

Techniques d'enquête

Plans de collecte de données adaptatifs visant à minimiser les effets du mode d'enquête – étude du cas de l'Enquête sur la population active des Pays-Bas

par Melania Calinescu et Barry Schouten

Date de diffusion : le 17 décembre 2015



Statistique
Canada

Statistics
Canada

Canada

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.

Vous pouvez également communiquer avec nous par :

Courriel à STATCAN.infostats-infostats.STATCAN@canada.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

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 sous « Contactez-nous » > « Normes de service à la clientèle ».

Note de reconnaissance

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

Signes conventionnels dans les tableaux

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^s 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 ($p < 0,05$)

Publication autorisée par le ministre responsable de Statistique Canada

© Ministre de l'Industrie, 2015

Tous droits réservés. L'utilisation de la présente publication est assujettie aux modalités de l'[entente de licence ouverte](#) de Statistique Canada.

Une [version HTML](#) est aussi disponible.

This publication is also available in English.

Plans de collecte de données adaptatifs visant à minimiser les effets du mode d'enquête – étude du cas de l'Enquête sur la population active des Pays-Bas

Melania Calinescu et Barry Schouten¹

Résumé

L'évaluation de l'impact des effets de mode sur les estimations d'enquête est devenue un objectif de recherche crucial en raison de l'utilisation croissante des plans de sondage à modes mixtes. Malgré les avantages inhérents aux plans de ce genre, comme la réduction des coûts et l'augmentation de la couverture, il y a une évidence manifeste à l'effet que le mode d'enquête peut avoir des effets importants relativement à la précision d'une enquête. Ces effets risquent d'aboutir à des statistiques qui ne sont pas comparables dans le temps ou entre des sous-groupes de population et d'accroître le biais. Les plans de collecte adaptatifs offrent un cadre mathématique souple qui permet d'optimiser l'équilibre entre la qualité et les coûts de l'enquête. Dans cet article, nous utilisons des plans adaptatifs afin de minimiser les effets de mode. Nous illustrons notre modèle d'optimisation au moyen de l'étude du cas de l'Enquête sur la population active des Pays-Bas. Nous mettons l'accent sur les effets de mode dépendant des items et nous évaluons l'impact sur la qualité des enquêtes par comparaison avec une norme de référence.

Mots-clés : Biais de sélection propre à un mode; biais de mesure propre à un mode; coûts d'enquête; qualité de l'enquête.

1 Introduction

Dans cet article, nous proposons de démontrer comment les plans de collecte adaptatifs aident à minimiser les effets de mode lorsqu'une enquête a une seule statistique ou un seul indicateur. Nous démontrons cette méthode à l'aide de l'Enquête sur la population active des Pays-Bas (EPA), dont l'indicateur clé est le taux de chômage.

L'émergence du Web comme mode d'enquête a ravivé les discussions sur les enquêtes multimodales. Les sociétés d'étude de marché n'ont pas tardé à intégrer le Web à leurs plans de sondage. Les instituts statistiques officiels ont été plus lents à adopter cette approche, mais ils envisagent des plans de sondage à modes mixtes incluant le Web, en raison notamment des coûts accrus des enquêtes en personne, de la couverture réduite des enquêtes par téléphone et du faible taux de participation aux enquêtes Web (Fan et Yan 2010). En conséquence, les organismes d'enquête convertissent graduellement leurs plans de sondage à mode unique en plans de sondage à modes mixtes. En 2012, un projet de grande envergure appelé *Data Collection for the Social Surveys* (DCSS) a été lancé dans le système statistique de l'Union européenne afin d'examiner les plans de sondage à modes mixtes pour l'Enquête sur la population active (Blanke et Luiten 2012).

Il est bien connu que le mode d'enquête a un impact sur les erreurs non dues à l'observation (non-réponse partielle, non-réponse totale et sous-dénombrement) ainsi que sur les erreurs d'observation (erreurs de mesure et de traitement). La différence globale entre deux modes est généralement qualifiée d'« effet de mode ». La différence entre les erreurs de mesure de deux modes est appelée « effet de mode pur » ou « effet de mesure », tandis que la différence dans le sous-dénombrement et la non-réponse est appelée « effet de sélection ». La question est examinée en détail dans de Leeuw (2005), Dillman, Phelps,

1. Melania Calinescu, Département de mathématiques, VU University Amsterdam, De Boelelaan 1081, 1081HV Amsterdam, Netherlands. Courriel : mcalinescu@gmail.com; Barry Schouten, Statistics Netherlands, C.P. 24500, 2490HA Den Haag, Netherlands. Courriel : jg.schouten@cbs.nl.

Tortora, Swift, Kohrell, Berck et Messes (2009), Vannieuwenhuyze (2013) et Klausch, Hox et Schouten (2013b). Il existe des données probantes (Jäckle, Roberts et Lynn 2010; Schouten, van den Brakel, Buelens, van der Laan, Burger et Klausch 2013b; Dillman et coll. 2009) selon lesquelles les effets de mode peuvent être importants. Ces effets peuvent produire des statistiques qui ne sont pas comparables dans le temps ou entre des sous-groupes de population. L'évaluation, la minimisation et la stabilisation de l'impact des effets de mode sur les estimations d'enquête sont devenues des objectifs importants.

Il y a quatre façons possibles de réduire l'impact des effets de mode dans le plan de sondage et l'estimation par sondage. Une approche rigoureuse de la conception des questionnaires et des plans de collecte de données devrait prévenir ces effets, tandis que la procédure d'estimation et le calage aident à tenir compte des effets de mode avec la pondération. Une conception minutieuse des questionnaires permet de réduire les différences de mesure entre les modes. On peut réduire ces différences en utilisant un plan d'enquête à mode unifié pour les questionnaires (Dillman et coll. 2009) ou en créant un stimulus équivalent par mode, (de Leeuw 2005). Certains effets de mesure sont toutefois inhérents au processus d'administration du mode d'enquête. Par exemple, une présentation orale plutôt que visuelle ou le rythme de l'interview rendent difficile ou impossible la suppression complète de ces effets. De plus, la conception des questionnaires ne permet pas de supprimer les effets de sélection, mais leur longueur, leur disposition et leur contenu peuvent fréquemment produire des effets de mesure et de sélection. Par ailleurs, l'historique des questions peut empêcher une refonte complète du questionnaire par mode, lorsque les utilisateurs de l'enquête ou les intervenants ne veulent pas raccourcir le questionnaire ou modifier le libellé des questions. En résumé, il restera toujours des effets de mode, même après un remaniement en profondeur du questionnaire. Si des estimations des effets de mesure et des effets de sélection sont disponibles, elles peuvent être utilisées pour concevoir la stratégie de collecte de données d'une enquête de manière à éviter ces effets, ou pour concevoir la stratégie d'estimation afin d'en tenir compte dans les futures enquêtes.

L'option de plan suppose que certains modes ou séquences de modes ne sont pas appliqués, parce qu'ils sont censés produire d'importants effets de mode par rapport à un plan de collecte repère, c'est-à-dire un plan de sondage considéré comme exempt d'effets de mode. L'espérance relative aux effets de mode importants se fonde idéalement sur des études pilotes, mais elle peut aussi être le fruit de l'expérience. Lorsque le choix de mode(s) n'est pas uniforme pour tout l'échantillon et qu'il se fonde plutôt sur les caractéristiques des personnes ou des ménages, l'option du plan de sondage revient à un plan de sondage adaptatif (Wagner 2008; Schouten, Calinescu et Luiten 2013a). Ces caractéristiques peuvent être disponibles avant le début de la collecte de données ou le devenir durant la collecte de données sous forme de paradonnées (données sur le processus de collecte des données [Kreuter 2013]), ce qui mène à des plans de collecte adaptatifs respectivement statiques et dynamiques. Le présent article est axé sur la prévention des effets de mode au moyen de plans de collecte adaptatifs.

L'option d'ajustement est particulièrement intéressante lorsqu'il y a une bonne raison d'approximer les valeurs réelles d'une statistique, c'est-à-dire lorsque l'accent est mis non seulement sur la comparabilité mais aussi sur l'exactitude des statistiques. Un inconvénient de l'option d'ajustement est qu'elle est plus coûteuse que l'option de plan, puisque des estimations précises des effets de mode sont requises pour que l'exactitude des statistiques résultantes ne soit pas affectée. Un avantage de cette option est qu'elle est plus souple. En effet, elle permet d'apporter différents ajustements à différentes variables d'enquête, tandis que l'option du plan de sondage nécessite un choix global. Une discussion de l'ajustement durant l'estimation figure dans Vannieuwenhuyze (2013), Klausch, Hox et Schouten (2013a) et Suzer-Gurtekin (2013).

Une autre option consiste à stabiliser les effets de mode, ce qui constitue une option de dernier recours utile. Comme on suppose que les effets de mode sont présents après l'administration du questionnaire, la collecte de données et l'estimation fondée sur le plan, ils peuvent être stabilisés au fil du temps par calage de la distribution des modes dans la réponse à une certaine répartition fixe des modes. Si la proportion moyenne d'un mode de réponse diffère d'un mois à l'autre, on attribue un poids plus important aux répondants à ce mode et un poids plus faible aux répondants à d'autres modes. Pour une discussion de cette méthode, voir Buelens et van den Brakel (2014).

Dans cet article, nous minimisons l'effet de méthode ajusté à un plan de collecte à mode repère en stratifiant la population en sous-groupes pertinents et en répartissant les différents sous-groupes entre différents modes ou séquences de modes. L'effet de méthode ajusté d'un plan est la différence entre la moyenne ajustée de la non-réponse de ce plan et celle du plan de collecte repère. L'ajustement suit les procédures standard, c'est-à-dire le calage de la réponse à une répartition de la population. Par conséquent, l'effet de méthode ajusté est l'hybride de l'effet de mesure entre les deux plans d'échantillonnage et l'effet de sélection résiduel entre les deux plans qui n'est pas éliminé par l'ajustement pour la non-réponse.

Les plans de collecte adaptatifs et les plans de collecte réactifs qui leur ressemblent beaucoup (Heeringa et Groves 2006; Kreuter 2013) sont généralement appliqués afin de réduire l'erreur due à la non-réponse. À notre connaissance, jusqu'ici, seuls Calinescu et Schouten (2013a) ont tenté d'axer les plans de collecte adaptatifs sur les erreurs de mesure ou la combinaison d'erreurs dues à la non-réponse et d'erreurs de mesure. Les principales raisons de ce choix sont, premièrement, que les plans de collecte adaptatifs et réactifs en sont encore à leurs débuts et que leur application est limitée et, deuxièmement, que les erreurs de mesure et effets de mesure sont intrinsèquement difficiles à mesurer. De nombreuses applications des plans de collecte adaptatifs impliquent un seul mode d'enquête dans lequel il est plausible que l'erreur de mesure soit relativement stable pour différents choix de conception. Lorsque le mode d'enquête est une des caractéristiques du plan de sondage, cette hypothèse n'est plus plausible. Le mode d'enquête est toutefois la caractéristique la plus intéressante des plans de collecte adaptatifs en raison de son écart qualité-coûts important.

Une complication qui survient lors de l'inclusion de l'erreur de mesure dans les plans de collecte adaptatifs est que, contrairement à l'erreur due à la non-réponse, elle ne résulte pas d'une simple décision de type oui ou non. Une unité d'échantillonnage fournit une réponse ou une non-réponse, tandis qu'une erreur de mesure a aussi une ampleur. L'ampleur de l'erreur de mesure peut varier selon la question du questionnaire d'enquête. Cela suppose qu'en présence de multiples questions d'enquête ou variables, le choix de modes est une décision multidimensionnelle. Calinescu et Schouten (2013a) tentent de réduire cette multidimensionnalité en utilisant des styles de réponse (ou des tendances de réponse). Lorsqu'une enquête compte seulement une ou quelques variables clés, ce qui est le cas de l'EPA, cette complication n'existe pas et l'accent peut être mis directement sur les principales variables. C'est le cheminement que nous suivons dans cet article.

Dans cet article, nous adoptons donc deux approches novatrices : nous incluons les effets dus aux modes dans les plans de collecte adaptatifs et nous mettons l'accent sur une seule variable clé. Dans notre démonstration utilisant l'EPA des Pays-Bas, nous examinons trois modes d'enquête – en ligne, par téléphone et en personne – et diverses séquences de ces modes. Au cours des dernières années, la conception de l'EPA des Pays-Bas a subi une série de modifications dans le cadre de la transition d'une enquête en personne à une enquête à modes mixtes. Une vaste base de connaissances et des données

d'enquête historiques sur l'interaction entre les caractéristiques des plans de sondage, le mode d'enquête en particulier, et le processus de réponse sont disponibles. Nous utilisons ces données pour estimer les différents paramètres requis pour le modèle d'optimisation.

L'article se présente comme suit. Dans la section 2, nous formulons le problème d'optimisation multimodal. Dans la section 3, nous décrivons un algorithme pour l'optimisation du problème de l'effet de mode. Dans la section 4, nous présentons les résultats d'optimisation. Enfin, dans la section 5, nous examinons les résultats obtenus. Les annexes A et B présentent des extensions aux résultats numériques de la section 4.

2 Le problème d'optimisation multimodal

Dans cette section, nous construisons le problème d'optimisation multimodal qui explique les effets de mode sur une seule variable d'enquête clé. Outre le mode d'enquête, nous considérons aussi les limitations du nombre d'appels téléphoniques et d'interviews en personne comme des caractéristiques du plan de sondage dans l'optimisation. Dans le modèle d'optimisation, nous permettons l'attribution de différentes caractéristiques du plan à différentes sous-populations. L'optimisation peut donc mener à un plan de sondage adaptatif lorsque les probabilités de répartition optimale diffèrent entre les sous-populations. Dans notre cas, les sous-populations sont établies à partir de données administratives couplées. Elles peuvent aussi être fondées sur des parodonnées recueillies aux premières étapes de l'enquête. La dernière composante du problème d'optimisation est un ensemble de fonctions explicites de qualité et de coût. Dans notre cas, les fonctions de qualité sont tirées des différences de mode dans le biais de sélection et de mesure et des exigences relatives à la précision des statistiques. Pour la fonction de coût, nous utilisons le total des coûts variables du plan de sondage. Dans les paragraphes qui suivent, nous discutons des composantes du problème d'optimisation.

Nous commençons par les caractéristiques du plan de sondage contenues dans l'ensemble de stratégies de collecte \mathcal{S} . Nous considérons des stratégies à mode unique et à mode mixte séquentiel, c'est-à-dire des stratégies où un suivi des non-répondants à un mode d'enquête donné est assuré dans un autre mode. Un mode unique serait désigné par la lettre M et un mode mixte séquentiel serait étiqueté $M_1 \rightarrow M_2$. Nous considérons les enquêtes en ligne, par téléphone et en personne comme les modes d'intérêt et les désignons par les abréviations *Web*, *Tel* et *F2F* (pour « Face-to-Face » en anglais). Des exemples de mode unique et de mode mixte séquentiel sont *Tel* et *Web* \rightarrow *F2F*, respectivement. Pour les modes d'interview, nous considérons également une limite k du nombre d'appels, dénoté par Mk . Par exemple, *F2F3* dénote une stratégie de collecte à mode unique qui prévoit un maximum de trois visites pour l'interview en personne. Nous représentons la stratégie de contrepartie par $Mk +$ lorsqu'il n'y a pas de limite explicite. Dans cet article, nous n'examinons pas les stratégies à modes mixtes concurrents (deux modes ou plus sont offerts simultanément aux unités de l'échantillon). Cette restriction n'entraîne aucune perte de généralité. Il serait simple d'appliquer la méthodologie à n'importe quel ensemble de stratégies à modes mixtes, y compris les formes hybrides de stratégies à modes mixtes séquentiels et à modes mixtes concurrents. Cependant, un ensemble élargi ou diffus de stratégies s'accompagne d'un plus grand nombre de paramètres d'entrée à estimer. L'ensemble de stratégies de collecte \mathcal{S} inclut explicitement la stratégie vide, dénotée par Φ , qui représente le cas où une unité de population n'est pas échantillonnée, c'est-à-dire qu'aucune mesure n'est prise pour obtenir une réponse de l'unité. Nous utilisons $\mathcal{S}^R = \mathcal{S} \setminus \{\Phi\}$ pour désigner l'ensemble de stratégies non vides réelles.

Les unités de population sont réparties entre $\mathcal{G} = \{1, \dots, G\}$ groupes, étant donné un ensemble de caractéristiques X telles que l'âge et l'ethnicité, qui peuvent être extraites des sources de données externes ou des parodonnées. Soit $p(s, g)$ la probabilité de répartition de la stratégie s pour le groupe g , c'est-à-dire qu'une proportion $p(s, g)$ de la sous-population g est échantillonnée et contactée au moyen de la stratégie s . En général, on peut considérer que plusieurs stratégies ont des probabilités d'affectation non nulles, de sorte que la sous-population est répartie entre plusieurs stratégies. Définissons la probabilité d'affectation $p(\Phi, g)$ comme étant la probabilité qu'une unité de la sous-population g ne soit pas incluse dans l'échantillon. Le ratio $p(s, g)/(1 - p(\Phi, g))$ est la probabilité qu'une unité échantillonnée soit affectée à la stratégie s . Par exemple, si seules les probabilités d'affectation à la stratégie vide $p(\Phi, g)$ varient et que les probabilités d'affectation $p(s, g), \forall s \in \mathcal{S}^R$ sont égales à condition que l'unité soit échantillonnée, le plan de sondage est alors stratifié mais non adaptatif. Les probabilités doivent satisfaire la condition suivante :

$$\sum_{s \in \mathcal{S}^R} p(s, g) + p(\Phi, g) = 1, \quad \forall g \in \mathcal{G}, \quad (2.1)$$

$$0 \leq p(s, g) \leq 1, \quad \forall s \in \mathcal{S}, g \in \mathcal{G}.$$

Les probabilités que des stratégies d'enquête soient appliquées à des sous-populations $p(s, g)$ définissent les variables de décision dans le modèle d'optimisation. De façon plus générale et comme dans le cas des plans d'échantillonnage, nous pourrions tenir compte des dépendances entre les unités de population échantillonnées et/ou affectées à des stratégies non vides $s \in \mathcal{S}^R$. Pour ne pas compliquer l'étude de cas, nous supposons ici que les unités sont indépendantes.

Nous examinons maintenant les fonctions de qualité et de coût. Nous supposons que nous avons intérêt à estimer les moyennes de population d'une variable d'enquête y . Vu que nous considérons le mode d'enquête comme une des caractéristiques du plan de sondage, nous considérons le biais ajusté pour la non-réponse de y entre le plan proposé et un plan de collecte repère spécifié BM comme la principale fonction de qualité. Ce biais peut être considéré comme l'effet de méthode ajusté pour BM, et il s'agit d'une combinaison de biais de mesure propres à un mode et des biais persistants de non-réponse propres à un mode après ajustement. Si le plan proposé et le plan de collecte repère sont tous deux à mode unique, le biais est un effet de mode (ajusté) réel. Si seulement un des plans est à modes mixtes, le biais représente une combinaison complexe d'effets de mode (voir par exemple Klausch, Hox et Schouten 2014).

Soit N_g la taille du groupe g , $w_g = N_g/N$ la proportion du groupe g dans la population de taille N , et $\rho(s, g)$ la propension à répondre du groupe g si la stratégie s est retenue. Pour un groupe précis, nous définissons l'effet de méthode ajusté comme étant la différence ajustée pour la non-réponse entre l'estimation d'enquête $\bar{y}_{s,g}$ et une estimation repère \bar{y}_g^{BM} de la moyenne de la population \bar{Y} , où l'estimation d'enquête $\bar{y}_{s,g}$ est obtenue par application de la stratégie $s \in \mathcal{S}^R$ à la sous-population $g \in \mathcal{G}$. Nous représentons cette différence par $D(s, g)$. L'effet de méthode ajusté s'exprime comme suit :

$$D(s, g) = \bar{y}_{s,g} - \bar{y}_g^{\text{BM}}, \quad \forall s \in \mathcal{S}^R, g \in \mathcal{G}. \quad (2.2)$$

Par souci de commodité, nous omettons le participe passé « ajusté » dans le reste de l'article et parlons simplement d'*effet de méthode* pour désigner $D(s, g)$.

Dans cet article, nous cherchons à minimiser l'effet de méthode global absolu prévu pour un plan de collecte repère BM, donné, qui correspond à la moyenne pondérée des effets de méthode $D(s, g)$ par strate et par stratégie de BM. L'effet de méthode global absolu prévu pour BM est égal à

$$\bar{D}^{\text{BM}} = \left| \sum_{g \in \mathcal{G}} w_g \frac{\sum_{s \in \mathcal{S}^R} p(s, g) \rho(s, g) D(s, g)}{\sum_{s \in \mathcal{S}^R} p(s, g) \rho(s, g)} \right|. \quad (2.3)$$

Cette fonction d'objectif représente la variation prévue des séries temporelles de la statistique d'enquête clé lorsque le plan de collecte repère est converti en plan de collecte adaptatif à l'aide des probabilités d'affectation $p(s, g)$. Si une enquête est nouvelle ou si le plan de collecte repère n'a jamais été mis en application, la fonction d'objectif représente le biais du plan de collecte adaptatif pour le plan de collecte repère. Il s'agit donc d'une fonction d'objectif très utile. Il est à noter que $\bar{y}_{s,g}$ est une estimation ajustée pour la non-réponse de \bar{Y} , tandis que $\rho(s, g)$ est une estimation non pondérée de la probabilité de réponse du groupe g dans la stratégie s . Nous supposons implicitement que l'ajustement pour la non-réponse n'influence pas la contribution de chaque groupe et stratégie à la réponse globale. Cela nous permet d'exprimer la fonction d'objectif comme en (2.4), alors qu'un ajustement pour la non-réponse dans le cadre d'optimisation pourrait créer un problème très complexe, sinon impossible à résoudre. Nous minimisons l'effet de méthode global \bar{D}^{BM} en faisant une affectation optimale des stratégies $s \in \mathcal{S}^R$ aux groupes $g \in \mathcal{G}$, c'est-à-dire

$$\underset{p(s,g)}{\text{minimiser}} \bar{D}^{\text{BM}}. \quad (2.4)$$

Idéalement, $\bar{D}^{\text{BM}} = 0$. Cette situation pourrait toutefois causer de graves problèmes pratiques, notamment parce qu'elle nécessiterait des ressources illimitées. Notre modèle prévoit un certain nombre de contraintes afin de tenir compte de différents aspects pratiques, comme le manque de ressources. Un budget limité B est disponible pour configurer et exécuter l'enquête. Soit $c(s, g)$ le coût unitaire de l'application de la stratégie s à une unité du groupe g . La contrainte de coût est formulée comme suit :

$$\sum_{s,g} N_g p(s, g) c(s, g) \leq B. \quad (2.5)$$

Pour que l'estimation d'enquête de \bar{Y} , soit suffisamment précise, il faut un nombre minimal R_g de répondants par groupe, ce qui se traduit par la contrainte suivante :

$$\sum_{s \in \mathcal{S}^R} N_g p(s, g) \rho(s, g) \geq R_g, \quad \forall g \in \mathcal{G}. \quad (2.6)$$

En plus de la fonction d'objectif, l'effet de méthode entre le plan de collecte proposé et le plan de collecte repère fait partie d'une contrainte dans le problème d'optimisation, à savoir une contrainte sur la comparabilité des sous-groupes de population. L'effet de méthode global en tant que fonction d'objectif risque de mener à une solution déséquilibrée. Supposons par exemple qu'une stratégie s est appliquée à un groupe g de sorte que l'effet de méthode $D(s, g)$ correspondant est une valeur négative élevée, et que des stratégies produisant des valeurs $D(s, h)$ positives sont appliquées aux autres groupes $h \in \mathcal{G} \setminus \{g\}$. La valeur négative élevée $D(s, g)$ est annulée, mais le groupe g aura un comportement très différent des autres groupes, ce qui complique les comparaisons entre les groupes. Pour

prévenir ce genre de situation, nous limitons la différence absolue dans l'effet de méthode entre deux groupes au moyen de la contrainte suivante :

$$\max_{g, h \in \mathcal{G}} \left\{ \frac{\sum_{s \in \mathcal{S}^R} p(s, g) \rho(s, g) D(s, g)}{\sum_{s \in \mathcal{S}^R} p(s, g) \rho(s, g)} - \frac{\sum_{s \in \mathcal{S}^R} p(s, h) \rho(s, h) D(s, h)}{\sum_{s \in \mathcal{S}^R} p(s, h) \rho(s, h)} \right\} \leq M. \quad (2.7)$$

Cependant, quand

$$\frac{\sum_{s \in \mathcal{S}^R} p(s, g) \rho(s, g) D(s, g)}{\sum_{s \in \mathcal{S}^R} p(s, g) \rho(s, g)} - \frac{\sum_{s \in \mathcal{S}^R} p(s, h) \rho(s, h) D(s, h)}{\sum_{s \in \mathcal{S}^R} p(s, h) \rho(s, h)} \leq M \quad (2.8)$$

est inclus dans le problème d'optimisation pour chaque paire $(g, h) \in \mathcal{G}$, la contrainte (2.7) est automatiquement satisfaite. Pour des raisons pratiques, c'est-à-dire pour éviter l'épuisement du cadre d'échantillonnage, nous imposons également une contrainte sur la taille maximale de l'échantillon S_{\max} , c'est-à-dire

$$\sum_{s, g} N_g p(s, g) \leq S_{\max}. \quad (2.9)$$

De plus, nous exigeons qu'au moins une probabilité $p(s, g)$ soit strictement positive,

$$\sum_{s \in \mathcal{S}^R} p(s, g) > 0, \forall g \in \mathcal{G}, \quad (2.10)$$

afin d'éviter les erreurs de calcul telles qu'une division par zéro en (2.8).

La fonction d'objectif (2.4), conjuguée aux contraintes (2.1), (2.5)–(2.10), forme le problème d'optimisation multimodal afin de minimiser les effets de méthode par rapport à une valeur repère au moyen de plans de collecte adaptatifs. Il s'agit d'un problème non linéaire non convexe.

3 Un algorithme pour résoudre le problème d'optimisation multimodal

Dans la section qui précède, nous avons introduit les fonctions de qualité et de coût et construit un problème d'optimisation multimodal. La contrainte de la comparabilité des sous-populations, c'est-à-dire la limite supérieure de la différence absolue maximale entre les effets de méthode de groupe, rend le problème non convexe et difficile à résoudre. En conséquence, lorsqu'ils essaient de résoudre le problème d'optimisation multimodal, la plupart des solveurs de programmation non linéaire à des fins générales ne peuvent pas faire mieux qu'un optimum local. Le choix des points de départ dans les solveurs joue donc un rôle important. C'est pourquoi nous proposons une approche en deux étapes. Dans la première étape, nous résolvons un problème de programmation linéaire (PL) en agissant sur les contraintes linéaires (2.1), (2.5), (2.6) et (2.9)–(2.10). Dans la deuxième étape, nous utilisons la solution optimale obtenue dans la première étape comme point de départ à un algorithme de recherche local pour résoudre le problème non linéaire non convexe.

Nous reformulons le problème d'optimisation afin d'en faciliter le calcul. Étant donné que $|f(x)| = \max\{f(x), -f(x)\}$, nous pouvons reformuler la fonction d'objectif au moyen d'une variable additionnelle t et imposer les contraintes $f(x) \leq t$ et $-f(x) \leq t$. De toute évidence, t doit être non

négatif. Les contraintes mêmes ne changent pas; elles sont simplement remplacées. Le problème d'optimisation multimodal est donné en (3.2).

Nous pouvons dériver la PL en supprimant les contraintes non linéaires de la comparabilité des effets de méthodes entre les sous-populations et en remplaçant la fonction d'objectif non linéaire par une des contraintes linéaires. Nous choisissons la minimisation des coûts comme objectif de la PL. Le problème de PL résultant est formulé comme suit :

$$\begin{aligned}
 & \underset{p(s,g)}{\text{minimiser}} && \sum_{s,g} N_g p(s,g) c(s,g) \\
 & \text{sous les contraintes} && \sum_{s \in \mathcal{S}^R} N_g p(s,g) \rho(s,g) \geq R_g, \forall g \in \mathcal{G} \\
 & && \sum_{s,g} N_g p(s,g) \leq S_{\max} \\
 & && 0 \leq p(s,g) \leq 1, \forall s \in \mathcal{S}, g \in \mathcal{G} \\
 & && \sum_{s \in \mathcal{S}} p(s,g) = 1, \forall g \in \mathcal{G} \\
 & && \sum_{s \in \mathcal{S}^R} p(s,g) > 0, \forall g \in \mathcal{G}.
 \end{aligned} \tag{3.1}$$

$$\begin{aligned}
 & \text{Minimiser} && t \\
 & \text{sous les contraintes} && \sum_{s,g} \frac{w_g p(s,g) \rho(s,g) D(s,g)}{\sum_{s' \in \mathcal{S}^R} p(s',g) \rho(s',g)} \leq t \\
 & && - \sum_{s,g} \frac{w_g p(s,g) \rho(s,g) D(s,g)}{\sum_{s' \in \mathcal{S}^R} p(s',g) \rho(s',g)} \leq t \\
 & && \sum_{s,g} N_g p(s,g) c(s,g) \leq B \\
 & && \sum_{s \in \mathcal{S}^R} N_g p(s,g) \rho(s,g) \geq R_g, \forall g \in \mathcal{G} \\
 & && \frac{\sum_{s \in \mathcal{S}^R} p(s,g) \rho(s,g) D(s,g)}{\sum_{s \in \mathcal{S}^R} p(s,g) \rho(s,g)} - \frac{\sum_{s \in \mathcal{S}^R} p(s,h) \rho(s,h) D(s,h)}{\sum_{s \in \mathcal{S}^R} p(s,h) \rho(s,h)} \leq M \\
 & && \sum_{s,g} N_g p(s,g) \leq S_{\max} \\
 & && 0 \leq p(s,g) \leq 1, \forall s \in \mathcal{S}, g \in \mathcal{G} \\
 & && \sum_{s \in \mathcal{S}} p(s,g) = 1, \forall g \in \mathcal{G} \\
 & && \sum_{s \in \mathcal{S}^R} p(s,g) > 0, \forall g \in \mathcal{G} \\
 & && 0 \leq t.
 \end{aligned} \tag{3.2}$$

Pour résoudre le problème linéaire, nous utilisons la méthode du simplexe disponible en R dans le module *boot*. Notre algorithme proposé en deux étapes traite donc (3.1) dans la première étape. x_{LP}^* dénote la solution optimale obtenue dans la PL. Dans la deuxième étape, la solution x_{LP}^* est soumise à un

algorithme d'optimisation non linéaire comme point de départ afin de résoudre (3.2). À cette étape, nous utilisons les algorithmes non linéaires disponibles dans NLOPT (voir Johnson 2013), une bibliothèque de source ouverte pour l'optimisation non linéaire qui peut être appelée à partir de R dans le module *nloptr*. La deuxième étape du problème non linéaire non convexe de l'algorithme est exécutée seulement si le budget minimal requis trouvé à la première étape de la PL est plus petit ou égal au budget B disponible. Si le budget minimal est plus grand, il n'y a pas de solution possible au problème d'optimisation.

Étant donné que la performance de ces algorithmes dépend du problème, nous avons choisi de combiner deux algorithmes locaux de recherche afin d'accroître la vitesse de convergence. Des algorithmes d'optimisation globale sont disponibles dans la bibliothèque NLOPT, mais leur performance dans la résolution de notre problème était considérablement inférieure à celle des algorithmes d'optimisation locale sélectionnés. Les deux algorithmes de recherche locale sélectionnés sont COBYLA (*Constrained Optimization by Linear Approximations*), proposé par Powell (1998) (voir Roy 2007 pour une mise en œuvre en \mathcal{C}) et l'algorithme lagrangien augmenté AUGLAG, décrit dans Conn, Gould et Toint (1991) et Birgin et Martinez (2008). La méthode COBYLA construit des approximations linéaires successives de la fonction d'objectif et des contraintes au moyen d'un simplexe de $n + 1$ points (en n dimensions), et optimise ces approximations dans une région de confiance à chaque étape. La méthode AUGLAG combine la fonction d'objectif et les contraintes non linéaires en une seule fonction, c'est-à-dire l'objectif en plus d'une pénalité pour toute contrainte non respectée. La fonction résultante est ensuite transférée à un autre algorithme d'optimisation en tant que problème sans contrainte. Si la solution de ce sous-problème enfreint les contraintes, les pénalités sont accrues et le processus est répété. Le processus finit par converger vers la solution souhaitée, si elle existe.

Nous avons choisi d'utiliser la méthode MMA (*Method of Moving Asymptotes*, introduite dans Svanberg 2002), comme optimiseur local pour la méthode AUGLAG, en raison de sa performance dans nos expériences numériques. La stratégie sous-tendant la méthode MMA est expliquée ci-après. À chaque point \mathbf{x}_k , MMA forme une approximation locale à la fois convexe et séparable en utilisant le gradient de $f(\mathbf{x}_k)$ et les fonctions de contrainte, ainsi qu'un terme de pénalité quadratique pour rendre les approximations prudentes, par exemple des limites supérieures pour les fonctions exactes. L'optimisation de l'approximation mène à un nouveau point candidat \mathbf{x}_{k+1} . Si les contraintes sont respectées, le processus continue à partir du nouveau point \mathbf{x}_{k+1} . Sinon, le terme de pénalité est accru et le processus est répété.

Nous utilisons deux algorithmes de recherche locaux, parce que la méthode AUGLAG est plus efficace pour trouver le voisinage de l'optimum global, tandis que la méthode COBYLA offre une plus grande exactitude dans la recherche de l'optimum. En conséquence, la solution optimale de la PL est d'abord soumise à AUGLAG, puis, après un certain nombre d'itérations, lorsque l'amélioration de la valeur objective est inférieure à un seuil spécifié, la solution de la méthode AUGLAG est traitée par la méthode COBYLA pour une plus grande exactitude. Pour notre étude de cas, étant donné les exigences en matière de précision des statistiques obtenues dans le cadre de l'enquête (0,5 %), les résultats sont considérés comme suffisamment exacts si la valeur objective obtenue se situe à moins de 10^{-4} de l'optimum global. Tout autre gain d'exactitude est complètement éliminé par la variation d'échantillonnage et l'exactitude des paramètres d'entrée mêmes. Les calculs peuvent prendre jusqu'à quelques heures. Comme il n'est pas nécessaire de résoudre le problème d'optimisation durant la collecte des données, cela ne posera pas de problèmes dans la pratique.

4 L'étude du cas de l'Enquête sur la population active des Pays-Bas

Dans cette section, nous discutons de l'étude du cas de l'Enquête sur la population active (EPA) des Pays-Bas couvrant les années 2010 à 2012. Nous commençons par décrire brièvement le plan de sondage de l'EPA, puis nous décrivons les caractéristiques sélectionnées du plan de sondage et les sous-groupes de population sélectionnés. Nous expliquons ensuite comment nous avons estimé les principaux paramètres d'entrée du problème d'optimisation, à savoir les propensions à répondre, les propensions à avoir un numéro de téléphone inscrit, les coûts variables et les effets de méthode ajustés en lien avec deux plans de collecte repères différents. Après avoir estimé les paramètres, nous présentons les principaux résultats d'optimisation. Nous concluons par une discussion allant de la sensibilité des plans de collecte optimaux à l'inexactitude des paramètres d'entrée. Pour plus de détails, nous renvoyons à Calinescu et Schouten (2013b).

4.1 L'EPA des Pays-Bas de 2010 à 2012 – plan de sondage et remaniement

L'EPA des Pays-Bas est une enquête-ménage mensuelle à panel rotatif comportant cinq vagues à intervalles trimestriels. Elle se fonde sur un échantillon d'adresses utilisant un plan à deux degrés où les unités de premier degré sont les municipalités et les unités de deuxième degré sont les adresses. Un échantillon aléatoire simple stratifié est prélevé en fonction de l'âge et de l'ethnicité des membres du ménage et des chômeurs inscrits. Tous les ménages, jusqu'à un maximum de huit, qui habitent à l'adresse sont invités à participer à l'enquête. Tous les membres du ménage de 15 ans ou plus sont admissibles, car ils forment la population active potentielle. Les questions de l'EPA portent sur différents sujets, dont la situation d'emploi, la profession, les heures de travail et le niveau de scolarité, mais la principale statistique est le taux de chômage.

Jusqu'en 2010, l'EPA était constituée d'une première vague d'interviews en personne, puis de vagues subséquentes d'interviews par téléphone. Pour différentes raisons, la plus importante étant les coûts, la première vague a subi un remaniement majeur. Les autres vagues sont restées inchangées, à part quelques modifications relativement mineures aux questionnaires. Le remaniement s'est déroulé en deux étapes correspondant aux deux modes d'enquête ajoutés : d'abord le téléphone, puis le Web. Lors de l'étape initiale, la première vague d'interviews en personne a été remplacée par un plan de collecte à modes simultanés où tous les ménages ayant au moins un numéro de téléphone inscrit ont été affectés au mode de collecte par téléphone et tous les autres ménages, au mode de collecte par interview en personne. Les numéros de téléphone inscrits comprennent les numéros de téléphone filaire et de téléphone mobile qui peuvent être achetés auprès de fournisseurs commerciaux. Dans la deuxième étape, la collecte simultanée par téléphone et en personne a été précédée d'une invitation en ligne, produisant une combinaison de plans de collecte à modes séquentiels et à modes simultanés. Tous les ménages ont reçu une invitation à participer par la voie d'un questionnaire en ligne. Les ménages non répondants ont été contactés par téléphone si un numéro inscrit était disponible et en personne autrement. La première étape a eu lieu en 2010 et la deuxième, en 2012. Durant ces deux années, de vastes échantillons parallèles ont été prélevés afin d'évaluer les effets de méthode entre les plans de collecte sur le taux de chômage. L'essai parallèle de 2010 comparait l'ancien plan de collecte au plan intermédiaire à modes simultanés, tandis que celui de 2012 comparait le plan intermédiaire au plan final à trois modes.

Le remaniement n'a pas modifié la stratégie de collecte de données par mode. Dans toutes les années, la stratégie de contact en personne pour la première vague de l'EPA prévoit un maximum de six visites à l'adresse, et les contacts varient selon les jours de la semaine et les heures de la journée. Si aucun contact

n'est établi après la sixième visite, l'adresse est traitée comme un non-contact. La stratégie de contact par téléphone prévoit trois séries de trois appels. Les trois séries sont désignées comme étant des tentatives de contact et représentent trois quarts de travail d'intervieweurs différents. Dans chaque quart, le numéro de téléphone est composé à trois reprises à intervalles d'environ une heure. La stratégie Web prévoit l'envoi d'une lettre initiale contenant un code de connexion à un site Web, puis de deux lettres de rappel envoyées à intervalles d'une semaine.

Nous utilisons les données de la première vague de l'EPA, qui s'est déroulée de 2010 à 2012, afin d'estimer différents paramètres d'entrée pour le modèle d'optimisation. Afin de ne pas compliquer l'exposition, et comme les vagues subséquentes n'ont pas été remaniées, nous nous limitons aux effets de méthode sur les estimations du taux de chômage fondées sur la première vague seulement. Cependant, le remaniement de la première vague pourrait très bien avoir influencé le recrutement et la réponse aux vagues 2 à 5. Dans des études de suivi réalisées à *Statistics Netherlands*, les propensions à participer aux vagues subséquentes étaient incluses dans le problème d'optimisation, mais nous n'abordons pas cette question ici. Les données de l'EPA ont été complétées par des données de deux registres administratifs : POLIS et UWV. Le registre POLIS contient des renseignements sur les emplois, les prestations, le revenu d'emploi et les avantages sociaux. Le registre UWV contient la liste des personnes qui se sont inscrites dans un bureau de chômage et qui ont demandé des prestations d'assurance-chômage. Les deux registres contiennent des variables pertinentes pour l'EPA et seront utilisés pour stratifier la population.

4.2 L'ensemble de stratégies

Les essais parallèles de l'EPA nous permettent d'examiner un problème d'optimisation multimodal selon différentes stratégies de collecte à mode unique et à modes mixtes séquentiels. Dans la section qui suit, nous désignons les modes de collecte par téléphone et en personne par les abréviations *Tel* et *F2F*, respectivement. Bien que la stratégie séquentielle *Web* → *F2F* ne s'applique qu'aux ménages nombreux et aux ménages sans numéro de téléphone inscrit, nous incluons cette stratégie dans l'optimisation.

Comme les contacts subséquents en personne et par téléphone coûtent relativement beaucoup plus cher que les interviews initiales, nous limitons le nombre d'appels. Pour *Tel*, nous limitons ce nombre à deux et pour *F2F*, à trois. Ces valeurs sont motivées par les données d'enquête historiques, par exemple après ces nombres d'appels, le coût par appel augmente rapidement. *Tel2* et *F2F3* dénotent les stratégies où le nombre d'appels est limité, tandis que *Tel2+* et *F2F3+* représentent les stratégies où aucune limite n'est imposée et où une stratégie de contacts réguliers est appliquée. Nous comprenons qu'imposer une limite n'est pas la même chose que limiter le nombre d'appels dans la pratique, particulièrement pour les interviews en personne. La réduction du nombre d'appels peut amener les intervieweurs ou leur personnel à modifier leur comportement et à espacer les appels différemment. *Statistics Netherlands* considère *Tel2* et *F2F3* comme des stratégies tronquées avec des périodes de collecte plus courtes, par exemple deux semaines plutôt que quatre. La charge de travail des intervieweurs est donc allégée après la période de collecte spécifiée à l'avance. De ce point de vue, il est plus raisonnable de supposer que la stratégie de contact optimale durant les deux premières semaines de *F2F3+* n'est pas si différente de celle de *F2F3*. Nous pouvons quand même nous attendre à ce que les propensions à répondre et les coûts réalisés des stratégies avec limite diffèrent de leurs propensions et de leurs coûts simulés. L'ensemble de stratégies devient alors :

$$S = \{Web, Tel2, Tel2+, F2F3, F2F3+, Web \rightarrow Tel2, \\ Web \rightarrow Tel2+, Web \rightarrow F2F3, Web \rightarrow F2F3+, \Phi\}, \quad (4.1)$$

où Φ dénote la stratégie autre que d'échantillonnage.

Les essais parallèles pour l'EPA de 2010 et 2012 étaient importants. Dans les deux années, la taille de l'échantillon de l'EPA a été doublée pendant une période de six mois. Les paramètres estimatifs restent sujets à la variation d'échantillonnage et, dans le cas des stratégies *Web* \rightarrow *F2F*, peut-être aussi au biais. Nous revenons là-dessus dans la section 4.6.

4.3 Les groupes de population

Afin de stratifier la population, nous avons utilisé les variables de pondération ordinaires de l'EPA comme point de départ : l'inscription au bureau de chômage, l'âge, la taille du ménage, l'ethnicité et l'emploi inscrit. Le recoupement des 5 variables a produit 48 strates de population (présence ou absence d'un chômeur inscrit dans le ménage x 3 catégories d'âge x 2 catégories de taille de ménage x 2 catégories d'ethnicité x présence ou absence d'un emploi inscrit dans le ménage). Ces strates ont été ramenées à neuf strates disjointes fondées sur le comportement de réponse et les effets de mode :

1. *Ménages avec chômeur inscrit* : ménages dont au moins un membre est inscrit dans un bureau de chômage (7,5 % de la population).
2. *Ménages de plus de 65 ans sans emploi* : ménages comptant un maximum de 3 personnes de 15 ans ou plus non inscrites dans un bureau de chômage, sans emploi, dont au moins un membre a 65 ans ou plus (19,8 % de la population).
3. *Ménages jeunes sans emploi* : ménages comptant un maximum de 3 personnes de 15 ans ou plus non inscrites dans un bureau de chômage, sans emploi, dont tous les membres ont moins de 65 ans et au moins un membre a entre 15 et 26 ans (2,4 % de la population).
4. *Ménages non occidentaux sans emploi* : ménages comptant un maximum de 3 personnes de 15 ans ou plus non inscrites dans un bureau de chômage, sans emploi, dont tous les membres ont entre 26 et 65 ans et au moins un membre est d'ethnicité non occidentale (1,5 % de la population).
5. *Ménages occidentaux sans emploi* : ménages comptant un maximum de 3 personnes de 15 ans ou plus non inscrites dans un bureau de chômage, sans emploi, dont tous les membres ont entre 26 et 65 ans et sont d'ethnicité occidentale (11,0 % de la population).
6. *Ménages jeunes avec emploi* : ménages comptant un maximum de 3 personnes de 15 ans ou plus non inscrites dans un bureau de chômage, avec au moins un emploi, dont tous les membres ont moins de 65 ans et au moins un membre a entre 15 et 26 ans (15,6 % de la population).
7. *Ménages non occidentaux avec emploi* : ménages comptant un maximum de 3 personnes de 15 ans ou plus non inscrites dans un bureau de chômage, avec au moins un emploi, dont tous les membres ont plus de 26 ans et au moins un membre est d'ethnicité non occidentale (3,9 % de la population).
8. *Ménages occidentaux avec emploi* : ménages comptant un maximum de 3 personnes de 15 ans ou plus non inscrites dans un bureau de chômage, avec au moins un emploi, dont tous les membres ont plus de 26 ans et sont d'ethnicité occidentale (33,5 % de la population).
9. *Ménages nombreux* : ménages comptant plus de 3 personnes de 15 ans ou plus non inscrites dans un bureau de chômage (4,9 % de la population).

Des étiquettes informelles ont été attribuées aux neuf strates de population afin de faciliter l'interprétation. Il faut toutefois noter que les strates 7, 8 et 9 peuvent inclure des membres du ménage qui ont plus de 65 ans. De plus, certains sous-groupes résultent de la réduction de certaines strates. Par exemple, nous pouvons trouver les ménages comptant au moins un emploi en combinant les strates 6, 7 et 8, et les ménages ne comptant pas plus de 3 membres de 15 ans ou plus en combinant toutes les strates de 1 à 8.

Dans le modèle d'optimisation, différentes stratégies et différentes probabilités d'affectation des stratégies ont été appliquées aux neuf strates. Nous avons imposé d'autres contraintes de précision suivant l'EPA ordinaire à une autre stratification. Il devait y avoir des nombres minimaux de répondants selon l'âge, l'ethnicité et l'inscription dans un bureau de chômage. Nous renvoyons de nouveau le lecteur à Calinescu et Schouten (2013b) pour plus de détails sur ces strates et les seuils de précision correspondants.

4.4 L'estimation des paramètres d'entrée

Les paramètres d'entrée du problème d'optimisation multimodal sont les propensions à répondre des sous-populations par stratégie, les propensions à avoir un numéro de téléphone inscrit des sous-groupes, les coûts des sous-groupes par unité d'échantillonnage et par stratégie et les effets de méthode ajustés des sous-groupes par stratégie. Nous faisons une estimation de chaque ensemble de paramètres dans les sous-sections qui suivent. De plus amples détails figurent à l'annexe A.

Trois configurations peuvent se dessiner au moment de l'estimation des paramètres d'entrée : 1) la stratégie est directement observée dans les données d'enquête historiques; 2) la stratégie est partiellement observée dans les données d'enquête historiques, c'est-à-dire seulement pour un sous-ensemble de l'échantillon; et 3) la stratégie n'est pas observée du tout.

Pour l'étude du cas de l'EPA, la première configuration s'applique aux stratégies *Web*, *Tel2+*, *F2F3+*, *Web* → *Tel2+*, la deuxième s'applique à *Web* → *F2F3+* et la troisième, à *Tel2*, *F2F3*, *Web* → *Tel2* et *Web* → *F2F3*. Les plans de sondage à modes mixtes séquentiels avec les interviews en personne comme mode de suivi sont observés uniquement pour les ménages sans numéro de téléphone inscrit et se rattachent aux configurations 2 ou 3 selon que le nombre d'appels est limité ou non. Nous avons essayé de composer avec la configuration 2 en modélisant les paramètres d'entrée en fonction des différences observées dans les paramètres entre *Tel2+* et *F2F3+*. Nous avons supposé que le ratio de la propension à répondre entre *F2F3(+)* et *Tel2+* pour les ménages ayant un numéro de téléphone inscrit peut s'appliquer à *Web* → *F2F3(+)* et *Web* → *Tel2+*. De plus, dans l'estimation, nous avons supposé que les stratégies limitant le nombre d'appels ressemblent aux stratégies simulées, en ce sens qu'elles limitent artificiellement les stratégies prévoyant le nombre intégral d'appels au plafond spécifié. Nous avons donc essayé de traiter la configuration 3 en censurant les stratégies. Calinescu et Schouten (2013b) expliquent en détail ces étapes de modélisation.

Pour l'effet de méthode $D(s, g)$, deux repères ont été sélectionnés : $BM_1 = \bar{y}_{F2F3+}$ et $BM_2 = 1/3 * (\bar{y}_{Web} + \bar{y}_{Tel2+} + \bar{y}_{F2F3+})$, où \bar{y}_{mode} représente le taux de chômage moyen estimé au moyen du mode d'enquête indiqué. Le premier repère suppose que le taux de chômage moyen estimé au moyen d'un plan unimodal de collecte en personne représente le taux de chômage cible. Le deuxième repère suppose qu'il n'y a pas de mode préféré et attribue donc un poids égal à chacun des trois modes. Le repère *F2F3+* a été choisi parce qu'il s'agit du mode de collecte traditionnel pour la première vague de l'EPA et qu'il détermine donc les séries temporelles de l'EPA jusqu'en 2010. Nous croyons également que ce

mode produit le biais de non-réponse le plus faible pour de nombreuses enquêtes (voir par exemple Klausch et coll. 2013a). Il est toutefois difficile de savoir si $F2F3+$ doit être considéré comme le mode ayant le biais de mesure le plus faible. Nous avons donc établi un deuxième repère afin d'examiner l'importance du choix de repère.

Les erreurs standard pour les paramètres d'entrée estimatifs ont été calculées par approximation au moyen d'un rééchantillonnage bootstrap par strate d'échantillonnage, d'après le modèle d'échantillonnage stratifié.

4.5 Les résultats d'optimisation

Dans cette section, nous explorons la répartition optimale et l'effet de méthode minimal pour différents niveaux budgétaires, entre les niveaux des effets de méthode par strate et les niveaux de taille de l'échantillon

$$B \in \{160\,000; 170\,000; 180\,000\}$$

$$M \in \{1\%; 0,5\%; 0,25\%\}$$

$$S_{\max} \in \{9\,500; 12\,000; 15\,000\}.$$

L'annexe B présente les effets de méthode minimaux pour les différents niveaux et/ou les deux plans de collecte repères, BM_1 et BM_2 . Par souci de concision, nous nous attardons surtout aux résultats pour BM_1 , qui est l'ancien plan de l'EPA. Les valeurs réelles pour le plan de l'EPA trimodal ordinaire non adaptatif sont

$$B = 170\,000 \quad M = 3,00\%$$

$$S_{\max} = 11\,000 \quad \bar{D}^{BM_1} = -0,15\%.$$

Nous pouvons tirer deux principales conclusions des résultats. Premièrement, le plan de collecte adaptatif peut réduire l'effet de méthode global absolu sur les deux repères tout en respectant une contrainte stricte sur l'effet de méthode maximal entre les strates et en maintenant le budget au niveau actuel. La seule contrainte qui doit être relâchée afin de réduire l'effet de méthode global est la taille maximale de l'échantillon. Deuxièmement, les effets de méthode globaux minimaux obtenus pour le plan repère BM_2 , sont plus faibles que ceux obtenus pour le plan BM_1 , à l'exception de $S_{\max} = 9\,500$. Cette différence est le résultat de valeurs généralement plus petites et plus semblables des effets de la méthode par strate $D(s, g)$. Nous pouvons examiner l'impact de la contrainte de la taille de l'échantillon en comparant les répartitions optimales pour $S_{\max} = 9\,500$ et $S_{\max} = 15\,000$. Supposons que les seuils sont établis à $B = 170\,000$, $M = 1\%$ et BM_1 . Les figures 4.1 et 4.2 présentent les probabilités de répartition optimale par strate et par stratégie en supposant qu'une unité soit échantillonnée. Chaque figure peut être considérée comme une matrice où chaque ligne représente une des stratégies en S^R et chaque colonne, une des neuf strates décrites dans la section 4.3, par exemple g_1 est la strate des chômeurs inscrits. Chaque cellule de la matrice, c'est-à-dire l'intersection d'une ligne et d'une colonne, montre la probabilité que la stratégie correspondante soit appliquée à la strate correspondante. Les probabilités sont illustrées sous forme de barres. Plus la barre est large, plus la proportion de la strate affectée à la stratégie est importante. La somme des probabilités est égale à un sur les stratégies, c'est-à-dire sur les lignes. Les valeurs exactes de 20 % ou plus sont données dans les barres. Les figures 4.1 et 4.2 montrent une transition évidente des probabilités d'affectation lorsque la taille de l'échantillon peut augmenter,

par exemple la strate 6 (ménages de jeunes avec emploi) est presque entièrement affectée au plan *Web*, tandis que les strates 8 (ménages occidentaux avec emploi) et 9 (ménages nombreux) passent d'une stratégie à modes séquentiels à une stratégie de collecte par interview en personne seulement.

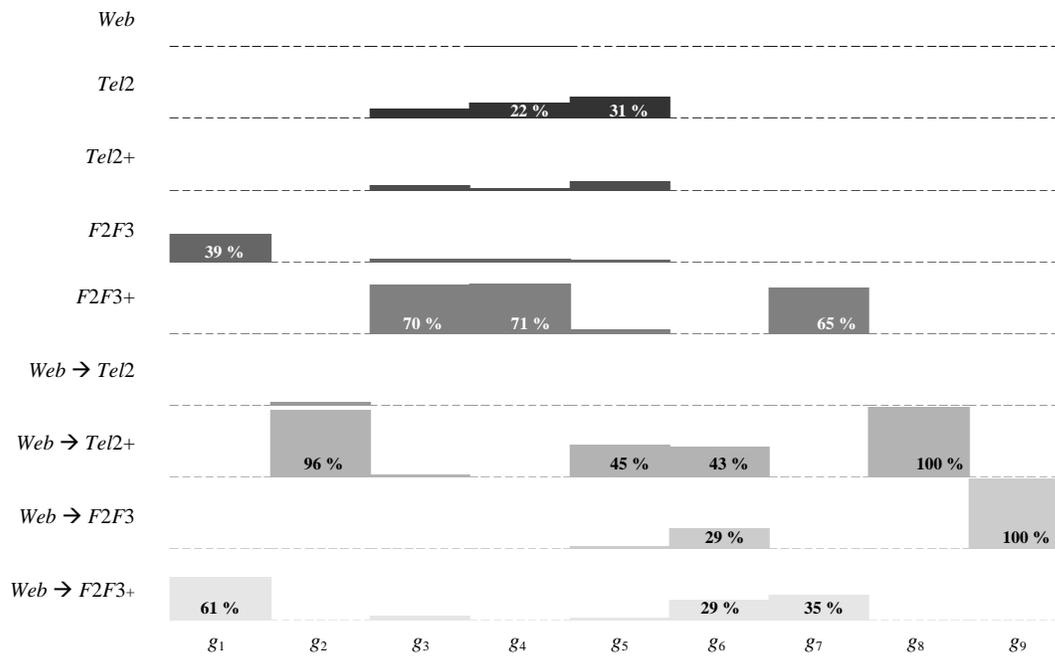


Figure 4.1 Répartition des stratégies étant donné une solution optimale pour $S_{\max} = 9500$, $B = 170000$, $M = 1\%$, BM_1 . Le trait pointillé indique que $p(s, g) = 0$.

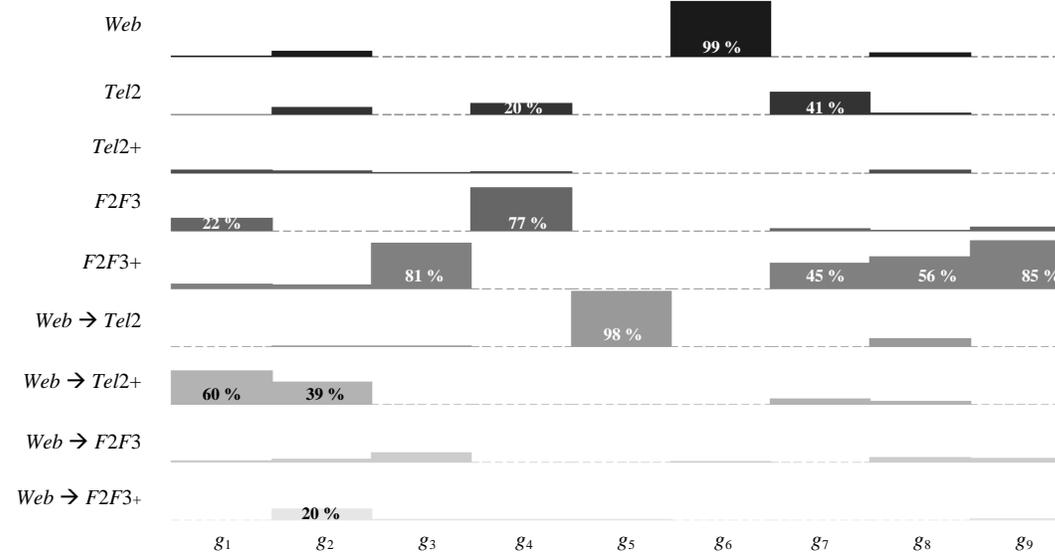


Figure 4.2 Répartition des stratégies étant donné une solution optimale pour $S_{\max} = 15000$, $B = 170000$, $M = 1\%$, BM_1 . Le trait pointillé indique que $p(s, g) = 0$.

L'impact du budget disponible est très clair pour $S_{\max} = 12\,000$ et BM_1 , où l'effet de méthode global minimal passe de 0,10 % pour $B = 160\,000$ à 0,01 % pour $B = 180\,000$. Les probabilités de répartition optimale sont indiquées dans les figures 4.3 et 4.4. Lorsque le budget augmente, nous observons une transition des stratégies de collecte par téléphone seulement à une combinaison de stratégies de collecte par interview en personne seulement et, ce qui est plutôt étonnant, de stratégies Web seulement.

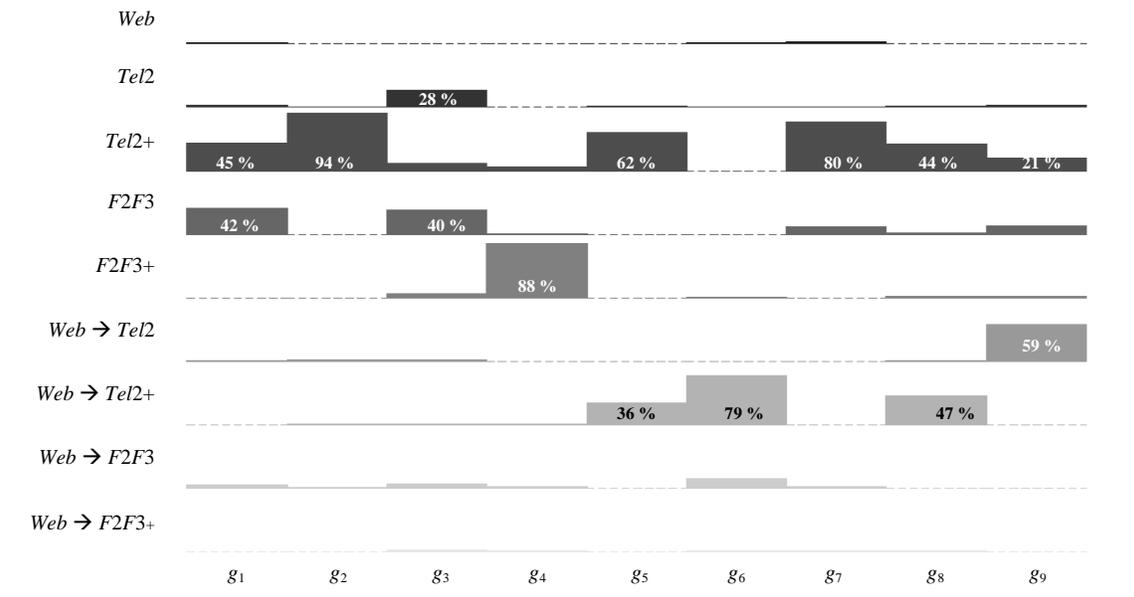


Figure 4.3 Répartition des stratégies étant donné une solution optimale pour $S_{\max} = 12\,000$, $B = 160\,000$, $M = 1\%$, BM_1 . Le trait pointillé indique que $p(s, g) = 0$.

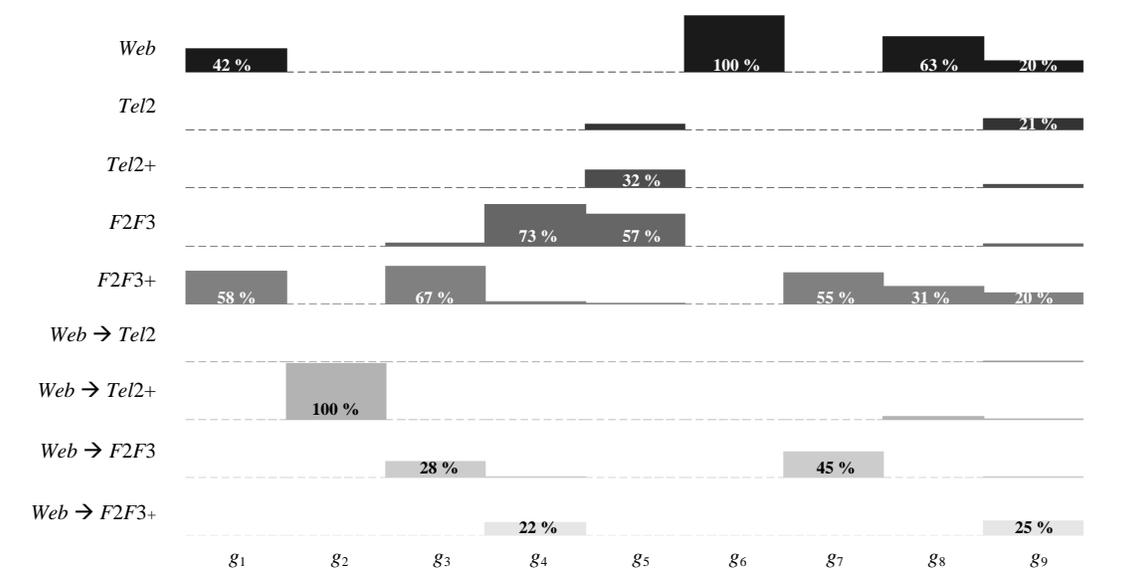


Figure 4.4 Répartition des stratégies étant donné une solution optimale pour $S_{\max} = 12\,000$, $B = 180\,000$, $M = 1\%$, BM_1 . Le trait pointillé indique que $p(s, g) = 0$.

Il est possible de tester différents scénarios en utilisant une vaste gamme de valeurs seuils, ce qui pourra faire l'objet d'autres articles. Nous concluons en mentionnant que les répartitions optimales dont les probabilités d'affectation sont nombreuses mais faibles aboutissent à des processus de données très difficiles à exécuter. On pourrait ajouter des seuils plus faibles de probabilités d'affectation afin d'éviter les stratégies qui s'appliquent seulement à un petit nombre de cas.

4.6 La robustesse des plans de collecte optimaux

Dans cette section, nous discutons brièvement de la robustesse des plans de collecte optimaux. Les analyses de sensibilité dépassent la portée de cet article et font actuellement l'objet de recherches.

Pour estimer les propensions à répondre, les propensions à avoir un numéro de téléphone inscrit, les coûts par unité d'échantillonnage et les effets de méthode ajustés, nous émettons quatre principales hypothèses, en plus de celles concernant la fonction de lien logistique entre la réponse et la non-réponse, la présence et l'absence de numéro de téléphone inscrit et les variables auxiliaires. Ces hypothèses sont les suivantes :

1. Modèle pour $Web \rightarrow F2F3$ et $Web \rightarrow F2F+$: Ces deux stratégies ont été utilisées seulement pour les ménages sans numéro de téléphone inscrit.
2. Estimation par censurement des stratégies limitant le nombre d'appels : Des stratégies limitant le nombre d'appels n'ont pas été mises en application, et nous supposons que leurs propensions à répondre et leurs coûts peuvent être estimés par censurement des stratégies englobant tous les modes de contact.
3. Affectation des coûts linéaires à des stratégies : Nous supposons que les coûts par unité d'échantillonnage ne dépendent pas de la taille de l'échantillon affecté à une stratégie.
4. Stabilité au fil du temps des effets de méthode entre 2010 et 2012 : Comme les essais parallèles ont été effectués en deux étapes, les effets de méthode pour certaines stratégies ont été estimés en deux étapes. Nous supposons implicitement que les effets de méthode pour ces plans n'ont pas changé entre 2010 et 2012.

Qui plus est, tous les paramètres d'entrée estimatifs sont assujettis à une variation d'échantillonnage. En conséquence, nous nous attendons à ce que les plans optimaux puissent subir certaines variations en raison de l'inexactitude des paramètres. Afin d'évaluer la robustesse des plans optimaux, nous proposons deux types d'analyses de sensibilité :

- Une optimisation répétée des paramètres d'entrée tirés des données rééchantillonnées. Autrement dit, toutes les données historiques sont rééchantillonnées à de multiples reprises et une optimisation est effectuée à chaque tirage. Les valeurs optimales résultantes pour la qualité et les coûts ainsi que la combinaison de stratégies des plans optimaux peuvent donc être comparées entre les différents tirages.
- L'évaluation de la performance du plan optimal pour les données rééchantillonnées. Autrement dit, une optimisation est effectuée en fonction des données historiques observées. Toutes les données historiques sont ensuite rééchantillonnées, et les paramètres d'entrée de l'optimisation sont recalculés pour chaque tirage. Le plan optimal est appliqué à chaque ensemble de paramètres d'entrée, et les valeurs de qualité et de coût correspondantes sont calculées. Enfin,

les propriétés statistiques des valeurs de qualité et de coût sont évaluées pour tous les tirages des paramètres d'entrée.

Des analyses exploratoires de la sensibilité indiquent que la combinaison de stratégies des plans optimaux affiche des variations relativement importantes, mais que les effets de méthode optimaux \bar{D}^{BM} sont très stables. Cela implique que l'effet de méthode, en tant que fonction d'objectif, est une fonction relativement lisse.

5 Discussion

Nous avons construit un problème d'optimisation multimodal qui étend le cadre des plans de collecte adaptatifs aux plans d'enquête à modes mixtes originaux ou remaniés. Ce cadre est particulièrement utile lorsqu'on s'attend à ce que les effets de méthode attribuables à un changement de plan aient un impact sur la comparabilité et l'exactitude des statistiques. À notre connaissance, il s'agit de la première tentative de recherche de ce genre et celle-ci peut servir de point de départ à la minimisation des effets de méthode sous réserve des coûts et d'autres contraintes.

Dans le modèle d'optimisation, nous avons inclus trois critères de qualité, un critère de coût et un critère logistique. Les critères de qualité sont le nombre de répondants dans les strates d'échantillonnage, qui indique le degré de précision, l'effet de méthode global ajusté absolu, qui est le changement de niveau causé par la différence entre le plan retenu et le plan repère et peut être considéré comme la comparabilité dans le temps, et la différence absolue maximale dans les effets de méthode sur les sous-populations importantes, qui peut être considérée comme la comparabilité entre les domaines de population. Le critère de coût est le budget total de l'enquête. Le critère logistique est la taille de l'échantillon, qui doit être limitée afin d'éviter un épuisement rapide du cadre d'échantillonnage. Le troisième critère de qualité, à savoir la différence absolue maximale dans les effets de méthode sur les sous-populations, est non linéaire dans les variables de décision (probabilités d'affectation des stratégies) et rend le problème d'optimisation difficile à résoudre en raison de la complexité des calculs. Ce critère complique le problème, mais il s'agit d'une contrainte utile qui est souvent mise de l'avant par les analystes et les utilisateurs des enquêtes. Dans les plans remaniés ordinaires, ce critère est souvent ignoré, et le plan à modes mixtes de l'EPA des Pays-Bas entraîne des différences relativement importantes dans les effets de méthode entre les sous-populations. De toute évidence, certains des critères peuvent être omis et d'autres critères logistiques, de qualité ou de coût peuvent être ajoutés. Dans un suivi de cette étude à *Statistics Netherlands*, plusieurs autres critères, surtout logistiques, sont pris en considération.

Dans le modèle d'optimisation, l'accent était mis sur la maximisation de la qualité, reflétée par la comparabilité dans le temps, sous réserve des contraintes de coût et d'autres contraintes logistiques et de qualité. L'objectif de l'optimisation peut toutefois être modifié, et chacune des contraintes pourrait faire fonction d'objectif. Nous pourrions, par exemple, minimiser le coût sous réserve des contraintes logistiques et de qualité. Nous pourrions aussi adopter une approche élargie et effectuer plusieurs optimisations pour différents niveaux de budget et de qualité afin de développer une optique multidimensionnelle informative pouvant servir de base à une décision.

Notre tentative doit être considérée comme un première étape vers des plans de sondage adaptatifs à modes mixtes. Il reste différentes questions de nature méthodologique et pratique à résoudre. Premièrement, notre approche convient aux enquêtes comportant seulement quelques statistiques clés.

Une optimisation peut être effectuée et une décision pondérée peut être prise pour chacune de ces statistiques. Cette approche n'est pas possible pour une enquête portant sur un vaste éventail de statistiques. Deuxièmement, l'optimisation dépend en grande partie de l'exactitude des paramètres d'entrée, c'est-à-dire des probabilités de réponse estimatives, des probabilités d'inscription du numéro de téléphone, des paramètres de coût et des effets de mode dans le cas qui nous intéresse. Il est important d'évaluer la sensibilité des résultats d'optimisation à l'exactitude de ces paramètres. Nous pouvons supposer que la fonction d'objectif est relativement lisse en ce qui concerne ces paramètres, mais il est quand même important de faire des analyses de sensibilité. Troisièmement, il est essentiel de tenir compte de la variation d'échantillonnage de la qualité et des coûts réalisés du plan optimisé lorsque de multiples vagues d'une enquête sont réalisées. Cette variation peut être importante et réduire la valeur d'une optimisation précise. Quatrièmement, une fois les critères non linéaires ajoutés au problème, il faut compter sur des solveurs avancés dans le logiciel statistique. Même lorsque de tels solveurs sont utilisés, la convergence vers l'optimum global n'est généralement pas garantie et il faut se contenter des optima locaux. C'est pourquoi il est important de choisir un ensemble utile de points de départ, y compris des points de départ qui correspondent aux plans de sondage actuels. Les questions pratiques concernent le nombre de strates de population, le nombre de stratégies et la coordination avec d'autres enquêtes. Les systèmes et outils d'administration des enquêtes peuvent appuyer les plans de collecte adaptatifs, mais ces plans sont plus difficiles à surveiller et à analyser. De plus, l'adaptation des modes d'enquête affecte l'importance et la forme de la charge de travail des intervieweurs, qui pourraient contacter seulement une tranche précise des sous-populations.

Un aspect important des plans de collecte adaptatifs est l'utilisation d'estimations pour toutes sortes de paramètres d'entrée tels que les propensions à répondre, les coûts variables par unité d'échantillonnage et les effets de méthode entre les plans. Ces estimations pourraient être difficiles à obtenir et n'être appuyées que par des données d'enquête historiques faibles. Il y a alors quatre options : chercher des enquêtes semblables appuyées par des données historiques, demeurer modestes et limités dans le choix des caractéristiques du plan de sondage, prévoir une période de transition durant laquelle des études pilotes et des essais parallèles sont réalisés, et élaborer un cadre d'apprentissage et de mise à jour des paramètres. En particulier, lorsque *Web* est un des modes d'enquête du plan, il pourrait y avoir un manque de données historiques à l'appui des estimations dans de nombreux pays (voir par exemple Mohorko, de Leeuw et Hox 2013). Soulignons également que les paramètres d'entrée peuvent changer graduellement au fil du temps et nécessiter une mise à jour continue. Cependant, cela ne diffère en rien d'une enquête non adaptative, sauf que des estimations sont maintenant nécessaires pour les sous-populations concernées plutôt que pour l'ensemble de la population. Enfin, nous constatons que les plans optimisés adaptatifs, comme les plans optimisés non adaptatifs, permettent d'obtenir la qualité et les coûts moyens prévus. En raison de la variation d'échantillonnage, la qualité et les coûts réalisés varieront, et des événements imprévus pourraient entraîner des écarts. Il reste donc nécessaire d'assurer une surveillance et de réagir aux événements imprévus.

Les futures recherches devront se pencher sur la robustesse des plans de collecte adaptatifs et examiner d'autres critères logistiques, de qualité et de coût. Il est également important de reproduire cette étude afin de déterminer si l'investissement en termes de collecte de données supplémentaires et d'optimisation explicite en vaut la peine. Le but ultime de cette recherche est une stratégie de collecte de données qui favorise l'apprentissage et la mise à jour des paramètres d'optimisation et d'entrée et qui appuie des

analyses coûts-avantages efficaces et efficaces dans les plans d'enquête à modes mixtes initiaux et remaniés. Une approche bayésienne semble plus prometteuse à cette fin.

Remerciements

Les auteurs désirent remercier M. Sandjai Bhulai (VU University Amsterdam) pour ses commentaires constructifs sur le cadre mathématique présenté dans cet article. Ils remercient également Boukje Janssen (CBS) et Martijn Souren (CBS) pour leur traitement des données d'échantillonnage brutes à analyser, ainsi que Joep Burger (CBS) pour ses commentaires qui ont aidé à améliorer cet article.

Annexe A

Estimations des paramètres d'entrée

Dans la section 4.4, nous expliquons l'estimation des paramètres d'entrée pour les stratégies observées seulement en partie lors des essais parallèles. Nous fournissons ici les estimations pour les propensions à répondre, les propensions à avoir un numéro de téléphone inscrit, les coûts variables par unité d'échantillonnage et les effets de méthode ajustés. Les erreurs types pour tous les paramètres ont été estimées par rééchantillonnage bootstrap.

Le tableau A2 présente les propensions à répondre estimatives $\rho(s, g)$ tirées des données disponibles et les erreurs types correspondantes. Le tableau A1 montre la propension estimative à avoir un numéro de téléphone inscrit $\lambda(g)$.

Tableau A1

Propensions estimatives des membres du groupe $g \in \mathcal{G}$ à avoir un numéro de téléphone inscrit avec les erreurs types correspondantes entre parenthèses

\mathcal{G}	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9
$\lambda(g)$	38,1 % (0,9)	76,4 % (1,6)	30,2 % (2,0)	22,4 % (2,2)	60,0 % (1,1)	38,9 % (0,7)	32,0 % (1,3)	53,4 % (0,6)	62,4 % (1,2)

Tableau A2

Propensions estimatives à répondre par stratégie s et par groupe g avec les erreurs types correspondantes entre parenthèses

$\rho(s, g)$	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9
<i>Web</i>	23,2 % (0,3)	23,6 % (0,6)	15,5 % (0,6)	10,8 % (0,6)	27,9 % (0,4)	27,7 % (0,2)	17,5 % (0,5)	36,7 % (0,2)	22,4 % (0,5)
<i>Tel2</i>	12,2 % (0,5)	31,4 % (1,1)	8,5 % (0,8)	4,7 % (0,8)	19,7 % (0,6)	13,3 % (0,4)	7,2 % (0,5)	18,1 % (0,4)	21,2 % (0,8)
<i>Tel2+</i>	20,8 % (0,6)	41,3 % (1,1)	15,2 % (1,0)	8,6 % (1,0)	31,1 % (0,7)	23,8 % (0,5)	14,3 % (0,7)	33,3 % (0,5)	37,5 % (0,9)
<i>F2F3</i>	43,5 % (1,5)	53,5 % (1,7)	42,2 % (2,4)	34,1 % (2,4)	45,1 % (1,1)	45,3 % (0,9)	35,9 % (1,5)	46,7 % (0,7)	54,6 % (1,4)
<i>F2F3+</i>	52,4 % (1,3)	58,3 % (1,6)	51,0 % (2,5)	41,2 % (2,2)	51,2 % (1,1)	54,9 % (0,8)	46,0 % (1,4)	56,8 % (0,7)	61,4 % (1,3)
<i>Web</i> → <i>Tel2</i>	28,3 % (0,4)	41,0 % (0,8)	20,2 % (0,7)	13,9 % (0,8)	36,3 % (0,4)	34,0 % (0,3)	20,8 % (0,5)	44,5 % (0,3)	23,1 % (0,5)
<i>Web</i> → <i>Tel2+</i>	32,8 % (0,4)	48,4 % (0,7)	23,8 % (0,8)	17,5 % (0,9)	42,1 % (0,5)	41,1 % (0,3)	25,8 % (0,6)	52,1 % (0,3)	24,4 % (0,5)
<i>Web</i> → <i>F2F3</i>	46,3 % (0,5)	57,7 % (1,0)	38,6 % (1,0)	32,7 % (1,0)	50,0 % (0,6)	51,0 % (0,4)	39,3 % (0,7)	58,9 % (0,4)	50,0 % (0,5)
<i>Web</i> → <i>F2F3+</i>	49,8 % (0,5)	58,3 % (0,9)	43,4 % (0,9)	36,6 % (0,9)	52,6 % (0,5)	54,7 % (0,4)	44,3 % (0,6)	62,0 % (0,4)	54,2 % (0,5)

Pour l'effet de méthode $D(s, g)$, deux valeurs repères ont été sélectionnées après consultation des praticiens, à savoir $BM_1 = \bar{y}_{F2F3+}$ et $BM_2 = 1/3 * (\bar{y}_{Web} + \bar{y}_{Tel2+} + \bar{y}_{F2F3+})$, où \bar{y}_{mode} représente le taux de chômage moyen estimé par le mode d'enquête indiqué. Les tableaux A3 et A4 présentent les effets de méthode estimatifs par rapport aux deux valeurs repères, avec leurs erreurs types.

Les estimations des coûts variables par unité d'échantillonnage avec les erreurs types estimatives sont présentées au tableau A5. Les coûts sont exprimés par rapport à la stratégie $F2F3+$, qui est fixée à un.

Tableau A3

Effets de méthode estimatifs par rapport à la valeur repère $BM_1 = \bar{y}_{F2F3+}$ avec les erreurs types correspondantes entre parenthèses

$D^{BM_1}(s, g)$	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9
<i>Web</i>	1,5 % (1,0)	0,0 % (0,5)	-2,3 % (1,5)	-4,5 % (3,1)	0,9 % (0,7)	-0,4 % (0,4)	-2,2 % (1,5)	0,6 % (0,5)	-0,4 % (0,6)
<i>Tel2</i>	-0,2 % (0,7)	-0,1 % (0,1)	-2,6 % (0,9)	-6,8 % (1,8)	-1,0 % (0,4)	-0,9 % (0,3)	-1,1 % (1,1)	0,2 % (0,4)	-1,3 % (0,4)
<i>Tel2+</i>	-0,1 % (0,7)	-0,1 % (0,1)	-2,3 % (0,8)	-4,9 % (1,7)	-0,6 % (0,4)	-1,0 % (0,3)	-0,8 % (1,0)	-0,2 % (0,3)	-1,2 % (0,4)
<i>F2F3</i>	-0,5 % (0,3)	-0,1 % (0,1)	0,0 % (0,4)	0,7 % (0,6)	-0,1 % (0,1)	0,0 % (0,1)	0,5 % (0,3)	0,3 % (0,1)	0,1 % (0,1)
<i>F2F3+</i>	0,0 % (0,0)								
<i>Web → Tel2</i>	0,9 % (1,0)	0,0 % (0,4)	-2,4 % (1,5)	-3,4 % (3,7)	-0,1 % (0,6)	-0,7 % (0,5)	-4,4 % (1,9)	0,9 % (0,5)	-0,7 % (0,6)
<i>Web → Tel2+</i>	0,9 % (0,9)	-0,1 % (0,3)	-3,7 % (1,4)	-1,7 % (3,2)	0,5 % (0,7)	-0,7 % (0,4)	-3,0 % (1,4)	0,6 % (0,5)	-0,4 % (0,6)
<i>Web → F2F3</i>	0,7 % (0,6)	0,0 % (0,3)	-1,2 % (0,8)	-1,6 % (1,4)	0,6 % (0,5)	-0,3 % (0,3)	-1,0 % (0,8)	0,5 % (0,3)	-0,2 % (0,3)
<i>Web → F2F3+</i>	0,9 % (0,6)	0,0 % (0,3)	-1,2 % (0,8)	-2,0 % (1,4)	0,6 % (0,5)	-0,3 % (0,3)	-1,2 % (0,8)	0,4 % (0,3)	-0,2 % (0,3)

Tableau A4

Effets de méthode estimatifs par rapport à la valeur repère $BM_2 = 1/3 * (\bar{y}_{Web} + \bar{y}_{Tel2+} + \bar{y}_{F2F3+})$ avec les erreurs types correspondantes entre parenthèses

$D^{BM_2}(s, g)$	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9
<i>Web</i>	1,0 % (0,5)	0,1 % (0,3)	-0,8 % (0,9)	-1,4 % (1,8)	0,8 % (0,4)	0,1 % (0,2)	-1,2 % (0,8)	0,5 % (0,2)	0,1 % (0,3)
<i>Tel2</i>	-0,6 % (0,3)	-0,1 % (0,2)	-1,0 % (0,6)	-3,7 % (1,4)	-1,2 % (0,2)	-0,5 % (0,2)	-0,1 % (0,8)	0,1 % (0,2)	-0,8 % (0,2)
<i>Tel2+</i>	-0,6 % (0,2)	-0,1 % (0,2)	-0,8 % (0,5)	-1,7 % (1,0)	-0,7 % (0,2)	-0,5 % (0,1)	0,2 % (0,5)	-0,3 % (0,1)	-0,6 % (0,2)
<i>F2F3</i>	-1,0 % (0,7)	-0,1 % (0,2)	1,6 % (0,8)	3,8 % (1,6)	-0,2 % (0,4)	0,5 % (0,2)	1,5 % (0,8)	0,2 % (0,3)	0,6 % (0,3)
<i>F2F3+</i>	-0,5 % (0,5)	0,0 % (0,2)	1,6 % (0,7)	3,1 % (1,4)	-0,1 % (0,4)	0,5 % (0,2)	1,0 % (0,7)	-0,1 % (0,3)	0,5 % (0,3)
<i>Web → Tel2</i>	0,4 % (0,5)	0,0 % (0,3)	-0,9 % (1,0)	-0,3 % (2,9)	-0,2 % (0,4)	-0,2 % (0,3)	-3,4 % (1,5)	0,7 % (0,3)	-0,1 % (0,4)
<i>Web → Tel2+</i>	0,5 % (0,4)	0,0 % (0,2)	-2,1 % (0,8)	1,5 % (2,0)	0,4 % (0,4)	-0,2 % (0,2)	-2,0 % (0,8)	0,5 % (0,2)	0,1 % (0,3)
<i>Web → F2F3</i>	0,3 % (0,2)	0,0 % (0,1)	0,4 % (0,3)	1,5 % (0,6)	0,5 % (0,2)	0,2 % (0,1)	0,0 % (0,3)	0,4 % (0,1)	0,3 % (0,1)
<i>Web → F2F3+</i>	0,4 % (0,1)	0,0 % (0,1)	0,4 % (0,3)	1,1 % (0,5)	0,5 % (0,2)	0,2 % (0,1)	-0,2 % (0,3)	0,3 % (0,1)	0,3 % (0,1)

Tableau A5

Coûts unitaires relatifs estimatifs (en euros) par stratégie s et par groupe g avec les erreurs types correspondantes entre parenthèses

$c(s, g)$	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9
<i>Web</i>	0,03 (0,0)	0,04 (0,0)	0,04 (0,0)	0,03 (0,0)	0,04 (0,0)	0,03 (0,0)	0,03 (0,0)	0,03 (0,0)	0,03 (0,0)
<i>Tel2</i>	0,11 (0,1)	0,15 (0,1)	0,10 (0,1)	0,09 (0,1)	0,13 (0,1)	0,11 (0,1)	0,09 (0,1)	0,12 (0,0)	0,14 (0,1)
<i>Tel2+</i>	0,13 (0,1)	0,17 (0,1)	0,11 (0,1)	0,10 (0,1)	0,15 (0,1)	0,14 (0,1)	0,11 (0,1)	0,16 (0,1)	0,20 (0,2)
<i>F2F3</i>	0,84 (0,4)	0,89 (0,5)	0,83 (0,5)	0,82 (0,8)	0,86 (0,3)	0,84 (0,2)	0,81 (0,5)	0,84 (0,2)	0,89 (0,5)
<i>F2F3+</i>	1,00 (0,6)	1,00 (0,6)	1,00 (0,7)	1,00 (1,1)	1,00 (0,4)	1,00 (0,3)	1,00 (0,6)	1,00 (0,2)	1,00 (0,5)
<i>Web → Tel2</i>	0,08 (0,0)	0,11 (0,1)	0,09 (0,1)	0,09 (0,1)	0,09 (0,0)	0,08 (0,0)	0,08 (0,0)	0,07 (0,0)	0,07 (0,0)
<i>Web → Tel2+</i>	0,09 (0,1)	0,12 (0,1)	0,10 (0,1)	0,10 (0,1)	0,10 (0,1)	0,09 (0,0)	0,09 (0,1)	0,08 (0,0)	0,07 (0,0)
<i>Web → F2F3</i>	0,60 (0,3)	0,66 (0,7)	0,64 (0,6)	0,70 (0,8)	0,59 (0,4)	0,56 (0,3)	0,65 (0,5)	0,51 (0,2)	0,61 (0,4)
<i>Web → F2F3+</i>	0,71 (0,4)	0,71 (0,7)	0,80 (0,9)	0,84 (1,2)	0,73 (0,6)	0,68 (0,4)	0,81 (0,8)	0,62 (0,3)	0,71 (0,6)

Annexe B

Aperçu des résultats d'optimisation

Dans la section 4.5, nous illustrons l'approche adoptée pour résoudre le problème d'optimisation multimodal pour un ensemble de paramètres d'entrée. Les tableaux B1 et B2 donnent un aperçu des résultats d'optimisation.

Tableau B1

Aperçu des résultats d'optimisation – formulation de la programmation linéaire – minimisation des coûts

Taille de l'échantillon (S_{\max})	Valeur objective (min coûts)	Repère	Effet de méthode (\bar{D}^{BM})	Différence maximale dans les effets de mode (M)	Taux de réponse
9 500	123 748,50	BM ₁	0,16 %	2,06 %	48,0 %
		BM ₂	0,29 %	3,31 %	
11 000	88 408,95	BM ₁	0,05 %	5,97 %	39,9 %
		BM ₂	0,19 %	2,98 %	
12 500	82 270,72	BM ₁	0,08 %	5,97 %	36,9 %
		BM ₂	0,21 %	2,98 %	
15 000	74 350,44	BM ₁	0,12 %	5,97 %	29,4 %
		BM ₂	0,25 %	2,39 %	

Tableau B2

Aperçu des résultats d'optimisation – problème non linéaire – minimisation de l'effet de méthode moyen dans l'EPA

S_{\max}	B	BM	M	\bar{D}^{BM}	M	\bar{D}^{BM}	M	\bar{D}^{BM}
9 500	160 000	BM ₁	1 %	0,155 %	0,5 %	Infaisable	0,25 %	Infaisable
		BM ₂		0,170 %				
	170 000	BM ₁	1 %	0,131 %	0,5 %	Infaisable	0,25 %	Infaisable
		BM ₂		0,170 %				
	180 000	BM ₁	1 %	0,100 %	0,5 %	Infaisable	0,25 %	Infaisable
		BM ₂		0,170 %				
12 000	160 000	BM ₁	1 %	0,097 %	0,5 %	0,119 %	0,25 %	0,123 %
		BM ₂		0,046 %		0,046 %		0,046 %
	170 000	BM ₁	1 %	0,076 %	0,5 %	0,093 %	0,25 %	0,101 %
		BM ₂		0,036 %		0,036 %		0,036 %
	180 000	BM ₁	1 %	0,009 %	0,5 %	0,058 %	0,25 %	0,095 %
		BM ₂		0,014 %		0,014 %		0,014 %
15 000	160 000	BM ₁	1 %	0,051 %	0,5 %	0,094 %	0,25 %	0,112 %
		BM ₂		0,006 %		0,006 %		0,006 %
	170 000	BM ₁	1 %	0,020 %	0,5 %	0,080 %	0,25 %	0,097 %
		BM ₂		0,004 %		0,004 %		0,004 %
	180 000	BM ₁	1 %	0,005 %	0,5 %	0,058 %	0,25 %	0,095 %
		BM ₂		0,000 %		0,000 %		0,000 %

Bibliographie

- Birgin, E.G., et Martinez, J.M. (2008). Improving ultimate convergence of an augmented lagrangian method. *Optimization Methods and Software*, 23, 177-195.
- Blanke, K., et Luiten, A. (2012). ESSnet project on data collection for social survey using multi modes (dcss). Papier pour l'UNECE Conference of European Statistics, 31 oct. au 2 nov., Genève, Suisse.
- Buelens, B., et van den Brakel, J. (2014). Measurement error calibration in mixed-mode surveys. À paraître dans la *Sociological Methods and Research*.
- Calinescu, M., et Schouten, B. (2013a). Adaptive survey designs that account for nonresponse and measurement error. Document de discussion, Statistics Netherlands.
- Calinescu, M., et Schouten, B. (2013b). Adaptive survey designs to minimize mode effects a case study on the dutch labour force survey. Document de discussion, Statistics Netherlands.
- Conn, A.R., Gould, N.I.M. et Toint, P.L. (1991). A globally convergent augmented lagrangian algorithm for optimization with general constraints and simple bounds. *SIAM Journal on Numerical Analysis*, 28, 545-572.

- de Leeuw, E.D. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, 21, 233-255.
- Dillman, D., Phelps, G., Tortora, R., Swift, K., Kohrell, J., Berck, J. et Messes, B. (2009). Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (ivr) and the internet. *Social Science Research*, 38, 1-18.
- Fan, W., et Yan, Z. (2010). Factors affecting response rates of the web survey: a systematic review. *Computers in Human Behavior*, 26, 132-139.
- Heeringa, S., et Groves, R. (2006). Responsive design for household surveys: tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 169(3).
- Jäckle, A., Roberts, C. et Lynn, P. (2010). Assessing the effect of data collection mode on measurement. *Revue Internationale de Statistique*, 78, 3-20.
- Johnson, S.G. (2013). The nlopt nonlinear-optimization package. Disponible en ligne au <http://ab-initio.mit.edu/nlopt>.
- Klausch, T., Hox, J. et Schouten, B. (2013a). Assessing the mode-dependency of sample selectivity across the survey response process. Document de discussion, Statistics Netherlands.
- Klausch, T., Hox, J. et Schouten, B. (2013b). Measurement effects of survey mode on the equivalence of attitudinal rating scale questions. *Sociological Methods & Research*, 42(3), 227-263.
- Klausch, L., Hox, J. et Schouten, B. (2014). Evaluating bias in sequential mixed-mode surveys against single- and hybrid-mode benchmarks the case of the crime victimization survey. Document de discussion, Statistics Netherlands.
- Kreuter, F. (2013). *Improving Surveys with Process and Paradata*. John Wiley & Sons, Inc.
- Mohorko, A., de Leeuw, E.D. et Hox, J. (2013). Internet coverage and coverage bias in Europe: Developments across countries over time. *Journal of Official Statistics*, 29, 609-622.
- Powell, M.J.D. (1998). Direct search algorithms for optimization calculations. *Acta Numerica*, 7, 287-336.
- Roy, J.S. (2007). Stochastic optimization - scipy project. Disponible en ligne au <http://js2007.free.fr/>.
- Schouten, B., Calinescu, M. et Luiten, A. (2013a). Optimiser la qualité de la réponse au moyen de plans de collecte adaptatifs. *Techniques d'enquête*, 39, 1, 33-66.
- Schouten, B., van den Brakel, J., Buelens, B., van der Laan, J., Burger, J. et Klausch, T. (2013b). Disentangling mode-specific selection and measurement bias in social surveys. *Social Science Research*, 42, 1555-1570.
- Suzer-Gurtekin, Z. (2013). *Investigating the Bias Properties of Alternative Statistical Inference Methods in Mixed-Mode Surveys*. Thèse de doctorat, University of Michigan.
- Svanberg, K. (2002). A class of globally convergent optimization methods based on conservative convex separable approximations. *SIAM Journal on Optimization*, 12, 555-573.
- Vannieuwenhuyze, J. (2013). *Mixed-Mode Data Collection: Basic Concepts and Analysis of Mode Effects*. Thèse de doctorat, Katholieke Universiteit Leuven.
- Wagner, J. (2008). *Adaptive Survey Design to Reduce Nonresponse Bias*. Thèse de doctorat, University of Michigan.