

## Article

# Optimiser la qualité de la réponse au moyen de plans de collecte adaptatifs

par Barry Schouten, Melania Calinescu et Annemieke Luiten

Juin 2013



## Comment obtenir d'autres renseignements

Pour toute demande de renseignements au sujet de ce produit ou sur l'ensemble des données et des services de Statistique Canada, visiter notre site Web à [www.statcan.gc.ca](http://www.statcan.gc.ca).

Vous pouvez également communiquer avec nous par :

**Courriel** à [infostats@statcan.gc.ca](mailto:infostats@statcan.gc.ca)

**Téléphone** entre 8 h 30 et 16 h 30 du lundi au vendredi aux numéros sans frais suivants :

- Service de renseignements statistiques 1-800-263-1136
- Service national d'appareils de télécommunications pour les malentendants 1-800-363-7629
- Télécopieur 1-877-287-4369

## Programme des services de dépôt

Service de renseignements 1-800-635-7943  
Télécopieur 1-800-565-7757

## Comment accéder à ce produit

Le produit n° 12-001-X au catalogue est disponible gratuitement sous format électronique. Pour obtenir un exemplaire, il suffit de visiter notre site Web à [www.statcan.gc.ca](http://www.statcan.gc.ca) et de parcourir par « Ressource clé » > « Publications ».

## Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site [www.statcan.gc.ca](http://www.statcan.gc.ca) sous « À propos de nous » > « Notre organisme » > « Offrir des services aux Canadiens ».

Publication autorisée par le ministre responsable de Statistique Canada

© Ministre de l'Industrie, 2013

Tous droits réservés. L'utilisation de la présente publication est assujettie aux modalités de l'entente de licence ouverte de Statistique Canada (<http://www.statcan.gc.ca/reference/licence-fra.html>).

This publication is also available in English.

## Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population du Canada, ses entreprises, ses administrations et les autres établissements. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques exactes et actuelles.

## Signes conventionnels

Les signes conventionnels suivants sont employés dans les publications de Statistique Canada :

- . indisponible pour toute période de référence
- .. indisponible pour une période de référence précise
- ... n'ayant pas lieu de figurer
- 0 zéro absolu ou valeur arrondie à zéro
- 0<sup>s</sup> valeur arrondie à 0 (zéro) là où il y a une distinction importante entre le zéro absolu et la valeur arrondie
- P provisoire
- r révisé
- X confidentiel en vertu des dispositions de la *Loi sur la statistique*
- E à utiliser avec prudence
- F trop peu fiable pour être publié
- \* valeur significativement différente de l'estimation pour la catégorie de référence ( $p < 0,05$ )

# Optimiser la qualité de la réponse au moyen de plans de collecte adaptatifs

Barry Schouten, Melania Calinescu et Annemieke Luiten<sup>1</sup>

## Résumé

Dans la plupart des enquêtes, le même traitement est réservé à toutes les unités échantillonnées et les mêmes caractéristiques de plan de collecte des données s'appliquent à toutes les personnes et à tous les ménages sélectionnés. Le présent article explique comment les plans de collecte des données peuvent être adaptés en vue d'optimiser la qualité, étant donné des contraintes de coûts. Ces types de plans sont appelés plans de collecte adaptatifs. Les éléments fondamentaux de ce genre de plans sont présentés, discutés et illustrés au moyen de divers exemples.

Mots clés : Coûts d'enquête; erreurs d'enquête; non-réponse; plan de collecte des données adaptatif.

## 1 Introduction

Dans la plupart des enquêtes, le même traitement est réservé à toutes les unités échantillonnées et les mêmes caractéristiques de plan de collecte des données s'appliquent à toutes les personnes et à tous les ménages sélectionnés. Lorsque des données auxiliaires sont disponibles sous forme de données de registre ou d'observations faites par les intervieweurs, les plans de collecte peuvent être adaptés afin d'optimiser les taux de réponse, de réduire la sélectivité de la non-réponse ou, plus généralement, d'améliorer la qualité. Bien qu'une terminologie générale fasse défaut dans la littérature, ce genre de plans sont habituellement appelés plans de collecte adaptatifs.

Le présent article a pour objectif de décrire les éléments fondamentaux des plans de collecte adaptatifs, de systématiser ces plans en fournissant un cadre mathématique, d'illustrer la possibilité qu'ils ont d'améliorer l'efficacité de la collecte des données, et de propager leur utilisation dans la pratique des sondages.

---

1. Barry Schouten, Statistics Netherlands et Université Utrecht, PO Box 24500, 2490HA La Haye, Pays-Bas. Courriel : [jg.schouten@cbs.nl](mailto:jg.schouten@cbs.nl); Melania Calinescu, VU University Amsterdam, Department of Mathematics, De Boelelaan 1081, 1081HV Amsterdam. Courriel : [melania.calinescu@vu.nl](mailto:melania.calinescu@vu.nl); Annemieke Luiten, Statistics Netherlands, PO Box 4481, 6401CZ Heerlen, Pays-Bas. Courriel : [a.luiten@cbs.nl](mailto:a.luiten@cbs.nl).

Les plans de collecte adaptatifs reposent sur le principe que différentes personnes ou différents ménages peuvent recevoir un traitement différent. Ces traitements sont définis avant le début de l'enquête, mais peuvent aussi être mis à jour au moyen de renseignements obtenus durant la collecte des données. Autrement dit, l'attribution des traitements se fonde sur des données liées à l'échantillon de l'enquête et sur des paradonnées. Ces dernières sont des données sur le processus de collecte des données, comme les observations des intervieweurs au sujet du quartier, du logement ou des répondants, ou sur le rendement des intervieweurs. Dans le présent article, le terme paradonnées est utilisé dans son sens le plus large pour désigner des données qui sont observées durant la collecte et qui renseignent sur le comportement de réponse des personnes ou des ménages échantillonnés.

Une introduction générale aux plans de collecte adaptatifs est donnée dans Wagner (2008). Ces plans adaptatifs tirent leur origine de la littérature sur les statistiques médicales, où des traitements distincts sont attribués à l'avance à différents groupes de patients, mais dépendent aussi de la réponse des patients, autrement dit de mesures faites durant la collecte des données. Voir, par exemple, Heyd et Carlin (1999), Murphy (2003) et Zajonc (2012).

Un cas particulier de plan de collecte adaptatif est le plan de collecte « réactif » (*responsive survey design*). Les plans de collecte « réactifs » ont été introduits par Groves et Heeringa (2006). Comme les plans de collecte adaptatifs généraux, les plans de collecte « réactifs » peuvent comprendre l'application de caractéristiques de plan de collecte différentielles aux unités échantillonnées. La principale distinction tient au fait que, dans le cas des plans « réactifs », les caractéristiques de plan de collecte ou traitements prometteurs et efficaces sont identifiées durant la collecte des données. Pour cela, on subdivise la collecte en plusieurs phases de conception. Une nouvelle phase emploie les résultats de contrastes randomisés entre unités échantillonnées aux phases antérieures pour faire la distinction entre les traitements efficaces et inefficaces, et pour déterminer les coûts associés aux traitements. Les contrastes randomisés sont les différences de taux de réponse entre les sous-populations pour des caractéristiques de plan de collecte

attribuées aléatoirement. Voir, par exemple, Mohl et Laflamme (2007), Laflamme et Karaganis (2010), Phillips et Tabuchi (2009) et Peytchev, Riley, Rosen, Murphy et Lindblad (2010). L'affectation des caractéristiques de plan de collecte doit se faire de façon que chaque phase atteigne sa capacité de phase, qui est le compromis optimal entre la qualité et les coûts. Le recours à des plans de collecte adaptatifs est motivé par des conditions d'enquête où l'on en sait peu à l'avance sur l'échantillon et où les données historiques fournissent à l'avance peu de renseignements au sujet de l'efficacité des traitements. Dans ces conditions, plusieurs phases sont nécessaires et les plans de collecte « réactifs » sont pratiques. Toutefois, si l'on considère la deuxième phase et les phases de conception de niveau plus élevé des plans de collecte « réactifs », le point de départ est le même que pour les enquêtes où l'on possède une information préalable importante au sujet des unités échantillonnées ou pour une enquête qui est répétée plusieurs fois. La seule distinction est que, durant les phases antérieures du plan de collecte, une partie de l'échantillon a déjà répondu. Dans le présent article, nous supposons que des données historiques sont disponibles, que les traitements efficaces ont été déterminés à l'avance et qu'il est spécifié que des données appariées et des paradosées seront utilisées pour adapter le plan de collecte.

Quelles sont les nouveautés dans le présent article ? Notre contribution est triple. Premièrement, nous établissons un cadre mathématique général pour optimiser la qualité de la réponse étant donné des contraintes de coûts. Deuxièmement, nous affectons explicitement différentes caractéristiques de plan de collecte à différentes unités échantillonnées à l'intérieur de ce cadre. Troisièmement, nous proposons d'optimiser les indicateurs de qualité pour l'erreur due à la non-réponse. Les deux derniers éléments ne sont, en soi, pas entièrement nouveaux. Des plans de collecte adaptatifs simples ont déjà été appliqués; par exemple, dans l'Enquête sur la population active des Pays-Bas, ni la collecte en ligne ni la collecte par téléphone n'est utilisée pour les grands ménages, et la réponse par personne interposée n'est permise que si elle est fournie par un membre du noyau du ménage. Les efforts en vue d'optimiser les plans d'enquête

afin de tenir compte de l'erreur due à la non-réponse remontent au moins à Hartley et Monroe (1979). En outre, il existe une littérature abondante sur l'optimisation du moment où sont faits les appels et du nombre de prises de contact dans les enquêtes avec intervieweur; par exemple, Kulka et Weeks (1988), Greenberg et Stokes (1990) et Kalsbeek, Botman, Massey et Liu (1994). L'aspect nouveau ici est le regroupement de tous ces éléments dans un cadre mathématique général qui fait abstraction des caractéristiques de plan de collecte uniques et qui permet d'appliquer des indicateurs de qualité généraux. Les principales raisons motivant la proposition de ce genre de cadre sont les pressions importantes en vue de réduire les coûts d'enquête et l'importance croissante d'Internet comme mode de collecte des données d'enquête. La collecte en ligne est caractérisée par un écart qualité-coûts important – elle est bon marché, mais les taux de réponse sont faibles – et elle ne possède pas les mêmes propriétés de mesure que les modes de collecte par intervieweur. Par conséquent, la collecte en ligne pose un défi en ce qui concerne le compromis entre la qualité et les coûts. Alors que les compromis entre les diverses erreurs étudiés au moment de la conception des enquêtes ont fait couler beaucoup d'encre, par exemple Lyberg, Biemer, Collins, De Leeuw, Dippo, Schwarz et Trewin (1997) et Dillman (2007), l'étude et la mise en œuvre de caractéristiques de plan de collecte différentielles en pratique se limitent encore, étonnamment, à quelques cas. Dans le présent article, nous espérons offrir un tremplin pour de futures études et discussions des plans de collecte adaptatifs.

À la section 2, nous décrivons la théorie et les concepts qui sous-tendent les plans de collecte adaptatifs. À la section 3, nous présentons un exemple fondé sur des données d'enquête virtuelles et, à la section 4, nous discutons d'une étude en simulation portant sur des données d'enquête réelles. Enfin, à la section 5, nous concluons par un résumé et une discussion.

## **2 Que sont les plans de collecte adaptatifs ?**

### **2.1 Plans de collecte adaptatifs en général**

À la présente section, nous exposons le cadre mathématique pour les plans de collecte

adaptatifs. Aux sections suivantes, nous mettons en relief et élaborons les composantes de ce cadre.

Soit la population composée des unités  $k = 1, 2, \dots, N$ . La population d'intérêt peut être constituée de toutes les unités d'une population, mais aussi de tous les membres recrutés d'un panel. À chaque unité sera attribuée une stratégie  $s$  provenant de l'ensemble de stratégies considérées  $S = \{\varphi, s_1, s_2, \dots, s_M\}$ . L'ensemble de stratégies d'enquête  $S$  comprend explicitement la stratégie vide  $\varphi$ . La stratégie vide signifie qu'aucune action n'a lieu, c'est-à-dire que l'unité de population n'est pas échantillonnée. Il s'agit du cadre le plus général. En pratique, on sépare souvent le plan d'échantillonnage de la répartition des stratégies et l'on considère l'échantillon comme donné et fixe. Cependant, on peut inclure la décision d'échantillonner une unité explicitement dans l'affectation globale des ressources.

En général, une stratégie  $s$  est un ensemble spécifié de caractéristiques de plan de collecte qui peut comprendre une série de traitements, où un traitement donné n'est appliqué que quand tous les traitements antérieurs ont échoué. Certaines de ces caractéristiques peuvent être séquentielles, telles que le type de mode de contact et le type de mode de collecte des données, mais les caractéristiques peuvent aussi décrire divers aspects d'un plan de collecte des données. Voici des exemples de stratégies :

$s_1$  = (lettre envoyée à l'avance 1, questionnaire en ligne, un appel);

$s_2$  = (lettre envoyée à l'avance 1, questionnaire en ligne, pas de rappel);

$s_3$  = (lettre envoyée à l'avance 2, questionnaire administré par ITAO, maximum de 6 tentatives d'appel);

$s_4$  = (lettre envoyée à l'avance 2, questionnaire administré par ITAO, maximum de 15 tentatives d'appel).

On trouve dans la littérature de nombreuses suggestions et évaluations de caractéristiques de plans de collecte; par exemple, Groves et Couper (1998) et Groves, Dillman, Eltinge et Little (2002). Nous renvoyons le lecteur à De Leeuw (2008) pour une discussion des modes de collecte

des données, à Dillman (2007) et De Leeuw, Callegaro, Hox, Korendijk et Lensvelt-Mulders (2007) pour une élaboration des lettres envoyées à l'avance et des rappels, à Wagner (2008) pour une discussion du protocole de prise de contact, à Barón, Breunig, Cobb-Clark, Gørgens et Sartbayeva (2009) pour une revue des primes d'incitation, à Kersten et Bethlehem (1984), Cobben (2009) et Lynn (2003) pour la recherche sur les questionnaires condensés, à Moore (1988) pour une discussion de la déclaration par personne interposée et à Cobben (2009) pour un exemple de tâches d'intervieweur.

Nous supposons dans le présent article que l'ensemble de stratégies  $S$  est connu et fixé une fois que l'attribution des stratégies a débuté. L'ensemble de stratégies peut être déterminé en se basant sur des données d'enquête historiques, l'expérience ou des études pilotes. Voir Schouten, Luiten, Loosveldt, Beullens et Kleven (2010) et Schouten, Shlomo et Skinner (2011) pour des lignes directrices et des exemples sur la façon de construire un ensemble de stratégies.

À chaque unité de population  $k$  est associé un vecteur de covariables  $X_k = (X_{k1}, X_{k2}, \dots, X_{kp})^T$ .  $X_k$  contient les caractéristiques connues avant le début de la collecte des données et avant l'attribution des stratégies. Les covariables doivent, par conséquent, être disponibles dans des registres ou des fichiers de données administratives qui peuvent être appariés à la base de sondage, ou directement dans la base de sondage. En plus de ces caractéristiques générales, il peut exister pour l'unité  $k$  un deuxième vecteur de covariables  $\tilde{X}_k = (\tilde{X}_{k1}, \tilde{X}_{k2}, \dots, \tilde{X}_{kq})^T$  qui reflète les caractéristiques observées durant la collecte des données pour les unités de population échantillonnées. Ces caractéristiques sont appelées paradonnées ou données sur le processus, parce qu'elles sont recueillies durant le processus de collecte des données par les intervieweurs et par le personnel de la collecte des données. Toutefois, en plus de la définition habituelle des paradonnées comme étant de l'information au sujet du processus, dans le contexte des plans de collecte adaptatifs,  $\tilde{X}_k$  comprend des observations au sujet de la personne ou du ménage échantillonné. Le vecteur  $X_k$  peut contenir, par exemple, le sexe, l'âge, le type de ménage ou le niveau de scolarité. Le vecteur  $\tilde{X}_k$  peut

contenir, par exemple, l'évaluation faite par l'intervieweur de la propension à répondre ou de la propension à être contacté, de l'état du logement ou du quartier, et de la présence d'un intercom.  $\tilde{X}_k$  est délibérément limité aux observations au sujet de l'échantillon qui permettent de choisir les caractéristiques de plan de collecte différentes. Il ne contient pas les valeurs des caractéristiques du plan de collecte proprement dites, tel que l'intervieweur qui a été affecté à l'adresse.

La différence importante entre  $X_k$  et  $\tilde{X}_k$  tient au niveau de disponibilité des données.  $\tilde{X}_k$  est connu uniquement pour les unités qui sont échantillonnées et ne peut pas être utilisé pour distinguer des sous-populations a priori. Soit  $q(x)$  la distribution de  $X_k$  dans la population et  $q(\tilde{x}, x)$  la distribution conjointe de  $X_k$  et  $\tilde{X}_k$  dans l'échantillon. En outre,  $q(\tilde{x} | x)$  désigne la distribution conditionnelle dans l'échantillon. On suppose que  $q(x)$  et  $q(\tilde{x}, x)$  sont connus d'avance. Dans les situations où l'on ne peut apparier que peu de données, voire aucune, l'attribution des stratégies doit être fondée entièrement sur les observations faites durant la collecte des données.

Les plans de collecte adaptatifs qui attribuent les stratégies en se basant sur les caractéristiques de la population pour lesquelles des données sont disponibles dans un registre ou dans la base de sondage sont considérés comme étant *statiques*, tandis que les plans de collecte adaptatifs qui attribuent les stratégies en se basant (en outre) sur des paradonnées sont appelés *dynamiques*. Il est important de souligner que les plans tant statiques que dynamiques comprennent un ensemble de stratégies qui est fixé avant le début de la collecte des données. Toutefois, dans le cas des plans dynamiques, on ne sait pas à l'avance quelles stratégies seront attribuées aux unités individuelles, parce que le choix de la stratégie dépend des données qui sont observées durant la collecte.

Soit  $\rho(x, s)$  la *propension à répondre* d'une unité présentant la caractéristique  $X = x$  et à laquelle est attribuée la stratégie  $s$ . On suppose que l'information sur  $\rho(x, s)$  est disponible dans les données historiques, c'est-à-dire des données qui proviennent de versions antérieures de la

même enquête, d'enquêtes possédant un thème et un plan similaires, ou de la phase initiale du plan de collecte. Manifestement, la propension à répondre attendue doit être une estimation qui s'approche de la propension réelle. Nous reviendrons sur cette composante essentielle des plans de collecte adaptatifs à la section 2.4.

Le *coût prévu* de l'affectation de la stratégie  $s$  à une unité ayant la caractéristique  $X = x$  est désigné par  $c(x, s)$ . Il s'agit d'une composante de coût individuel. Les coûts d'enquête sont décrits dans la littérature comme comprenant de nombreuses composantes dont certaines sont des coûts indirects et d'autres, des coûts individuels; voir, par exemple, Groves (1989). Les fonctions de coût sont discutées à la section 2.3.

Soit  $p(s | x)$  la *probabilité d'affectation* d'une unité de population ayant la caractéristique  $x$  à la stratégie  $s$ , et soit  $p(s | x, \tilde{x})$  la probabilité d'affectation à cette stratégie sachant que des parodontées  $\tilde{x}$  sont également observées. Les conditions suivantes doivent être vérifiées :

$$0 \leq p(s | x) \leq 1, \quad 0 \leq p(s | x, \tilde{x}) \leq 1 \quad (2.1)$$

$$\sum_s p(s | x) = 1, \quad \sum_s p(s | x, \tilde{x}) = 1, \quad (2.2)$$

c'est-à-dire que toutes les unités sont affectées à une stratégie. En général, les probabilités d'affectation peuvent avoir des valeurs comprises entre 0 et 1. Autrement dit, les sous-population ayant les mêmes scores sur  $x$  et  $\tilde{x}$  peuvent être affectées (aléatoirement) à des stratégies différentes. Par exemple, on pourrait n'approcher de nouveau au suivi qu'une partie seulement des non-répondants. Permettre que les probabilités d'affectation soient comprises entre 0 et 1 donne plus de souplesse pour atteindre les niveaux de qualité ou respecter les contraintes de coûts. Dans la suite de l'exposé,  $p$  désigne la matrice des probabilités d'affectation, c'est-à-dire  $p = \{p(s_j | x, \tilde{x})\}_{1 \leq j \leq M, x, \tilde{x}}$  et contient les variables décisionnelles pour l'optimisation.

Les propensions à répondre  $\rho_x$  peuvent être calculées en se basant sur les propensions à répondre selon la stratégie et sur les probabilités d'affectation par

$$\rho_x(x) = \sum_{s \in S} \sum_{\tilde{x}} q(\tilde{x} | x) p(s | x, \tilde{x}) \rho(s, x, \tilde{x}). \quad (2.3)$$

Les stratégies, les covariables, les propensions à répondre, les fonctions de coûts et les probabilités d'affectation constituent les éléments fondamentaux des plans de collecte adaptatifs. Partant de ces éléments, on peut formuler le problème d'optimisation du plan de collecte adaptatif. Toutefois, deux éléments manquent encore, à savoir une fonction de qualité et une fonction de coût global. Soit  $Q(p)$  un indicateur de qualité et  $C(p)$  une évaluation du coût total. La dépendance des deux fonctions à l'égard des probabilités d'affectation est soulignée, car les probabilités sont les variables décisionnelles dans l'optimisation.

Le problème d'optimisation peut maintenant être formulé sous la forme

$$\max_p Q(p) \text{ sachant que } C(p) \leq C_{\max} \quad (2.4)$$

ou sous la forme

$$\min_p C(p) \text{ sachant que } Q(p) \geq Q_{\min}, \quad (2.5)$$

où  $C_{\max}$  représente le budget d'une enquête et  $Q_{\min}$ , les contraintes de qualité minimales. Les problèmes (2.4) et (2.5) sont appelés problèmes d'optimisation double, quoique leurs solutions puissent différer selon les contraintes de qualité et de coût.

Il est important de souligner que l'optimisation de la qualité ou du coût n'est effectuée qu'une seule fois, avant que débute la collecte des données, et qu'elle n'est pas répétée durant la collecte. Donc, c'est la stratégie, et non pas l'optimisation proprement dite, que l'on adapte à l'unité de population et, dans le cas d'un plan dynamique, aux paradonnées au sujet de l'unité en question. L'optimisation est fondée sur des données d'enquête historiques qui comprennent les paradonnées devenues disponibles dans une enquête. La densité de probabilité conjointe  $q(x, \tilde{x})$ , les probabilités de réponse  $\rho(x, \tilde{x}, s)$  et la fonction de coût  $c(x, \tilde{x}, s)$  sont toutes estimées d'après les données d'enquête historiques et sont considérées comme étant données. Puisqu'en pratique, les paradonnées ne deviennent disponibles que durant la collecte des données, les stratégies considérées pour les unités comprises dans une même strate  $x$  sont les mêmes jusqu'au moment où les paradonnées  $\tilde{x}$  deviennent disponibles. Par exemple, on pourrait avoir les quatre

stratégies suivantes : 1) deux tentatives d'appel téléphonique et aucun suivi, 2) deux tentatives d'appel téléphonique et un suivi avec prime d'incitation, 3) trois tentatives d'appel téléphonique et aucun suivi, et 4) trois tentatives d'appel téléphonique et un suivi avec prime d'incitation. La décision de faire deux ou trois tentatives d'appel est fondée sur  $x$ , tandis que le suivi est décidé en se servant d'une observation  $\tilde{x}$  tirée des parodonnées téléphoniques. Donc, on estime à l'avance combien d'unités seront comprises dans la strate  $(x, \tilde{x})$  et combien feront l'objet d'un suivi, mais ce n'est que quand  $\tilde{x}$  est mesuré que la stratégie complète est connue pour les unités individuelles.

## 2.2 Fonctions d'objectif de qualité

Tels qu'ils sont décrits dans la littérature, les plans de collecte adaptatifs sont habituellement axés sur l'erreur due à la non-réponse. À la présente section, nous commençons par présenter une classification générale des fonctions de qualité, puis nous passons aux fonctions de qualité pour l'erreur due à la non-réponse. En général, la concentration sur l'erreur due à la non-réponse donne une perspective trop restreinte, surtout quand le mode de collecte des données est l'une des caractéristiques du plan de collecte prise en considération dans le plan de collecte adaptatif. Ici, nous ne discutons toutefois pas explicitement des autres erreurs d'enquête, mais nous revenons à cette question dans la discussion. Nous renvoyons à Calinescu, Schouten et Bhulai (2012) pour une extension des plans de collecte adaptatifs à l'erreur de mesure et à Beaumont et Haziza (2011) pour une discussion des plans de collecte adaptatifs et de la variance de non-réponse.

### 2.2.1 Fonctions de qualité fondées sur les covariables et fondées sur l'item

Lorsque l'on optimise la qualité conformément à (2.4), les fonctions de qualité appliquent à l'échantillon pour lequel on dispose de données appariées, de parodonnées et de réponses aux items une valeur unique qui peut être interprétée et optimisée. Lorsque les coûts sont minimisés sous des contraintes de qualité comme dans (2.5), la qualité peut être multidimensionnelle (mais les fonctions de coût doivent être unidimensionnelles).

En général, on distingue deux types de fonctions de qualité, celles qui emploient des covariables provenant de données appariées et de paradonnées seulement, et celles qui emploient également les réponses concernant les variables cibles (ou items) de l'enquête. Nous les désignons comme étant *fondées sur les covariables* et *fondées sur l'item*, respectivement. Une fonction de qualité fondée sur l'item est une fonction de la distribution des réponses à un item de l'enquête et de la distribution attendue, estimée, dans la population complète, étant donné les données appariées et les paradonnées disponibles. La principale distinction entre les fonctions de qualité fondées sur les covariables et celles fondées sur l'item est que les secondes nécessitent des hypothèses. Évidemment, les réponses des non-répondants manquent. Donc, l'évaluation de la qualité doit être fondée sur les relations entre les variables cibles et les covariables telles qu'elles sont observées dans les réponses. Par conséquent, les fonctions de qualité fondées sur l'item comportent un risque découlant directement du phénomène qu'elles essaient de mesurer. Les relations entre les variables cibles et les covariables pourraient être différentes chez les non-répondants et la qualité fondée sur l'item pourrait produire une image incomplète. De surcroît, dans les enquêtes comportant de nombreuses variables cibles, différentes variables peuvent mener à des décisions différentes et des plans de collecte de données optimaux différents. Cependant, contrairement aux fonctions de qualité fondées sur des covariables, les fonctions de qualité fondées sur l'item adaptent les plans de collecte de données spécifiquement aux thèmes de l'enquête. Les fonctions de qualité fondées sur les covariables ne peuvent être reliées qu'au biais de non-réponse des covariables qui sont incluses.

### **2.2.2 Optimisation de la qualité de la réponse**

Pour commencer, nous décrivons brièvement un certain nombre de fonctions de qualité relevées récemment dans la littérature. Ensuite, nous discutons du choix d'une fonction de qualité.

La fonction de qualité fondée sur des covariables la mieux connue pour la non-réponse est le taux de réponse. Il ne s'agit pas d'une vraie fonction de qualité fondée sur des covariables en ce sens qu'il dépend de données appariées ou de paradonnées. Cependant, puisque l'indicateur de réponse 0-1 peut être considéré comme la forme la plus simple de paradonnées, nous le désignons comme une fonction de qualité fondée sur des covariables. Le taux de réponse est représenté comme la propension moyenne à répondre,

$$\text{Taux de réponse : } \quad Q(p) = \bar{p} = \sum_{x, \tilde{x}, s} q(x, \tilde{x}) p(s | x, \tilde{x}) \rho(x, \tilde{x}, s) \quad (2.6)$$

Schouten, Cobben et Bethlehem (2009) proposent deux fonctions de qualité fondées sur des covariables, l'indicateur  $R$  et une mesure qu'ils appellent le biais maximal ou le pire biais de non-réponse. Le nom donné à ce deuxième indicateur est trompeur, car il s'agit seulement d'un estimateur du biais maximal de la moyenne non corrigée des réponses, et non du biais maximal réel. Le coefficient de variation des propensions à répondre est une meilleure appellation que nous utiliserons ici. Les mesures peuvent s'écrire

$$\text{Indicateur } R : \quad Q(p) = R(\rho_Z) = 1 - 2S(\rho_Z) \quad (2.7)$$

$$\text{Coefficient de variation : } \quad Q(p) = CV(\rho_Z) = \frac{S(\rho_Z)}{\bar{p}}, \quad (2.8)$$

où la représentativité peut être évaluée par rapport aux données appariées seulement,  $Z = X$ , ou par rapport à un vecteur contenant à la fois les données appariées et les paradonnées,  $Z = (X, \tilde{X})^T$ . Les écarts-types de propension à répondre,  $S(\rho_X)$  et  $S(\rho_{X, \tilde{X}})$ , peuvent s'écrire en fonction des probabilités d'affectation des stratégies sous la forme

$$S(\rho_X) = \sqrt{\sum_x q(x) \left( \sum_{\tilde{x}, s} q(\tilde{x} | x) p(s | x, \tilde{x}) \hat{p}(x, \tilde{x}, s) - \bar{p} \right)^2} \quad (2.9a)$$

$$S(\rho_{X, \tilde{X}}) = \sqrt{\sum_{x, \tilde{x}} q(x, \tilde{x}) \left( \sum_s p(s | x, \tilde{x}) \hat{p}(x, \tilde{x}, s) - \bar{p} \right)^2}. \quad (2.9b)$$

Särndal et Lundström (2010) et Särndal (2011a et b) proposent des indicateurs dont la définition et la nature sont très similaires à (2.7) et (2.8). Ces indicateurs ont été dérivés dans la

perspective du calage et de la réponse dite équilibrée, et peuvent être utilisés pour remplacer (2.7) ou (2.8).

Un exemple de fonction de qualité fondée sur l'item pour la non-réponse est présenté par Groves et Heeringa (2006). Pour une variable cible particulière  $Y$ , Groves et Heeringa (2006) proposent le biais de non-réponse de la moyenne non corrigée des réponses,

$$\text{Biais de non-réponse estimé : } Q(p) = \frac{\text{cov}(Y, \rho_X)}{\bar{p}}, \quad (2.10)$$

où  $\text{cov}(Y, \rho_X)$  est la covariance de réponse entre la variable cible et les propensions à répondre sachant les covariables  $X$ . Elle peut s'écrire sous la forme

$$\text{cov}(Y, \rho_X) = \frac{\sum_x q(x) \rho_X(x) (\rho_X(x) - \bar{\rho}_X) (y(x) - \bar{y}_R)}{\bar{p}}, \quad (2.11)$$

avec  $\rho_X(x)$  comme dans (2.3),  $y(x)$  désignant la valeur moyenne de  $Y$  pour  $X = x$  et  $\bar{y}_R$ , la moyenne attendue des réponses. De nouveau, l'expression (2.11) peut être étendue de manière à inclure les paradonnées  $\tilde{X}$ .

Toutes les fonctions de qualité décrites à la présente section sont définies comme des paramètres de population. En pratique, elles doivent être estimées d'après des données d'enquête. Les propensions à répondre réelles  $\rho_X(x)$  doivent être remplacées par les estimateurs  $\hat{\rho}_X(x)$ , fondés sur une certaine forme de régression, et les sommations sur la population doivent être remplacées par des sommations pondérées par les poids de sondage sur l'échantillon. Nous revenons sur l'estimation des propensions à répondre à la section 2.4.

Maintenant, comment choisir une fonction de qualité ? Toutes les fonctions de qualité mentionnées ici ont pour objectif de mesurer l'effet de la non-réponse au-delà de la simple réduction de la taille de l'échantillon. Elles le font en se basant sur des covariables provenant de données appariées et de paradonnées. La raison qui justifie l'optimisation de ces fonctions de qualité est que des traces plus fortes d'erreur de non-réponse sur ces covariables pourraient

impliquer une erreur de non-réponse la plus importante sur d'autres variables également; les fonctions de qualité sont considérées comme des indicateurs de la qualité du processus plutôt que de la qualité du produit. Quoiqu'intéressante, cette conjecture requiert clairement un soutien empirique. Le choix d'une fonction de qualité pour la non-réponse doit être fondé sur l'ensemble de variables étudiées clés, les paramètres de population d'intérêt et les estimateurs qui seront utilisés. Le taux de réponse et l'indicateur R ne ciblent ni un paramètre de population ni un estimateur particulier. Le coefficient de variation est axé sur les moyennes de population, mais n'est particulier à aucune variable étudiée ni correction de la non-réponse, tandis que le biais de non-réponse estimé naît de la même perspective, mais ne s'applique qu'à une seule variable étudiée. Si une enquête comporte plusieurs variables étudiées clés, il faut éviter d'utiliser une fonction de qualité fondée sur l'item pour la non-réponse, car cela pourrait mener à des problèmes d'optimisation contradictoires. Si une enquête ne comporte qu'une seule variable clé, de bons résultats peuvent être obtenus en utilisant la fonction de qualité fondée sur l'item ou en se limitant aux covariables les plus pertinentes seulement dans des fonctions de qualité fondées sur des covariables. Dans le cas d'une enquête à usage multiple, il est, selon nous, trop contraignant de se concentrer sur un paramètre de population et un estimateur particulier, et nous privilégions l'indicateur R plutôt que toute autre fonction de qualité. Si l'on peut s'attendre à ce que les utilisateurs se concentrent sur les moyennes ou sur les totaux de population, il faut, à notre avis, donner la préférence au coefficient de variation, car il ne suppose pas l'utilisation d'une méthode d'ajustement particulière.

Cependant, l'ensemble de données appariées et de parodonnées qui est entré dans l'indicateur est encore plus important que le choix de l'indicateur proprement dit. Si une enquête ne comporte qu'une variable clé ou quelques-unes seulement, les données appariées et les parodonnées choisies peuvent, et doivent, être en rapport avec ces variables. Toutefois, si une enquête comprend une grande gamme de variables étudiées, on doit nécessairement se limiter aux variables auxiliaires qui distinguent généralement les personnes ou les ménages.

Jusqu'à présent, nous avons ignoré l'effet de la non-réponse sur la précision, alors que des exigences en matière de précision des estimations doivent être énoncées explicitement. Il existe deux options pour inclure la précision dans l'optimisation. Premièrement, on peut ajouter une contrainte supplémentaire. Le choix direct serait une contrainte concernant le nombre minimal de répondants, éventuellement pour un certain nombre de sous-groupes de population, évitant ainsi de spécifier les paramètres de population et les estimateurs. Deuxièmement, la variance due à la non-réponse peut être incluse dans la fonction de qualité proprement dite, c'est-à-dire que l'on considérerait des indicateurs de l'erreur quadratique moyenne. Cette option est proposée et élaborée par Beaumont et Haziza (2011). Sous cette seconde option, il faut de nouveau prendre en considération l'ensemble de variables étudiées clés, le paramètre de population et la stratégie d'estimation, car la précision est propre à un estimateur pour un paramètre de population donné pour une variable étudiée unique.

### **2.3 Fonctions de coût**

Les fonctions de coût sont le pendant des fonctions de qualité. Elles comprennent plusieurs composantes. Il est important de limiter la spécification des coûts aux caractéristiques du plan de collecte que l'on fait varier dans le plan de collecte adaptatif. Par exemple, si c'est la prime d'incitation que l'on fait varier pour différentes sous-populations, il n'est pas nécessaire de spécifier ni de détailler les coûts relatifs aux durées de déplacement des intervieweurs. Quand le protocole concernant le moment de la prise de contact est la caractéristique du plan de collecte que l'on adapte, les durées et les coûts de déplacement jouent manifestement un rôle prépondérant.

Si un grand nombre de caractéristiques du plan de collecte sont optionnelles, les fonctions de coût prennent des formes complexes comportant de nombreuses composantes de coûts indirects et de coûts variables. En général, les coûts variables dépendent de la taille de l'échantillon, mais non les coûts indirects. Les composantes de coûts indirects peuvent être associées au personnel de

collecte des données, à l'échantillonnage et au traitement des échantillons. Les coûts variables découlent, par exemple, de la formation des intervieweurs, de l'impression et de l'envoi par la poste des questionnaires et des rappels, du traitement des questionnaires papier, du taux de rémunération horaire et des déplacements des intervieweurs, des primes d'incitation et de l'appariement des numéros de téléphone, de l'utilisation du téléphone et des serveurs informatiques.

Dans le problème d'optimisation, on peut préciser deux composantes de coût, l'une fixe et l'autre variable. La composante variable dépend de l'affectation des unités de population aux stratégies, tandis que la composante fixe comprend tous les autres coûts. Il convient de souligner que la composante fixe diffère selon la caractéristique de collecte sur laquelle est axé le plan de collecte des données adaptatif. La fonction de coût  $C(p)$  est la somme de deux composantes

$$C(p) = C_F + C_V(p), \quad (2.12)$$

dont seulement la seconde, la composante variable, dépend des probabilités d'affectation.

En général, pour une stratégie de collecte  $s$ , les coûts  $c(\tilde{x}, x, s)$  sont associés aux unités de population provenant du groupe  $(\tilde{x}, x)$ . La fonction de coût individuelle peut être une fonction des propensions à répondre, voire même plus spécifiquement, des propensions à la prise de contact et à la participation. Par exemple, les coûts d'intervieweur dans différents protocoles concernant le moment de la prise de contact dépendent des taux de prise de contact et des sous-population sélectionnées. La fonction de coût  $c(\tilde{x}, x, s)$  est une fonction de coût relative car elle décrit uniquement la contribution de la stratégie à la composante de coût variable  $C_V(p)$

$$C_V(p) = \sum_{\tilde{x}, x, s} q(\tilde{x}, x) p(s | x, \tilde{x}) c(\tilde{x}, x, s). \quad (2.13)$$

Trois remarques s'imposent. Premièrement, l'établissement des composantes de coût fixe et variable est compliqué quand un organisme d'enquête exécute de nombreuses enquêtes en parallèle. D'une part, l'interaction entre les enquêtes rend difficile la ventilation des coûts par

enquête, surtout quand les stratégies sont taillées sur mesure. Par ailleurs, quand plusieurs enquêtes sont réalisées, certaines composantes de coût variable peuvent être désignées comme étant fixes. Ainsi, si un nombre relativement faible seulement d'unités de population sont attribuées au mode d'interview sur place, on peut supposer que les frais de déplacement restent constants, car les adresses sont regroupées avec celles d'autres enquêtes. La deuxième remarque concerne l'aspect multidimensionnel des coûts. Mis à part le budget global, on peut demander que le taux d'activité des intervieweurs soit proche de 1 au cours du temps ou qu'aucun intervieweur ne doive faire plus d'heures supplémentaires qu'un nombre fixe déterminé. La fonction de coût devient alors un vecteur et la contrainte, un vecteur de contraintes. La troisième remarque a trait à la validité des fonctions de coût. Puisque ces dernières sont difficiles à construire en pratique, l'optimisation peut parfois être trop optimiste. Il est important de surveiller de près la collecte des données et d'intégrer des indicateurs dans les stratégies.

## **2.4 Estimation des probabilités de réponse**

Après les paramètres de coût et les fonctions de qualité vient un autre élément important des plans de collecte adaptatifs, à savoir l'ensemble de propensions à répondre pour les diverses stratégies. Les données sur ces propensions à répondre doivent provenir d'enquêtes antérieures, de préférence la même enquête ou une enquête similaire. Ou bien, comme le proposent Groves et Heeringa (2006), on peut s'appuyer sur les premières phases de la collecte de données pour dériver ces propensions à répondre. Cela se fait aux dépens de l'efficacité, puisqu'une partie de l'enquête est déjà réalisée. Néanmoins, l'information recueillie directement est réinjectée dans l'enquête courante.

La littérature sur les enquêtes-ménages fournit une longue liste de modèles de réponse qui comprennent les caractéristiques du plan de collecte. Le dénominateur commun à tous ces modèles est que les propensions à répondre sont estimées en s'appuyant sur un certain nombre d'hypothèses quant à la nature réelle du mécanisme de formation des données manquantes dues à

la non-réponse. En général, ces modèles sont des simplifications. Conséquemment, les propensions à répondre attendues  $\rho(x, s)$  présentent une erreur-type, et peuvent même être elles-mêmes biaisées lorsqu'elles sont fondées sur des enquêtes similaires, mais différentes. Dans le problème d'optimisation, cette incertitude peut être prise en compte en permettant aux propensions à répondre d'être des variables aléatoires plutôt que des quantités fixes. Le caractère aléatoire demande des analyses de sensibilité, ainsi que des évaluations de la robustesse de l'optimisation permettant de mieux comprendre la variation de la qualité des coûts quand l'enquête est réalisée plusieurs fois (dans les mêmes circonstances).

## 2.5 Le problème d'optimisation

L'optimisation de (2.3) et de (2.4) peut se faire selon deux approches : une approche par tâtonnement ou une optimisation mathématique. Dans le présent article, nous nous concentrons sur un cadre et une optimisation mathématiques, mais on peut aussi être plus modeste et introduire les plans de collecte adaptatifs progressivement au moyen d'études pilotes et d'essais sur le terrain.

Les fonctions de qualité (2.6), (2.7), (2.8) et (2.10) sont toutes des fonctions des probabilités d'affectation de la stratégie  $p$ . Le taux de réponse est une fonction linéaire des probabilités d'affectation, de sorte qu'il est relativement facile de l'optimiser en se servant d'un logiciel d'optimisation standard (par exemple, le progiciel *linprog* en R ou tout autre logiciel permettant de résoudre des problèmes de programmation linéaire). Néanmoins, autant que nous sachions, en raison de la forte dimensionnalité de  $p$ , il n'existe aucune solution explicite de (2.4), même pour les problèmes linéaires. Cependant, en général, les fonctions de qualité sont des fonctions non linéaires non convexes en ce qui concerne les probabilités d'affectation et elles ne peuvent pas être optimisées sans recourir à des méthodes numériques ou de Monte Carlo. La complexité du problème augmente rapidement avec le nombre de stratégies considérées et le nombre de sous-groupes fondés sur les données appariées et les paradonnées.

Les logiciels statistiques courants contiennent des procédures ou des modules externes qui permettent de traiter les problèmes d'optimisation non linéaires, comme *nlm* ou *nlinb* en R ou *proc optmodel* en SAS. Cependant, les problèmes non linéaires non convexes nécessitent parfois de longs temps de calcul ou peuvent converger vers des optima locaux. Pour cette raison, il est recommandé d'utiliser des logiciels d'optimisation spécialisés, tels que Xpress, Baron ou AMPL.

Dans les exemples des sections 3 et 4, nous exécutons un certain nombre d'optimisations. Le problème d'optimisation de la section 3 est relativement simple; la fonction d'objectif de qualité est l'indicateur R qui est évalué sur deux sous-groupes de population. Pour les deux sous-groupes, l'optimisation peut être réécrite sous forme d'un problème de programmation linéaire. Dans le cas de l'exemple de la section 4, nous avons pu construire un algorithme qui converge vers la solution optimale en un petit nombre d'étapes. Toutes les optimisations ont été programmées en R et le code peut être obtenu sur demande.

### **3 Un plan de collecte adaptatif dynamique : réaffectation des intervieweurs dans une enquête de suivi**

À la présente section, nous donnons un exemple de plan adaptatif dynamique ayant trait à la réaffectation des intervieweurs en s'appuyant sur des observations de la propension à coopérer. L'exemple est fondé sur des propensions à répondre et des fonctions de coût hypothétiques. Les intervieweurs sont affectés aux cas qui ont refusés de répondre une fois, d'après une évaluation de la propension à répondre faite durant une première phase de l'enquête. L'évaluation est faite pour les répondants et pour les personnes qui refusent de répondre, mais elle n'est pas disponible pour les unités échantillonnées avec lesquelles un contact n'a pas été pris durant la première phase. Elle fournit un jugement de la propension qu'auront les unités échantillonnées à participer à l'enquête lorsque l'on prendra de nouveau contact avec elles. L'évaluation est faite sur une échelle de trois points : *forte*, *moyenne*, *faible*. Une propension forte signifie qu'il existe une probabilité élevée que l'unité échantillonnée réponde s'il est de nouveau pris contact avec elle.

Après une première phase de collecte des données, les résultats intermédiaires de la collecte sont évalués et les unités échantillonnées sont réparties en trois catégories, à savoir répondants, cas de refus et cas de non-contact. Un traitement différent est appliqué aux cas de refus. Les intervieweurs sont évalués en se basant sur leur rendement historique et groupés en *bons* et *moins bons* intervieweurs. Les cas de refus sont réaffectés à l'un des deux groupes d'intervieweurs. Comme il n'existe pas d'évaluation disponible pour les cas de non-contact, le traitement de ce groupe n'est pas modifié.

Nous utilisons l'indicateur  $R$  donné par (2.7) comme fonction d'objectif de qualité. Nous répartissons l'échantillon en utilisant  $X = (\text{âge})$  en deux groupes, nommés *personnes jeunes* et *personnes âgées*. L'objectif à la deuxième phase est d'affecter les cas de refus aux deux groupes d'intervieweurs de manière que l'indicateur  $R$  en ce qui concerne l'âge soit maximisé.

Soit  $n$  la taille de l'échantillon de l'enquête. Les proportions de population des deux sous-populations, *personnes jeunes* et *personnes âgées*, sont désignées par  $q(1)$  et  $q(2)$ . Soit  $q(\tilde{x} | x)$  la probabilité conditionnelle qu'une unité échantillonnée provenant de la sous-population d'âge  $x$  soit de type  $\tilde{x}$ , où  $\tilde{x} \in \{\text{forte, moyenne, faible}\}$ . En outre, soit  $\lambda(x, \tilde{x})$  la probabilité qu'une unité échantillonnée de type  $\tilde{x}$  provenant de la sous-population d'âge  $x$  soit un cas de refus. Si une personne n'est pas un cas de refus, alors  $\mu(x, \tilde{x})$  est la probabilité que cette personne était un répondant après la première phase ou qu'elle devienne un répondant si elle était un cas de non-contact après la première phase.

Le nombre total d'intervieweurs est  $M$  et  $p_s M$  représente le nombre d'intervieweurs ayant la compétence  $s \in S = \{\text{bon, moins bon}\}$ ,  $0 \leq p_s \leq 1$  et  $p_{\text{bon}} + p_{\text{moins bon}} = 1$ . L'ensemble  $S$  forme l'ensemble de stratégies, c'est-à-dire que nous voulons attribuer chaque cas de refus à un bon intervieweur ou à un moins bon intervieweur. Nous supposons que chaque intervieweur peut traiter au plus  $c$  cas de refus à la deuxième phase de l'enquête. La probabilité qu'un cas de refus de type  $\tilde{x}$  provenant de la sous-population  $x$  réponde s'il est contacté par un intervieweur ayant

la compétence  $s$  est désignée par  $\rho(s, x, \tilde{x})$  et nous supposons de nouveau qu'elle est connue d'après les enquêtes antérieures.

Soit  $\{p(s | x, \tilde{x})\}_{x, \tilde{x}}$  l'ensemble de variables décisionnelles, où  $p(s | x, \tilde{x})$  représente la probabilité qu'une unité échantillonnée de type  $\tilde{x}$  soit attribuée à un intervieweur de compétence  $s$  sachant qu'elle appartient à la sous-population  $x$ . Autrement dit, nous permettons l'affectation aléatoire des unités échantillonnées aux deux groupes d'intervieweurs.

Dans cet exemple, nous exprimons les coûts en fonction des taux globaux d'activité des intervieweurs. Puisque les intervieweurs peuvent traiter au plus  $c$  cas, il existe deux contraintes,

$$n \sum_{x, \tilde{x}} q(x)q(\tilde{x} | x)p(s | x, \tilde{x})\lambda(x, \tilde{x}) \leq Mp_s c, \quad \forall s \in S.$$

Autrement dit, le nombre total de cas de refus qui peuvent être attribués aux intervieweurs de compétence  $s$  est limité à la charge de travail maximale possible pour le groupe ayant cette compétence.

La propension à répondre d'une unité provenant de la sous-population  $x$  peut maintenant être dérivée sous la forme

$$\sum_{\tilde{x}} q(\tilde{x} | x) \left[ (1 - \lambda(x, \tilde{x}))\mu(x, \tilde{x}) + \lambda(x, \tilde{x}) \sum_s p(s | x, \tilde{x})\rho(s, x, \tilde{x}) \right],$$

et représente la valeur entrée dans l'indicateur R.

Considérons maintenant les données d'entrée suivantes pour l'exemple : une taille d'échantillon de  $n = 2\,000$ , un total de 80 intervieweurs,  $M = 80$ , une charge de travail maximale de 30 cas par intervieweur,  $c = 30$ , une distribution selon l'âge égale à  $q(1) = q(2) = 0,5$ , les distributions conditionnelles de type de refus  $q(\tilde{x} | 1) = (0,2; 0,3; 0,5)'$  et  $q(\tilde{x} | 2) = (1/3; 1/3; 1/3)'$  et la classification de 25 % des intervieweurs comme étant bons,  $p_1 = 0,25 = 1 - p_2$ .

Les tableaux 3.1 et 3.2 donnent les probabilités de réponse hypothétique  $\rho(s, x, \tilde{x})$  pour les deux sous-groupes quand on applique la conversion des cas de refus, de même que les probabilités de coopération  $\mu(x, \tilde{x})$  et les probabilités de refus  $\lambda(x, \tilde{x})$ .

Nous optimisons l'indicateur R par rapport aux deux groupes d'âge. Dans le cas de deux strates, on peut montrer que l'indicateur R est maximal quand la distance absolue entre les propensions à répondre dans les deux strates est minimale. La valeur optimale de l'indicateur R s'avère être de 0,827. Le tableau 3.3 donne les valeurs optimales des variables décisionnelles; toutes ces variables sauf une,  $p(s|x, \tilde{x})$ , ont une valeur de 0 ou de 1, c'est-à-dire que les réaffectations sont principalement non probabilistes. Fait exception la sous-population des personnes jeunes pour lesquelles la propension à répondre est évaluée comme étant moyenne.

**Tableau 3.1**

**Probabilités de réponse quand la conversion des refus est appliquée aux personnes jeunes et âgées qui refusent de répondre sachant l'évaluation de la propension à répondre**

|                         | Refus d'une personne jeune |         |        |                        |         |        |
|-------------------------|----------------------------|---------|--------|------------------------|---------|--------|
|                         | Bon intervieweur           |         |        | Moins bon intervieweur |         |        |
|                         | Forte                      | Moyenne | Faible | Forte                  | Moyenne | Faible |
| $\rho(s, 1, \tilde{x})$ | 0,8                        | 0,6     | 0,4    | 0,7                    | 0,5     | 0,3    |
|                         | Refus d'une personne âgée  |         |        |                        |         |        |
|                         | Bon intervieweur           |         |        | Moins bon intervieweur |         |        |
|                         | Forte                      | Moyenne | Faible | Forte                  | Moyenne | Faible |
| $\rho(s, 2, \tilde{x})$ | 0,9                        | 0,7     | 0,5    | 0,8                    | 0,6     | 0,4    |

**Tableau 3.2**

**Probabilités de refus et de coopération à la première phase de collecte des données**

|                         | Personne jeune |         |        | Personne âgée |         |        |
|-------------------------|----------------|---------|--------|---------------|---------|--------|
|                         | Forte          | Moyenne | Faible | Forte         | Moyenne | Faible |
| $\lambda(x, \tilde{x})$ | 0,5            | 0,6     | 0,7    | 0,2           | 0,3     | 0,4    |
| $\mu(x, \tilde{x})$     | 0,85           | 0,8     | 0,76   | 0,95          | 0,93    | 0,91   |

**Tableau 3.3**

**Affectation optimale des cas aux intervieweurs**

|                        | Personne jeune |         |        | Personne âgée |         |        |
|------------------------|----------------|---------|--------|---------------|---------|--------|
|                        | Forte          | Moyenne | Faible | Forte         | Moyenne | Faible |
| Bon intervieweur       | 1              | 0,83    | 1      | 0             | 0       | 0      |
| Moins bon intervieweur | 0              | 0,17    | 0      | 1             | 1       | 1      |

Il est utile de comparer l'affectation optimale à une affectation aléatoire des intervieweurs afin d'évaluer le gain d'efficacité. Si nous affectons aléatoirement les cas de refus aux intervieweurs, la valeur de l'indicateur R est égale à 0,749. Donc, l'affectation optimale donne lieu à une augmentation considérable de la valeur de l'indicateur R. Les taux de réponse sont, respectivement, de 72,0 % et de 70,1 % pour les affectations optimale et aléatoire.

Si nous augmentons le nombre d'intervieweurs tout en fixant le nombre maximal de cas par intervieweur ainsi que les autres paramètres, pour tout nombre d'intervieweurs supérieur à  $M = 84$ , l'indicateur R ne s'améliore pas. Les deux groupes d'intervieweurs sont suffisamment grands pour traiter tous l'échantillon et la contrainte de coût n'est plus vraiment une contrainte. L'indicateur R pour  $M = 84$  est égal à 0,830 et le taux de réponse est de 72,1 %. Si nous maximisons le taux de réponse plutôt que l'indicateur R, l'affectation des intervieweurs convergera vers l'affectation de tous les cas aux *bons* intervieweurs seulement.

#### **4 Un plan de collecte adaptatif statique : affectation des intervieweurs travaillant par téléphone**

À la présente section, nous présentons une étude en simulation où la caractéristique du plan de collecte d'intérêt est l'affectation des intervieweurs travaillant par téléphone. Les probabilités de réponse utilisées dans l'exemple sont estimées après les données d'une enquête téléphonique réelle.

L'Enquête sur la satisfaction des consommateurs (ESC) des Pays-Bas est une enquête téléphonique mensuelle sur les sentiments des ménages au sujet de leur situation économique et de leurs dépenses. L'enquête donne un aperçu du développement économique à court terme et des indicateurs précoces des différences dans les tendances de consommation. Chaque mois, 1 500 ménages sont échantillonnés. Les deux causes les plus influentes de non-réponse dans l'enquête sont la non-prise de contact et le refus de participer. De l'échantillon sélectionné, 95 % sont contactés et des ménages contactés, 71 % participent. Le taux de réponse est de 67 %.

L'intervieweur est l'un des facteurs qui influent le plus sur la participation. Le rendement des intervieweurs peut varier considérablement quand il s'agit d'obtenir une réponse. En tout, 60 intervieweurs ont travaillé à l'ESC en 2005. Cela signifie qu'un intervieweur a pris contact avec 280 ménages en moyenne. Les taux de participation par intervieweur variaient de 50 % à 79 %. Le taux le plus faible, soit 50 %, était toutefois exceptionnel, car l'avant-dernier taux de participation était de 61 %. Le taux de participation moyen par intervieweur était de 67 %. Les ménages ont été affectés aléatoirement aux intervieweurs dans le système de gestion de l'ITAO. Donc, en ce qui concerne l'intervieweur, les données ont été randomisées (ou entremêlées). Dans la suite de l'exposé, l'intervieweur sera la caractéristique d'intérêt du plan de collecte. L'ensemble de stratégies de collecte  $S$  comprend soixante stratégies,  $S = \{s_1, s_2, \dots, s_{60}\}$ .

Partant des variables auxiliaires disponibles, nous avons sélectionné un vecteur  $X$  contenant l'ethnicité, la composition du ménage en ce qui concerne le sexe (masculin, féminin ou mélange), l'âge moyen du noyau du ménage par tranche de cinq ans, le type de ménage, le degré d'urbanisation du quartier de résidence et la valeur moyenne des logements dans le quartier. L'âge, la valeur moyenne des logements et le type de ménage sont surtout reliés aux statistiques clés déduites de l'ESC. Aucune parodonnée n'était disponible dans la présente étude. Par conséquent, le plan de collecte adaptatif est statique. Dans le problème d'optimisation, les probabilités d'affectation  $p(s_k | x)$  doivent être choisies, c'est-à-dire qu'il faut décider à quel intervieweur sont affectées les sous-populations en se basant sur  $X$  (de sorte que  $\sum_k p(s_k | x) = 1$ ).

Le coefficient de variation des propensions à répondre  $\rho_x$  défini par (2.8) est sélectionné comme étant la fonction de qualité cible. Pour estimer les propensions à répondre  $\rho(s_k, x)$  pour les intervieweurs, nous utilisons un modèle multiniveaux comportant une fonction de lien identité, c'est-à-dire une régression linéaire comprenant deux niveaux. Les intervieweurs forment le premier niveau du modèle et les ménages, le second. Le modèle multiniveaux est utilisé pour séparer les propensions à répondre individuelles et les propensions à répondre selon

l'intervieweur. La logique est qu'en faisant cette distinction, l'effet d'intervieweur peut être isolé et l'affectation des intervieweurs peut être optimisée. Nous choisissons un modèle linéaire car il est facile à optimiser. Puisque les propensions ne sont jamais proches de 0 ni de 1, le modèle linéaire produit presque les mêmes estimations qu'un modèle logit ou probit.

En ce qui concerne l'effet d'intervieweur, nous avons d'abord cherché à déterminer s'il était suffisant d'utiliser un modèle multiniveaux à pente fixe, c'est-à-dire un modèle dans lequel l'intervieweur est ajouté comme un effet principal seulement et dans lequel il n'y a aucune interaction avec les variables auxiliaires. Toutes les covariables présélectionnées ont contribué de manière significative au modèle multiniveaux, mais aucune des interactions avec l'intervieweur n'était significative au seuil de signification de 5 %. Par conséquent, nous nous sommes limités au modèle d'effet principal suivant

$$\rho(s_k, x_i) = \beta_0 + \beta_x x_i + \beta_k \quad (4.1)$$

où  $x_i$  est le vecteur de covariables du ménage  $i$ ,  $1 \leq i \leq n$ ,  $n$  est la taille d'échantillon,  $\beta_k$  est l'effet (fixe) d'intervieweur pour l'intervieweur  $k$ ,  $\beta_0$  est le terme constant ou ordonnée à l'origine, et  $\beta_x$  est le paramètre de pente. Soit  $\rho(x_i) = \sum_k p(s_k | x_i) \rho(s_k, x_i)$  la propension à répondre de l'unité échantillonnée  $i$ .

Le modèle (4.1) a été ajusté au jeu de données de l'ESC. Puis, nous nous sommes servis de l'effet d'intervieweur estimé  $\beta_k$  pour optimiser le coefficient de variation, sous deux contraintes de coût : la durée totale de l'interview ainsi que le nombre individuel d'appels faits par chaque intervieweur devaient être les mêmes que dans le plan original. Puisque le système de gestion des appels téléphoniques traite les appels, la durée de l'interview est la composante dominante du coût. Si nous fixons la durée totale de l'interview, nous contraignons le coût d'être le même que pour l'ESC ordinaire. Puisque les intervieweurs ne peuvent traiter qu'un certain nombre d'appels, nous devons aussi fixer le nombre d'appels qui leur sont affectés. La première contrainte implique que nous fixons le taux de réponse, car la durée totale d'interview est égale au produit

de la durée individuelle moyenne d'interview et du nombre de répondants. Le questionnaire de l'ESC est simple et ne contient aucun groupe d'items emboîtés. Par conséquent, la durée individuelle d'interview ne varie pour ainsi dire pas sur les sous-groupes de population. La deuxième contrainte est égale à

$$\sum_i p(s_k | x_i) = n_k, \quad (4.2)$$

où  $n_k$  est le nombre spécifié préalablement d'appels pour l'intervieweur  $k$  et  $\sum_k n_k = n$ .

Nous optimisons le coefficient de variation en répartissant les  $\beta_k$  entre les ménages. Étant donné la nature additive du modèle, il est facile de montrer que toute permutation des intervieweurs attribués aux cas mène à la même propension moyenne à répondre et, donc, aux mêmes durée et coût d'interview. La propension moyenne à répondre est donnée par

$$\bar{p} = \frac{1}{n} \sum_i p(x_i) = \frac{1}{n} \sum_{i,k} p(s_k | x_i) (\beta_0 + \beta_x x_i + \beta_k) = \beta_0 + \frac{1}{n} \sum_i \beta_x x_i + \frac{1}{n} \sum_k n_k \beta_k,$$

qui ne dépend pas de l'ensemble de probabilités d'affectation  $p(s_k | x)$ . Par conséquent, optimiser le coefficient de variation revient à optimiser la variance des propensions à répondre  $S^2(\rho_x)$ .

Si nous nous limitons à des variables décisionnelles de type 0-1, c'est-à-dire  $p(s_k | x) \in \{0, 1\}$ ,  $\forall x, k$ , alors il est relativement facile de montrer que l'affectation optimale consiste à unir les meilleurs intervieweurs aux unités échantillonnées les plus difficiles et inversement. Autrement dit, les unités échantillonnées sont triées en classant les propensions à répondre individuelles sans l'effet d'intervieweur,  $\beta_0 + \beta_x x_i$ , par ordre croissant, et les intervieweurs sont triés par ordre décroissant en se basant sur l'effet d'intervieweur,  $\beta_k$ . Si deux unités échantillonnées  $i$  et  $j$  sont affectées à deux intervieweurs différents, disons  $k$  et  $l$ , et que  $\beta_x x_i < \beta_x x_j, \beta_k < \beta_l$  et  $p(s_k | x_i) = p(s_l | x_j) = 1$ , alors il est optimal de permuter les deux intervieweurs, c'est-à-dire  $p(s_l | x_i) = p(s_k | x_j) = 1$ . Nous pouvons le montrer comme il suit. La différence de variance  $S^2(\rho_x)$  est proportionnelle à

$$\begin{aligned}
\Delta S^2(\rho_X) &= (\beta_0 + \beta_X x_i + \beta_k - \bar{\rho})^2 \\
&\quad + (\beta_0 + \beta_X x_j + \beta_l - \bar{\rho})^2 - (\beta_0 + \beta_X x_i + \beta_l - \bar{\rho})^2 - (\beta_0 + \beta_X x_j + \beta_k - \bar{\rho})^2 \quad (4.3) \\
&= 2(\beta_l - \beta_k)(\beta_0 + \beta_X x_j - \bar{\rho}) - 2(\beta_l - \beta_k)(\beta_0 + \beta_X x_i - \bar{\rho}) \\
&= 2(\beta_l - \beta_k)(\beta_X x_j - \beta_X x_i) > 0.
\end{aligned}$$

Partant de (4.3), nous pouvons conclure à une diminution de la variance et, donc, du coefficient de variation si nous permutons les deux intervieweurs pour les cas  $i$  et  $j$ . De cet argument, il découle facilement que la solution optimale est celle proposée. D'une manière similaire, mais qui nécessite plus d'opérations algébriques, on peut montrer que la solution optimale pour les affectations probabilistes,  $p(s_k | x) \in [0, 1]$ , est la même.

Les deux premières lignes du tableau 4.1 contiennent la propension moyenne à répondre et le coefficient de variation avant et après la réaffectation des intervieweurs. Le coefficient de variation est passé de 0,117 à 0,035. Afin d'avoir une idée du degré de signification de la variation de la fonction de qualité, nous avons calculé les erreurs-types par la méthode du bootstrap. Pour chaque bootstrap, nous avons effectué la réaffectation des intervieweurs. Les erreurs sont données au tableau 4.1.

**Tableau 4.1**

**Propension moyenne à répondre et coefficient de variation pour l'ESC ordinaire, l'ESC après réaffectation des intervieweurs avec et sans correction pour la durée de l'interview (les erreurs-types bootstrap sont entre parenthèses)**

| ESC           | Correction pour la durée de l'interview ? | $\bar{\rho}$ | $Q(p)$        |
|---------------|---|--------------|---------------|
| Ordinaire     | -   | 70,8 %       | 0,117 (0,005) |
| Réaffectation | Non                                       | 70,8 %       | 0,035 (0,003) |
| Réaffectation | Oui                                       | 70,8 %       | 0,034 (0,003) |

Le lecteur aura peut-être remarqué que des nombres fixes de cas par intervieweur n'impliquent pas des nombres fixes d'interview par intervieweur. En fait, en réarrangeant les intervieweurs, ceux qui sont bons feront moins d'interviews car ils obtiennent les cas difficiles, alors que ceux qui sont moins bons feront plus d'interviews. Par conséquent, les bons

intervieweurs travailleront un plus petit nombre d'heures qu'ils ne le feraient dans le cadre de l'ESC ordinaire et les moins bons intervieweurs travailleront plus. Cela serait un effet secondaire indésirable qui peut, toutefois, être corrigé assez facilement. En partant de la solution optimale et en triant de nouveau les unités échantillonnées en fonction de leur propension à répondre individuelle sans tenir compte de l'effet d'intervieweur, nous pouvons déplacer les cas avoisinants des intervieweurs moins bons aux bons intervieweurs. Nous effectuons ce déplacement de telle façon que la durée totale d'interview par intervieweur ne dépasse pas celle de l'ESC ordinaire. On peut de nouveau prouver que cette procédure donne lieu à une nouvelle solution optimale où la contrainte du nombre fixe de cas présente dans (4.2) est remplacée par la contrainte du nombre fixe d'interviews

$$\sum_i p(s_k | x_i) \rho(s_k, x_i) = r_k, \quad (4.4)$$

où  $r_k$  est le nombre d'interviews spécifié préalablement. Le tableau 4.1 donne le coefficient de variation pour la solution optimale étant donné (4.4). Le taux de réponse demeure fixe et le coefficient de variation est légèrement plus faible.

En 2009, l'ESC a été utilisée comme instrument pour tester un plan de collecte adaptatif statique. Nous renvoyons le lecteur à Luiten et Wetzels (2009) et à Luiten et Schouten (2013) pour des renseignements détaillés. L'affectation des intervieweurs était l'une des principales caractéristiques du plan de collecte qui a été adaptée. Les autres caractéristiques du plan étaient le mode de collecte des données et le protocole de prise de contact. Outre l'interview par téléphone, l'interview en ligne a également été choisie comme mode possible de collecte. Les unités échantillonnées dont la probabilité de contact estimée était faible ont été affectées à des protocoles de prise de contact plus intensifs et ont été classées par ordre de priorité. Les probabilités de prise de contact et les probabilités de réponse ont été estimées en se basant sur les données historiques de l'ESC. L'essai pilote a montré qu'il était possible d'améliorer significativement le coefficient de variation, tout en fixant le taux de réponse et le budget.

À la présente section, nous avons présenté une étude en simulation dans laquelle les bons intervieweurs par téléphone se voient attribuer les cas les plus difficiles. En pratique, cela pourrait irriter ces intervieweurs. Lorsqu'un plan de ce genre est mis en œuvre, il convient d'en informer les intervieweurs à l'avance. Durant l'essai pilote de l'ESC de 2009, cela n'a entraîné aucun commentaire négatif de la part des intervieweurs. Dans les enquêtes sur place, une réaffectation des intervieweurs ne peut pas avoir lieu aussi facilement, car les frais de déplacement peuvent varier considérablement. Néanmoins, dans les régions à forte densité de population, la réaffectation des intervieweurs peut être une option.

## 5 Discussion

Le présent article décrit les plans de collecte de données d'enquête dans lesquels différents traitements ou stratégies de collecte sont appliqués à différentes unités de la population. Les différences entre les unités de population sont reflétées par des covariables pour lesquelles on dispose de données de registre appariées ou de paradonnées. Les stratégies de collecte des données sont définies comme étant différentes spécifications des caractéristiques du plan de collecte. Les plans de ce genre sont appelés plans de collecte adaptatifs car ils adaptent (taille sur mesure) la collecte des données à la population d'intérêt. Les éléments fondamentaux des plans de collecte adaptatifs sont les stratégies de collecte, les covariables de population, les propensions à répondre, les fonctions de coût et de qualité, et les probabilités d'affectation des stratégies. Les plans de collecte adaptatifs ont pour objectif d'optimiser la qualité des réponses en affectant différentes stratégies à différentes unités de population. Les probabilités d'affectation des stratégies représentent les variables décisionnelles dans l'optimisation.

Selon nous, le présent article contribue à la littérature de trois façons : il présente un cadre général, il donne explicitement l'option de choisir parmi un ensemble de stratégies en se basant sur un compromis qualité-coût, et il se concentre sur des indicateurs de l'erreur de non-réponse. Les deux dernières composantes sont déjà décrites dans la littérature sur les enquêtes; c'est la

généralisation à plusieurs caractéristiques du plan de collecte et à l'erreur de non-réponse qui est nouvelle. Sous leur forme la plus modeste, les plans de collecte adaptatifs représentent une affectation des stratégies d'enquête stratifiée sur différents sous-groupes de population. Dans leur forme la plus ambitieuse, les plans sont des extensions des plans d'échantillonnage à des stratégies multiples en mettant l'accent sur l'erreur de non-réponse. Cependant, même sous leur forme la plus modeste, les plans de collecte adaptatifs peuvent englober les modes de collecte des données, les primes d'incitation, les rappels, la durée du travail sur le terrain dans les enquêtes avec interview sur place, les tâches des intervieweurs et le type de déclaration.

Les plans de collecte adaptatifs se prêtent le mieux aux circonstances dans lesquelles les enquêtes sont exécutées de manière répétée pendant une longue période. C'est dans ces conditions que les données historiques sont les plus puissantes. Les plans adaptatifs conviennent aussi pour les organismes qui réalisent de nombreuses enquêtes dont les thèmes et le budget sont relativement similaires. Les nouvelles enquêtes et les enquêtes ponctuelles demandent de la modestie et de la prudence. Cependant, cela est vrai aussi pour les plans à stratégie unique. Les plans de collecte adaptatifs peuvent tenir compte du manque de données historiques puissantes en permettant une incertitude dans les propensions à répondre et d'autres paramètres, et en introduisant une période d'apprentissage ou une phase de conception initiale.

À notre avis, la concentration sur l'erreur de non-réponse est un élément important du cadre. Dans le présent article, notre objectif est la représentativité de la réponse. Cet objectif vient de notre conviction que la non-réponse est toujours de type ne manquant pas au hasard (*not-missing-at-random*). Nous observons les écarts plus importants par rapport au mécanisme de formation de données manquantes entièrement au hasard (*missing-completely-at-random*) pour les variables auxiliaires pertinentes comme des indices d'une non-réponse produite au hasard plus forte sur les variables étudiées étant donné ces variables auxiliaires. Théoriquement, il ne doit pas nécessairement en être ainsi. Considérons une simple question à réponse binaire oui-non et un taux de non-réponse de 50 %. Ces cas extrêmes se produiraient si les non-répondants disaient

tous soit oui soit non. Ils peuvent le faire quel que soit le choix des variables auxiliaires et par conséquent le biais de non-réponse maximal sur cette question est le même indépendamment du choix des variables auxiliaires. Donc, la recherche fournira un soutien empirique pour la concentration sur les mesures indirectes de l'erreur de non-réponse.

Il est également nécessaire de poursuivre les travaux de recherche portant sur les plans de collecte adaptatifs et « réactifs » dans le contexte d'autres questions. La recherche devrait viser à étendre les plans à plusieurs erreurs d'enquête et à étudier la robustesse des plans à l'erreur de spécification des modèles de propension à répondre. Jusqu'à présent, les plans de collecte adaptatifs et « réactifs » ont été axés sur l'erreur de non-réponse, en ignorant l'erreur de réponse ou de mesure. Or, il est bien connu que certaines caractéristiques du plan de collecte des données, par exemple le mode de collecte des données ou les intervieweurs, peuvent avoir un effet important sur l'erreur de réponse et, par conséquent, sur l'erreur totale d'enquête. Les plans de collecte adaptatifs doivent, par conséquent, tenir compte de l'erreur de mesure également, quand on peut s'attendre à ce que les caractéristiques du plan auront un effet différentiel important sur l'erreur de réponse. L'optimisation tenant compte d'erreurs multiples représente un important domaine de recherche future.

Les plans de collecte des données adaptatifs devraient dans tous les cas être modestes en ce qui concerne le nombre de stratégies employées, afin d'éviter un processus de collecte exagérément complexe et l'optimisation sur des propensions et des fonctions de coût qui sont sujettes à incertitude. Néanmoins, il convient toujours de privilégier une méthode structurée d'examen; les plans de collecte adaptatifs fournissent ce genre de cadre et permettent une recherche structurée de plans de collecte des données améliorés.

## **Remerciements**

Les auteurs remercient James Wagner, Mick Couper, Fannie Cobben et Mariëtte Vosmer de leurs commentaires utiles au sujet de la version provisoire du présent article. Les auteurs

remercient aussi le rédacteur adjoint et les examinateurs de leurs remarques qui leur ont permis d'améliorer considérablement l'article.

## Bibliographie

- Barón, J.D., Breunig, R.V., Cobb-Clark, D., Gørgens, T. et Sartbayeva, A. (2009). Does the effect of incentive payments on survey response rates differ by income support history. *Journal of Official Statistics*, 25, 483-507.
- Beaumont, J.-F., et Haziza, D. (2011). A theoretical framework for adaptive collection designs. Document présenté au 5<sup>e</sup> atelier de l'International Total Survey Error, 21 au 23 juin, Québec, Canada.
- Calinescu, M., Schouten, B. et Bhulai, S. (2012). Adaptive survey designs that minimize nonresponse and measurement risk, document de travail, Statistics Netherlands. Disponible au [www.cbs.nl](http://www.cbs.nl).
- Cobben, F. (2009). Nonresponse in sample surveys. Methods for analysis and adjustment, Thèse de doctorat, University of Amsterdam.
- De Leeuw, E.D. (2008). Choosing the Method of Data Collection. Dans *International Handbook of Survey Methodology*, (Éds., E.D. De Leeuw, J.J. Hox et D.A. Dillman), Lawrence Erlbaum Associates, New York, États-Unis, 113-135.
- De Leeuw, E., Callegaro, M., Hox, J., Korendijk, E. et Lensvelt-Mulders, G. (2007). The influence of advance letters on response in telephone surveys. *Public Opinion Quarterly*, 71, 413-443.
- Dillman, D.A. (2007). Mail and internet surveys: The tailored design method, Wiley, Hoboken, New Jersey, États-Unis.
- Greenberg, B.S., et Stokes, S.L. (1990). Developing an optimal call scheduling strategy for a telephone survey. *Journal of Official Statistics*, 6, 421-435.
- Groves, R.M. (1989). Survey errors and survey costs, Wiley, New York, États-Unis.
- Groves, R.M., et Couper, M.P. (1998). Nonresponse in Household Interview Surveys, Wiley series in probability and statistics, Survey methodology section.
- Groves, R.M., Dillman, D., Eltinge, J. et Little, R. (2002). Survey Nonresponse, New York : Wiley Series in Probability and Statistics.
- Groves, R.M., et Heeringa, S.G. (2006). Responsive design for household surveys: Tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society, Série A*, 169, 439-457.
- Hartley, H.O., et Monroe, H. (1979). Interviewer assignments which minimize the effects of nonsampling errors. *Proceedings of the Survey Research Methods Section*, American Statistical Association, 216-220.

- Heyd, J.M., et Carlin, B.P. (1999). Adaptive design improvements in the continual reassessment method for Phase I studies. *Statistics in Medicine*, 18, 1307-1321.
- Kalsbeek, W.D., Botman, S.L., Massey, J.T. et Liu, P.W. (1994). Cost-efficiency and the number of allowable call attempts in the National Health Interview Survey. *Journal of Official Statistics*, 10, 133-152.
- Kersten, H.M.P., et Bethlehem, J.G. (1984). Exploring and reducing the non-response bias by asking the basic question, BPA 627-84-M1, CBS, Voorburg.
- Kulka, R.A., et Weeks, M.F. (1988). Towards the development of optimal calling protocols for telephone surveys: A conditional probabilities approach. *Journal of Official Statistics*, 4, 319-322.
- Laflamme, F., et Karaganis, M. (2010). Implementation of responsive collection design for CATI surveys at Statistics Canada, document présenté à la Q2010, 3 au 6 mai, Helsinki, Finland.
- Luiten, A., et Schouten, B. (2013). Adaptive fieldwork design to increase representative household survey response. *Journal of the Royal Statistical Society, Séries A*, 176, 1, 169-190.
- Luiten, A., et Wetzels, W. (2009). Indicators and data collection control. Work plan and preliminary findings pilot Statistics Netherlands, RISQ deliverable 8.2, [www.risq-project.eu](http://www.risq-project.eu).
- Lyberg, L.E., Biemer, P., Collins, M., de Leeuw, E.D., Dippo, C., Schwarz, N. et Trewin, D. (1997). Survey measurement and process quality, Wiley Series in Probability and Statistics, Wiley, Hoboken, New Jersey, États-Unis.
- Lynn, P. (2003). PEDAksi: Methodology for Collecting Data about Survey Non-Respondents. *Quality and Quantity*, Vol. 37, 3, 239-261.
- Mohl, C., et Laflamme, F. (2007). Research and responsive design options for survey data collection at Statistics Canada. Proceedings of ASA Joint Statistical Meeting, Section 293, 29 juillet au 2 août, Salt Lake City, États-Unis.
- Moore, J. (1988). Self/Proxy response status and survey response quality. A review of literature. *Journal of Official Statistics*, 4, 2, 155-172.
- Murphy, S.A. (2003). Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society, Séries B*, 65, 331-355.
- Peytchev, A., Riley, S., Rosen, J., Murphy, J. et Lindblad, M. (2010). Reduction of Nonresponse Bias in Surveys through Case Prioritization, *Survey Research Methods*, 4, 1, 21-29.
- Phillips, O., et Tabuchi, T. (2009). Un plan de sondage réactif pour l'Enquête sur la dynamique du travail et du revenu. *Recueil : Symposium 2009, Les enquêtes longitudinales : de la conception à l'analyse*, Statistique Canada.
- Särndal, C.-E. (2011a). Three factors to signal non-response bias with applications to categorical auxiliary variables. *Revue Internationale de Statistique*, 79, 2, 233-254.
- Särndal, C.-E. (2011b). The 2010 Morris Hansen Lecture: Dealing with survey nonresponse in data collection, in estimation. *Journal of Official Statistics*, 27, 1, 1-21.

- Särndal, C.-E., et Lundström, S. (2010). Plan d'estimation : détermination de vecteurs auxiliaires en vue de réduire le biais de non-réponse. *Techniques d'enquête*, 36, 2, 141-156.
- Schouten, B., Cobben, F. et Bethlehem, J. (2009). Indicateurs de la représentativité de la réponse aux enquêtes. *Techniques d'enquête*, 35, 1, 107-121.
- Schouten, B., Luiten, A., Loosveldt, G., Beullens, K. et Kleven, Ø. (2010). Monitoring and changing data collection through R-indicators and partial R-indicators, document de travail du RISQ. Disponible à l'adresse [www.risq-project.eu](http://www.risq-project.eu).
- Schouten, J.G., Shlomo, N. et Skinner, C. (2011). Indicators for monitoring and improving representativeness of response. *Journal of Official Statistics*, 27, 2, 231-253.
- Wagner, J. (2008). Adaptive survey design to reduce nonresponse bias, Thèse de doctorat, University of Michigan, États-Unis.
- Zajonc, T. (2012). Bayesian inference for dynamic treatment regimes: Mobility, equity and efficiency in student tracking. *Journal of the American Statistical Association*, 107, 80-92.