 événements WinDATI dans le cas de la Labour Force Survey et après 15 événements WinDATI dans celui de la Household Finances;
- le biais de non-réponse semble se stabiliser après 15 événements WinDATI dans le cas de la Labour Force Survey et après 10 à 15 événements WinDATI dans le cas de la Household Finances;
- les estimations (globales) du ratio de chômeurs semblent se stabiliser après 15 événements WinDATI.

Les résultats indiquent clairement qu'il est possible de réduire le nombre courant d'événements WinDATI (tentatives d'appel) sans accroître le biais de non-réponse. Malgré cela, à l'heure actuelle, Statistics Sweden n'est

pas prêt à adopter une stratégie de conception adaptative pour la collecte des données. D'autres analyses sont nécessaires. Par exemple, dans le cas de la LSF, qui fournit d'importantes données d'entrée pour les Comptes nationaux et d'autres statistiques secondaires importantes, nous devons approfondir l'analyse pour certains domaines d'intérêt. Il ne saurait être question de compromettre les statistiques essentielles.

En outre, nous avons besoin de parodonnées plus nombreuses et de meilleure qualité. À titre d'exemple, mentionnons :

- établir un lien entre l'événement WinDATI et le numéro de téléphone (numéro au domicile/numéro de téléphone cellulaire/numéro au travail);
- obtenir des données sur le temps d'intervieweur selon le code d'événement WinDATI;
- créer une base de parodonnées indépendante du système WinDATI;
- produire des codes d'événements WinDATI spéciaux pour les enquêtes par panel.

La prochaine étape des travaux consistera à exécuter l'expérience décrite à la section 4.

Bibliographie

- Isaksson, A., Lundquist, P. et Thorburn, D. (2008). Optimalt antal kontaktförsök i en telefonundersökning, *Research and Development*, 2008:1, Statistics Sweden [en suédois].
- Japéc, L. et Hörngren J. (2005). Effects of Field Efforts on Nonresponse Bias and Costs in the Swedish Labour Force Survey, *Quality Issues in Interview Surveys – Some Contributions*, (Thèse de doctorate de L. Japéc), Department of Statistics, Stockholm University, Article IV.
- Japéc, L. et Lundquist, P. (2000). Bortfallet - påverkas det av intervjuarnas attityder och strategier?, rapport inédit, Stockholm, Suède; Statistics Sweden. [en suédois]
- Karlberg, M., Björnram, A., Boynton, I-M., Göransson, B. et Lundquist, P. (2002). Embedded Experiments in Sample Surveys at Statistics Sweden: The Development of an Experimentation Manual. *Proceedings of the Section on Survey Research Methods*, American Statistical Association.
- Lundquist, P. et coll. (2007). Utveckling för system av kontaktstrategier i intervjuundersökningar med individer och hushåll Delrapport I, rapport inédit, Stockholm, Suède; Statistics Sweden. [en suédois]
- Westling, S. (2008). Utveckling för system av kontaktstrategier i intervjuundersökningar med individer och hushåll Delrapport II, rapport inédit, Stockholm, Suède; Statistics Sweden. [en suédois]