Nº 12-001-X au catalogue
ISSN 1712-5685

Techniques d’enquête

Plan d’enquête adaptatif robuste pour les variations temporelles des propensions à répondre dans un contexte multimodal

par Shiya Wu, Harm-Jan Boonstra, Mirjam Moerbeek et Barry Schouten

Date de diffusion : le 20 décembre 2024

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 infostats@statcan.gc.ca

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

- Service de renseignements statistiques 1-800-263-1136
- Service national d'appareils de télécommunications pour les malentendants 1-800-363-7629
- Télécopieur 1-514-283-9350

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.

Publication autorisée par le ministre responsable de Statistique Canada

© Sa Majesté le Roi du chef du Canada, représenté par le ministre de l'Industrie, 2024

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.

Techniques d'enquête, décembre 2024
Vol. 50, nº 2, pp. 473-505
Statistique Canada, nº 12-001-X au catalogue

Plan d'enquête adaptatif robuste pour les variations temporelles des propensions à répondre dans un contexte multimodal

Shiya Wu, Harm-Jan Boonstra, Mirjam Moerbeek et Barry Schouten[1]

Résumé

Les plans d'enquête adaptatifs (PEA) permettent d'adapter les protocoles de recrutement aux sous-groupes de population qui présentent un intérêt pour une enquête. Ces dernières années, l'optimisation efficace d'un PEA a fait l'objet de recherches et de plusieurs applications. Toutefois, le rendement d'un PEA optimisé au fil du temps réagit aux variations temporelles des propensions à répondre. On ne comprend pas encore entièrement la façon dont les stratégies d'adaptation peuvent s'adapter à une telle variation au fil du temps. Dans la présente étude, nous proposons une approche d'optimisation robuste dans le contexte d'enquêtes séquentielles multimodales à l'aide d'une analyse bayésienne. Cette approche s'exprime sous la forme d'un problème de programmation mathématique qui tient explicitement compte de l'incertitude attribuable à la variation temporelle. Des décisions en matière de PEA peuvent alors être prises en tenant compte de la variation chronologique de la propension à répondre conditionnelle au mode et des corrélations de la propension à répondre entre les modes. La démonstration de cette approche fait appel à une étude de cas : l'Enquête sur la santé aux Pays-Bas de 2014 à 2017. Nous évaluons la sensibilité du rendement des PEA 1) au niveau budgétaire et 2) à la durée des données de série chronologique historiques applicables. Nous concluons que la dépendance au niveau budgétaire n'est que modérée et que la dépendance aux données historiques est tempérée par le degré de saisonnalité au cours de l'année.

Mots-clés : Analyse de séries chronologiques; approche bayésienne; modèle de propension à répondre; optimisation; plans d'enquête adaptatifs.

1. Introduction

Les plans d'enquête adaptatifs (PEA) (Wagner, 2008; Schouten, Peytchev et Wagner, 2017) sont progressivement devenus un choix viable dans les enquêtes contemporaines; on ne propose plus un seul protocole d'enquête pour toutes les personnes et tous les sous-groupes, mais on l'adapte afin d'obtenir des réponses des personnes de manière efficace, en fonction de caractéristiques de population connues et de caractéristiques observées au cours du travail sur le terrain. Cette transition a été accélérée par la baisse persistante des taux de réponse, les budgets limités, une plus grande variété de sources de données, l'émergence de toutes sortes d'appareils mobiles et la migration progressive vers des enquêtes multimodales. Cette évolution suppose une plus grande urgence et davantage d'options en ce qui concerne les plans d'enquête.

Un élément clé du PEA est la stratégie d'optimisation, c'est-à-dire l'ensemble de règles de décision. De telles stratégies reposent sur les intrants de propensions à répondre et d'autres paramètres du plan d'enquête. Parmi les principales approches en matière d'optimisation figurent la priorisation des cas (Peytchev, Riley, Rosen, Murphy et Lindblad, 2010; Wagner, 2013; Wagner et Hubbard, 2013), l'approche par essais et erreurs, ainsi que l'optimisation mathématique et statistique (van Berkel, van der Doef et Schouten, 2020;

1. Shiya Wu, Utrecht University, Department of Methodology and Statistics; Harm-Jan Boonstra, Statistics Netherlands, Department of Statistical Methods; Mirjam Moerbeek, Utrecht University, Department of Methodology and Statistics; Barry Schouten, Statistics Netherlands, Department of Statistical Methods and Utrecht University, Department of Methodology and Statistics. Courriel : jg.schouten@cbs.nl.

Calinescu, Bhulai et Schouten, 2013; Schouten, Calinescu et Luiten, 2013); voir Schouten et coll. pour connaître les avantages et les inconvénients de chaque approche. Cependant, dans le cadre de ces contributions, l'inexactitude des propensions à répondre estimées à partir des données historiques a été la plupart du temps ignorée. Dans la programmation mathématique, les objectifs peuvent être paramétrés sous forme de fonctions de propension à répondre agissant comme l'un des principaux intrants de l'optimisation. Une erreur serait introduite dans la prise de décisions en cas de variation des réelles propensions à répondre au fil du temps, lorsque cette variation n'est pas prise en compte dans l'estimation. Par conséquent, cette inexactitude rend tout PEA sous-optimal, ou, pire, inefficace. Placer l'optimisation du PEA dans un contexte bayésien est naturel pour répondre à cet enjeu; pourtant, les études pertinentes sur cette méthodologie d'enquête en sont toujours à leurs débuts. Récemment, Ma (2021) a élaboré une méthodologie pour optimiser efficacement une stratification en tenant compte de l'exactitude des estimations des propensions à répondre de façon bayésienne, en fonction des données historiques les plus récentes.

La variation temporelle des propensions à répondre et des estimations inexactes de données historiques compromet la robustesse de l'optimisation des PEA; voir Schouten et coll. (2017) ainsi que Chun, Heeringa et Schouten (2018) pour une analyse plus approfondie. Récemment, les travaux de recherche en matière de PEA ont commencé à mettre l'accent sur l'élaboration de modèles de propension à répondre et l'amélioration de l'exactitude des prédictions; Schouten, Mushkudiani, Shlomo, Durrant, Lundquist et Wagner (2018) ont été les premiers à élaborer des méthodes de mise à jour bayésiennes afin de contrer ce biais en utilisant statistiquement les données d'enquête accumulées et les données historiques générées à partir de mises en œuvre passées d'une même enquête. En étant les plus instructives, les croyances *a priori* découlant de données d'enquête antérieures peuvent améliorer les données actuelles à des fins de prédiction. Manifestement, traduire des sources de données externes en croyances *a priori* est nécessaire pour élaborer des modèles de propension à répondre. Pour ce faire, l'analyse documentaire (West, Wagner, Coffey et Elliott, 2023) et l'obtention de connaissances spécialisées (Coffey, West, Wagner et Elliott, 2020; Wu, Schouten, Meijers et Schouten, 2022) sont des approches récentes permettant d'obtenir des renseignements *a priori*. Les spécialistes de la recherche sur les enquêtes traitent le sujet de l'actualité des données historiques de façon incomplète et considèrent globalement les propensions à répondre à différentes étapes de l'enquête, alors que des faits, notamment des taux de réponse constamment en baisse au fil des ans, indiquent que l'exactitude des estimations des propensions à répondre dépend du temps et que les propensions à répondre à des enquêtes séquentielles sont probablement corrélées.

Les travaux les plus pertinents menés par Wu, Boonstra, Moerbeek et Schouten (2023) ont porté sur la décomposition de la variation temporelle des propensions à répondre à de multiples niveaux afin d'étudier l'influence de la durée des données d'enquête historiques applicables sur l'exactitude des prédictions de la propension à répondre. Dans ce cas précis, seule l'étape de collecte de données des interviews Web assistées par ordinateur (IWAO) dans le cadre de l'Enquête sur la santé aux Pays-Bas (GEZO) a été prise en compte et non l'étape d'interview sur place assistée par ordinateur (IPAO) applicable aux non-répondants aux IWAO. Dans la présente étude, nous généralisons l'élaboration du modèle de prédiction de la propension à

répondre aux multiples étapes de la collecte de données, en particulier au cas des étapes de collecte de données par IWAO et IPAO de l'enquête GEZO. Cela permet, par exemple, d'évaluer l'exactitude conditionnelle des prédictions des propensions à répondre par IPAO étant donné la réalisation de la réponse par IWAO sur une certaine période. Comme dans Wu et coll. (2023), nous adoptons une approche bayésienne permettant une quantification complète de l'incertitude des propensions à répondre et des quantités dérivées. La deuxième extension de la présente étude par rapport à Wu et coll. (2023) est l'analyse du rendement des PEA selon diverses contraintes externes, en tenant compte de l'incertitude des prédictions des propensions à répondre.

Dans l'ensemble, la présente étude vise deux contributions aux modèles séquentiels multimodaux (MM) : prédire aussi exactement que possible la propension à répondre de chaque mode d'enquête et prendre des décisions adaptatives de manière aussi optimale que possible. Pour satisfaire cette ambition en utilisant des données de séries chronologiques historiques dans l'évaluation, nous posons trois questions de recherche :

- Comment construire des modèles de série chronologique pour améliorer l'exactitude de la prédiction de la propension à répondre dans un plan séquentiel multimodal ?
- Dans quelle mesure le rendement des PEA est-il sensible au niveau budgétaire donné ?
- Dans quelle mesure le rendement des PEA dépend-il de la durée des données historiques applicables ?

En réponse à la première question, nous étendons les modèles hiérarchiques binomiaux de série chronologique pour un seul mode proposés par Wu et coll. (2023) à des modèles de série chronologique hiérarchiques multinomiaux pour de multiples modes, et nous illustrons cela au moyen d'une application à l'enquête GEZO, pour laquelle nous utilisons des données relatives à la période de 2014 à 2017. Cette extension tient également compte de l'intégration de paramètres de corrélation entre les modes à la modélisation des propensions à répondre des modes d'IWAO et d'IPAO.

Pour ce qui est des deuxième et troisième questions, une stratégie tenant compte de l'incertitude quant aux intrants de l'optimisation lors de l'optimisation de répartitions probabilistes est très recherchée. Le rendement du plan d'enquête est suivi en évaluant la représentation de caractéristiques contextuelles pertinentes couplées au moyen de données administratives. Cette représentation est opérationnalisée par le coefficient de variation (c.v.) des propensions à répondre (Schouten, Cobben et Bethlehem, 2009). Nous étalonnons le rendement des PEA par rapport au rendement des plans par IWAO uniquement et des plans non adaptatifs pour veiller à ce que les répartitions déterminées améliorent le rendement des PEA. Pour déterminer la sensibilité du rendement des PEA, nous menons deux expériences. Dans la première expérience, nous sélectionnons un trimestre de l'année et diminuons progressivement le niveau budgétaire. Cette expérience nous permet d'étudier la sensibilité du rendement des PEA aux contraintes budgétaires. Dans la deuxième expérience, le niveau budgétaire est fixe et la fenêtre de la série chronologique des données historiques se déplace vers le nouveau trimestre suivant de collecte de données. Des données

historiques sont d'abord utilisées pour établir une probabilité *a priori*. Ensuite, au fil des nouveaux trimestres, la probabilité *a priori* est mise à jour de manière répétée en une probabilité *a posteriori* servant de probabilité *a priori* pour le trimestre suivant. Nous évaluons la façon dont la probabilité *a priori* varie ainsi que les effets que cela a sur les répartitions des unités d'échantillonnage.

La présente étude s'articule comme suit : nous commençons par construire le modèle de série chronologique pour les plans séquentiels multimodaux dans la deuxième section. Dans la troisième section, nous décrivons le problème d'optimisation. Nous introduisons l'étude de cas dans la quatrième section et répondons aux questions de recherche. Dans la dernière section, nous traitons des avantages et des inconvénients de notre méthode et concluons avec des idées de travaux de recherche futurs.

2. Méthodes

Dans la présente section, un modèle de série chronologique multivarié est élaboré pour les propensions à répondre de plans séquentiels multimodaux. Nous élargissons le modèle hiérarchique de série chronologique suggéré par Wu et coll. (2023) en introduisant des propensions conditionnelles à répondre de modes de suivi.

2.1 Modélisation des propensions à répondre dans des plans d'enquête séquentiels multimodaux

Le modèle de série chronologique de Wu et coll. (2023) produit des estimations précises des propensions à répondre pour des plans d'enquête à un seul mode, ou pour le premier mode des enquêtes multimodales au cours du travail sur le terrain. Ici, l'objectif évolue pour permettre des prédictions fiables pour chaque mode d'enquêtes multimodales afin d'étendre l'attrait du modèle. On considère notamment des données de série chronologique à valeurs discrètes, y compris la taille d'un échantillon et le nombre de répondants pour chaque mode, d'une répartition multinomiale, alors que Wu et coll. (2023) considéraient une répartition binomiale des données.

La propension à répondre (PR) est la propension théorique d'un sujet échantillonné à répondre pour un mode d'interview particulier selon un ensemble de caractéristiques connues. Ces caractéristiques peuvent comprendre des paradonnées recueillies à une étape particulière ou un mode particulier de l'enquête. Ce sujet peut être soit une personne, soit un groupe de personnes bien défini. Il convient de mentionner qu'un groupe est établi en recoupant plusieurs variables auxiliaires considérées comme étant des prédicteurs puissants des variables d'enquête. Au sein de ce groupe, les unités présentent des attributs démographiques homogènes, comme l'âge. Un tel ensemble de groupes peut varier au fil du temps ou des variations de conception (voir Schouten et coll., 2017 pour en savoir plus sur la stratification), mais nous supposons ici que la stratification est constante et précisée avant l'ajustement du modèle.

Pour modéliser la propension à répondre au niveau du mode, nous supprimons, dans la présente section, les indices indiquant un groupe particulier dans le paramètre de propension ainsi qu'un moment particulier; toutefois, dans la section suivante, nous devons préciser cet indice pour décomposer une série chronologique en effets fixes ou aléatoires au niveau de la strate, du temps et/ou du mode.

Supposons qu'une enquête multimodale comprend $M-1$ modes de collecte de données. Nous ajoutons un M^{e} « mode » correspondant à la non-réponse, c'est-à-dire la catégorie représentant une non-réponse aux $M-1$ modes. Soit un échantillon aléatoire de taille n connu avant le début de la collecte de données et r_j, le nombre observé de répondants pour le j^{e} mode, où $j \in \{1, \ldots, M\}$. Supposons une répartition multinomiale en M modes avec une propension à répondre ρ_j pour le j^{e} mode, où $\rho_j \in [0,1]$. ρ_M est la propension de non-réponse; toutefois, celle-ci n'est plus explicitement modélisée plus tard.

Le vecteur $\mathbf{r} = (r_1, \ldots, r_M)$ suit une répartition multinomiale dont la taille d'échantillon est n et la propension à répondre est $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_M)$, c'est-à-dire la répartition conjointe de $\mathbf{r}$ est une généralisation multivariée d'une répartition binomiale,

$$\text{mult}(\mathbf{r} \mid n, \boldsymbol{\rho}) = \frac{n!}{\prod_{j=1}^{M} r_j!} \prod_{j=1}^{M} \rho_j^{r_j}. \tag{2.1}$$

Linderman, Johnson et Adams (2015) ont eu recours à une transformation par rupture de bâton pour reformuler la répartition multinomiale en un produit de répartitions binomiales, au sein de laquelle les paramètres créés sont dépendants. Cela a donné l'occasion de reformuler l'équation (2.1) à m dimensions de façon récursive en termes de $M-1$ binômes. Dans la représentation par rupture de bâton, le vecteur de propension $\boldsymbol{\rho}$ sert de bâton divisé de manière récursive en deux morceaux pour créer une variable binomiale $\tilde{\boldsymbol{\rho}} = (\tilde{\rho}_1, \ldots, \tilde{\rho}_{M-1})$. Pour fournir une dérivation, posons que la variable r_j de réponse pour le j^{e} mode suit une densité binomiale avec les paramètres n_j et $\tilde{\rho}_j$, c'est-à-dire $\text{bin}(r_j \mid n_j, \tilde{\rho}_j)$, où n_j et $\tilde{\rho}_j$ représentent la taille restante de l'échantillon et la fraction de la probabilité restante approchée par le mode j^{e},

$$n_j = n - \sum_{k<j} r_k, \tag{2.2}$$

$$\tilde{\rho}_j = \frac{\rho_j}{1 - \sum_{k<j} \rho_k}, \tag{2.3}$$

où $j \in \{2, \ldots, M\}$. Lorsque $j = 1$, le paramètre $n = n_1 = \sum_{j \in \{1, \ldots, M\}} r_j$ et le paramètre $\tilde{\rho}_j = \rho_1$. À l'aide de (2.3), nous obtenons

$$\tilde{\rho}_j^{r_j} (1 - \tilde{\rho}_j)^{n_j - r_j} = \rho_j^{r_j} \frac{1}{\left(1 - \sum_{k<j} \rho_k\right)^{n_j}} \left(1 - \sum_{k \le j} \rho_k\right)^{n_j - r_j}. \tag{2.4}$$

Notons que la somme de r_j est n sur j et $n_j = n_{j-1} - r_{j-1}$ pour tout $j \in \{2, \ldots, M\}$. Cela signifie que les paires de termes $\left(1 - \sum_{k<j} \rho_k\right)^{n_j}$ s'annuleront dans le produit de (2.4) sur j, menant au même format que le terme exponentiel multinomial dans (2.1),

$$\prod_{j=1}^{M-1} \tilde{\rho}_j^{r_j} \, (1-\tilde{\rho}_j)^{n_j - r_j} = \prod_{j=1}^{M} \rho_j^{r_j}. \tag{2.5}$$

Les constantes de normalisation suivent le même raisonnement. Combinée aux termes exponentiels dans (2.5), l'équation (2.1) peut être réécrite sous la forme

$$\text{mult}(\mathbf{r} \mid n, \boldsymbol{\rho}) = \prod_{j=1}^{M-1} \text{bin}(r_j \mid n_j, \tilde{\rho}_j). \tag{2.6}$$

Nous utilisons la représentation par rupture de bâton du modèle multinomial à des fins pratiques : pour cette représentation, il existe un échantillonneur de Gibbs simple et efficace pour le modèle (hiérarchique) multinomial. Comme l'expliquent Linderman et coll. (2015), ils emploient la même méthodologie d'augmentation de données Polya-Gamma (Polson, Scott et Windle, 2013) utilisée pour les modèles binomiaux dans Wu et coll. (2023). La représentation par rupture de bâton permet d'échantillonner les coefficients du modèle pour toutes les $M-1$ catégories modélisées dans un bloc, améliorant ainsi la convergence de l'échantillonneur de Gibbs. Cette représentation présente également un inconvénient : la définition de $\tilde{\rho}_j$ rend l'interprétation des coefficients de corrélation sous-jacents plus difficile, en particulier l'interprétation des coefficients de corrélation dans les modèles décrits en détail dans la section 2.2.

Dans la section 2.2, un modèle de série chronologique structurel est employé pour décomposer une série chronologique observée en des composantes temporelles sous-jacentes.

2.2 Modèle hiérarchique multinomial de série chronologique

Pour mesurer la dépendance des propensions à répondre parmi les modes, nous élargissons les modèles de Wu et coll. (2023) en introduisant un nouveau paramètre hiérarchique qui désigne les coefficients de corrélation. Une telle dépendance répartie sur les composantes d'intérêt de la série chronologique est similaire à ce qu'ont adopté Wu et coll. (2023); il est suggéré de consulter cette étude pour obtenir de plus amples précisions sur chaque définition de composante ainsi que des renseignements techniques.

Pour décrire chaque composante du modèle au niveau le plus détaillé, posons que le vecteur de paramètre de propension dépendante des modes séquentiels est associé à une strate de PEA particulière et à un moment particulier, c'est-à-dire $\tilde{\boldsymbol{\rho}}_{g,t} = \{\tilde{\rho}_{g,t,j} \mid j \in \{1, \ldots, M-1\}\}$, où la j^{e} entrée désigne le paramètre de propension du j^{e} mode dans la strate de PEA g au moment t, comme le définit (2.3). Les nombres de strates de PEA, de moments et de modes d'enquête sont respectivement G, T et $M-1$. Il convient de mentionner que dans la présente étude, nous utilisons le terme « strate de PEA » pour désigner un groupe de population pouvant recevoir un traitement différent. Cela ne doit pas être confondu avec des strates d'échantillon recevant des probabilités d'inclusion différentes. Nous omettons la référence au PEA dans les strates de PEA dans la majeure partie de ce qui suit.

Dans la présente étude, nous supposons que le choix des strates de PEA a déjà été effectué et que la sélection de variables elle-même ne fait pas partie de la stratégie de modélisation. Les strates seront fondées

sur des renseignements auxiliaires prédictifs des variables d'intérêt de l'enquête ou de la non-réponse à l'enquête. Toutefois, en cas de volume important de renseignements auxiliaires, une stratégie de sélection de variables devrait être intégrée à l'ajustement du modèle et à la stratégie d'optimisation du PEA. Nous revisitons cet enjeu dans la section sur l'analyse.

Soit $\theta_{g,t,j} = \text{logit}(\tilde{\rho}_{g,t,j})$. Une fonction de lien logit est utilisée pour convertir l'échelle limitée d'une probabilité à l'échelle non limitée d'un prédicteur linéaire, c'est-à-dire une combinaison linéaire de composantes de modèle. Soit $\theta_{g,t,j} = \text{logit}(\tilde{\rho}_{g,t,j})$, le prédicteur linéaire. Par conséquent, la fonction de probabilité multinomiale dans l'équation (2.1) peut être reformulée en substituant la transformation inverse à $\tilde{\rho}_{g,t,j}$,

$$\text{mult}(\mathbf{r} \mid n, \boldsymbol{\rho}) \propto \prod_{j=1}^{M-1} \left(\frac{e^{\theta_{g,t,j}}}{1+e^{\theta_{g,t,j}}} \right)^{r_j}. \tag{2.7}$$

Les modèles hiérarchiques considérés pour modéliser le prédicteur linéaire $\theta_{g,t,j}$ prennent la forme générale d'une décomposition additive, qui fait référence à une fonction de la somme des composantes de la série chronologique. Donc,

$$\theta_{g,t,j} = \beta_j + \beta'_{xj} x_g + \delta'_s s_t + u_{t,j} + v_{g,j} + z_{g,t,j} + e_{g,t,j}, \tag{2.8}$$

dont les trois premiers et les quatre derniers termes sont respectivement modélisés en tant qu'effets fixes et effets aléatoires.

Les premiers effets fixes de régression β_j sont des ordonnées à l'origine propres au mode, mesurant l'effet principal sur $\theta_{g,t,j}$. Les deuxièmes effets fixes β_{xj} sont des coefficients de régression propres au mode associés à la covariable x_g de vecteur p, correspondant aux caractéristiques démographiques particulières associées à la strate g. Les troisièmes effets fixes δ_s sont des coefficients de régression propres à la saison associés à la variable d'indicateur de saison s_t de vecteur q. Voir la définition des strates et des saisons à l'annexe A. Actuellement, les effets de la saison et du mode sont communs à toutes les strates. Plus globalement, ces effets fixes peuvent être propres à la strate. Dans l'application présente tout au long de cette étude, x_g et s_t sont des vecteurs binaires correspondant à des variables catégoriques, mais des variables ordinales ou numériques, voire variant dans le temps, peuvent également être prises en compte, le cas échéant.

Chaque terme d'effet aléatoire dans l'équation (2.8) permet implicitement la corrélation entre les modes d'enquête. Veuillez consulter l'étude de Wu et coll. (2023) pour obtenir une description des composantes d'effet aléatoire. Comme nous l'avons souligné, ces termes sont désormais croisés avec le mode, c'est-à-dire que des paramètres de variance distincts pour chaque mode et des paramètres de corrélation entre les modes sont introduits. La tendance temporelle globale $\mathbf{u}$, l'ordonnée à l'origine aléatoire pour la strate $\mathbf{v}_g$ et la tendance propre à la strate $\mathbf{z}_g$ respectent cette règle. Des effets aléatoires de bruit blanc $e_{g,t,j}$ sont également croisés avec le mode, mais nous utilisons un seul paramètre de variance commun pour tous les modes et aucune corrélation.

Nous adoptons une approche bayésienne pour estimer le modèle dans l'équation (2.8) et obtenir des prédictions fiables des propensions à répondre au niveau du mode, de la strate et du temps. Comme nous l'avons mentionné, les probabilités *a priori* sont les mêmes pour les coefficients correspondant aux différents modes. Pour simplifier la notation, nous supprimons les indices g, t et j dans chaque terme de composante de modèle. Les effets fixes β et δ se voient attribuer des probabilités *a priori* faiblement informatives à répartition normale de moyenne nulle et de matrice de variance diagonale, où l'erreur-type prend une valeur relativement grande de 10. On suppose que chaque vecteur d'effets aléatoires respecte la probabilité *a priori* normale multivariée de moyenne $\mathbf{0}$ et de matrice de covariance définie comme le produit de Kronecker de deux matrices de covariance $\mathbf{A}$ et $\mathbf{V}$. Ici, $\mathbf{V}$ est une matrice de covariance entièrement paramétrée, à laquelle est attribuée une loi *a priori* de Wishart inversée et mise à l'échelle (Gelman et Hill, 2007; O'Malley et Zaslavsky, 2008) et $\mathbf{A}$ est une matrice fixe, qui peut être une matrice diagonale simple pour des effets non structurés ou une matrice structurée correspondant à des marches aléatoires au fil du temps. De plus amples renseignements techniques relatifs à la spécification de la probabilité *a priori* et de la stratégie d'estimation, notamment l'échantillonneur de Gibbs et les répartitions conditionnelles complètes nécessaires, figurent dans les études de Boonstra et van den Brakel (2019) et de Wu et coll. (2023).

Un modèle plus parcimonieux peut être obtenu en omettant l'interaction axée sur le mode et en remplaçant la matrice de covariance entièrement paramétrée $\mathbf{V}$ par une matrice diagonale, si les effets entre les modes sur les prédictions de propension présentent peu d'intérêt. On appelle ce modèle le modèle sans corrélation; il fournit un modèle de base par rapport auquel on peut évaluer le rendement du modèle de corrélation complet.

L'estimation au moyen du modèle est effectuée à l'aide du paquet R *mcmcsae* (Boonstra, 2022). La convergence des résultats de la simulation de la méthode de Monte Carlo par chaîne de Markov (MCMC) est traitée en utilisant le tracé de trajectoire et la réduction scalaire potentielle ou diagnostic R-hat (Gelman et Rubin, 1992). Ces diagnostics montrent que l'échantillonneur de Gibbs converge rapidement, à la fois pour les modèles incluant et excluant des corrélations entre les modes. Dans une large mesure, cela est attribuable à l'échantillonnage de tous les effets fixes et aléatoires dans un seul bloc, ce qui est possible en raison de la représentation par rupture de bâton de la répartition multinomiale en combinaison avec l'augmentation de données Polya-Gamma, comme nous l'avons mentionné précédemment.

2.3 Extensions à des plans d'enquête complexes plus généraux

Le modèle présenté à la section 2.2 prend en charge des plans d'enquête d'échantillonnage aléatoire stratifié, mais pas d'autres caractéristiques d'échantillonnage complexe. Cela est justifiable pour l'application de l'enquête GEZO sur laquelle nous nous concentrons. Même si l'enquête GEZO repose sur un plan de sondage à deux degrés (les municipalités étant les unités d'échantillonnage de premier degré), les effets de catégorisation des résultats obtenus sont très mineurs, car la plupart des municipalités sont sélectionnées.

De plus, les probabilités d'inclusion aux premier et deuxième degrés sont telles que les probabilités d'inclusion générale sont égales pour toutes les personnes. La stratification utilisée dans la présente étude est choisie de sorte que les probabilités de réponse soient raisonnablement homogènes au sein d'une strate.

Nous allons maintenant décrire brièvement dans quelle mesure notre méthode peut être étendue pour prendre en charge les enquêtes à plans d'échantillonnage complexe plus généraux. Si les probabilités d'échantillonnage sont inégales, l'une peut, dans certains cas, tout de même définir une stratification de telle sorte que les probabilités soient (approximativement) égales au sein d'une strate. Si cela n'est pas possible, une analyse fondée sur un modèle au niveau de l'unité serait plus appropriée. Le modèle multinomial au niveau de la strate devient alors un modèle catégorique ou Multinoulli au niveau de la personne, c'est-à-dire le cas spécial de la répartition multinomiale avec $n_i = 1$ pour chaque personne i. Il convient de souligner que les dérivations relatives à la représentation par rupture de bâton de la section 2.2 sont toujours valides dans ce cas. Pour l'essentiel, les indices de strate g dans l'équation (2.8) deviendraient des indices de personne i.

Des covariables au niveau de l'unité expliquant la variation à la fois de l'échantillonnage et des probabilités de réponse devraient être incluses dans le modèle pour atténuer le biais autant que possible. En particulier, les probabilités d'inclusion elles-mêmes ou les variables sous-jacentes dont elles dépendent sont des covariables précieuses. Il peut également être utile de modéliser la dépendance envers les probabilités d'inclusion de manière souple; voir par exemple Chen, Elliott et Little (2010). Une stratification peut toujours être définie expressément aux fins du PEA, de sorte que les probabilités de réponse soient raisonnablement homogènes au sein d'une strate. Il n'est pas nécessaire qu'une telle stratification coïncide avec une stratification possible utilisée pour l'échantillonnage. Cela peut être géré par un modèle au niveau de l'unité en incluant des indicateurs de strate dans le modèle. De la même manière, pour un plan par grappe, le modèle au niveau de l'unité peut tenir compte des effets de grappe en incluant des indicateurs de grappe. Toutefois, puisque généralement seul un sous-échantillon de grappes est inclus dans l'échantillon et que le nombre d'observations par grappe observée peut être petit, les coefficients de grappe correspondants devraient être modélisés en tant qu'effets aléatoires (Scott et Smith, 1969). Enfin, nous soulignons que les extensions de modèle décrites pour prendre en charge des plans d'échantillonnage plus complexes peuvent être traitées à l'aide du même cadre d'échantillonneur de Gibbs employé dans l'application de l'enquête GEZO. En particulier, la même approche d'augmentation des données peut être utilisée pour la famille catégorique ou Multinoulli, puisqu'il s'agit d'un cas particulier de la répartition multinomiale. La seule différence est que l'on peut devoir intégrer d'autres covariables et leurs effets fixes correspondants, ainsi que des effets aléatoires supplémentaires pour les grappes, possiblement à diverses étapes. De tels modèles au niveau de l'unité peuvent être ajustés de la même façon, par exemple à l'aide du paquet R *mcmcsae*, même si les temps de calcul augmenteront du fait de la taille supérieure des données au niveau de l'unité et de la taille du modèle.

3. Optimisation de la répartition de modes dans le modèle de série chronologique hiérarchique bayésien

La présente section vise à étudier un problème de répartition rendant compte de l'incertitude en matière d'actualité et de mise en œuvre de PEA. La section 3.1 souligne les principaux éléments de construction et de fonctionnement de ce problème dans un cadre bayésien. Une stratégie est proposée à la section 3.2 pour évaluer l'intérêt de répartitions adaptatives par rapport à des répartitions non adaptatives en matière de réduction du risque de biais de non-réponse en suivant une mesure du risque de biais.

3.1 Principaux éléments

En général, une optimisation mathématique fait intervenir la sélection des « meilleures valeurs disponibles » d'une fonction objectif quelconque par rapport à un nombre de contraintes en choisissant des valeurs d'intrants parmi un ensemble autorisé. Établir des modèles d'optimisation fait intervenir trois éléments principaux : des variables de décision visant à optimiser l'objectif, des objectifs à minimiser ou à maximiser et des contraintes sur les variables de décision. Du fait de l'optimisation relativement au paramètre bayésien, nous insistons sur le fait que tous les paramètres statistiques mentionnés sont considérés comme des variables aléatoires ayant des valeurs variant dans le temps. Par conséquent, des fonctions objectifs et des fonctions de contraintes sont également des variables aléatoires. Dans ce qui suit, les principaux éléments sont d'abord présentés pour un cadre non bayésien et sont ensuite développés pour le cadre bayésien.

Les *variables de décision* sont des représentations symboliques d'une intervention décidée par le décideur. Elles représentent des parties inconnues d'une fonction objectif pouvant être manipulées et peuvent prendre toute valeur possible au sein d'un ensemble autorisé s'il est précisé. Dans la présente étude, une intervention est censée attribuer des modes d'interview à des strates lorsque les modes précédents ne permettent pas d'obtenir les données. Les variables de décision font par conséquent référence à des probabilités de répartition indiquant dans quelles mesures les non-répondants sont susceptibles d'être approchés par un mode de suivi. La probabilité de répartition $s_{g,t} \in [0,1]$ prend une décision sur la taille des candidats au suivi dans la strate g au moment t, où $s_{g,t} = 0$ signifie que la collecte de données pour la strate g cesse, et $s_{g,t} = 1$ signifie que tous les non-répondants de la strate g des modes précédents sont affectés au mode suivant.

La *fonction objectif* définit le critère permettant d'évaluer les valeurs des variables de décision candidates. Mis à part les variables de décision, elle dépend uniquement de quantités observées et estimées. Notre objectif d'optimisation est de minimiser le risque attendu de biais de non-réponse grâce à une répartition optimale. Puisque la non-réponse ne peut pas être observée directement, la présente étude tient compte d'un indicateur substitut du biais de non-réponse qui est une fonction des propensions à répondre. Nous employons le coefficient de variation (c.v.) des propensions à répondre (Schouten et coll., 2009). Le c.v. réel de la population limite le biais normalisé absolu des moyennes des répondants. Toutefois, le c.v.

estimé sur un ensemble particulier de variables auxiliaires respectera uniquement une partie de ce biais global. Même si d'autres indicateurs existent (voir Moore, Durrant et Smith (2018) ainsi que Nishimura, Wagner et Elliott (2016)), ce sont principalement les renseignements auxiliaires disponibles qui jouent un rôle décisif. Dans des enquêtes polyvalentes, la disponibilité peut ne pas être un réel enjeu, puisque l'on peut se concentrer sur une représentation générale et que toute amélioration sera utile. Dans des enquêtes ne comprenant que quelques statistiques clés, la disponibilité de variables auxiliaires pertinentes est essentielle.

Le c.v. est l'écart-type pondéré divisé par le taux de réponse pondéré :

$$\text{c.v.}(s,t) \;=\; \frac{\sqrt{\sum_g d_{g,t}\left(\rho_{g,t}-\overline{\rho}_t\right)^2}}{\overline{\rho}_t}. \tag{3.1}$$

Le poids $d_{g,t}$ est la proportion de l'échantillon de la taille de strate g au moment t par rapport à la taille globale au moment t, c'est-à-dire $d_{g,t} = n_{g,t} \Big/ \sum_g n_{g,t}$. Cette notation suppose implicitement que le plan de sondage donne lieu à des poids d'inclusion égaux; dans le cas contraire, les poids de sondage devraient alors être également intégrés. Cet ajout est simple, mais rend la notation difficile. La propension à répondre mixte $\rho_{g,t}$ désigne la propension générale pour les modes, qui est la somme de la propension marginale à répondre du mode de départ et des propensions conjointes à répondre du mode $j \geq 2$, en supposant que la strate g n'a pas répondu aux derniers $j-1$ modes. (Ici, nous supposons implicitement que tous les non-répondants d'un mode sont admissibles au suivi. En pratique, certains types de non-réponse, comme celle attribuable à une maladie physique ou mentale, peuvent ne pas être admissibles.)

$$\rho_{g,t} = \rho_{g,t,1} + \sum_{j\in\{2,\ldots,m-1\}} \rho_{g,t,j} \prod_{i\leq j-1} (1-\rho_{g,t,k}). \tag{3.2}$$

Les propensions individuelles (conditionnelles) $\rho_{g,t,j}$ pour un mode j sont estimées par des modèles multinomiaux dans la section 2. L'équation (3.2) suppose que tous les non-répondants aux modes précédents seront recrutés par le mode j pour une enquête non adaptative; cependant, cela peut être modifié en une enquête adaptative en réduisant la propension conjointe à la variable de décision $s_{g,t,j} \in [0,1]$; l'équation mise à jour devient donc

$$\rho_{g,t} = \rho_{g,t,1} + \sum_{j\in\{2,..,m-1\}} s_{g,t,j}\, \rho_{g,t,j} \prod_{i\leq j-1} (1-\rho_{g,t,k}). \tag{3.3}$$

Il est clair que l'équation (3.3) est équivalente à l'équation (3.2) lorsque tous les $s_{g,t,j} = 1$.

Les dénominateurs de l'équation (3.1), appelés les taux de réponse pondérés sur les strates, indiquent le niveau estimé de propensions inconnues, définies comme la somme pondérée des propensions mixtes de l'équation (3.2) et de l'équation (3.3). Nous disons que le c.v. est non adaptatif lorsque tous les $s_{g,t,j} = 1$ et adaptatif, lorsqu'au moins un $s_{g,t,j} \neq 1$.

Nous utilisons un seul indicateur motivé d'enquêtes polyvalentes, c'est-à-dire présentant un éventail important et diversifié de variables d'enquête cibles. Toutefois, dans les enquêtes comptant une ou quelques

variables d'enquête cibles, des indicateurs plus ciblés peuvent être utilisés. Un bon exemple est l'indicateur $H1$ de Särndal et Lundström (2010). Faire cela modifierait l'utilisation de données d'enquête historiques et inclurait également des associations avec la ou les variables d'enquête cibles.

Les *contraintes* sont des inégalités ou équations fonctionnelles représentant des restrictions logiques quant aux variables de décision autorisées. Des contraintes peuvent garantir, par exemple, une recherche approfondie de solutions possibles à partir d'un espace de solutions finies. Dans le contexte des plans d'enquête, une contrainte peut être une limite imposée sur la qualité de l'enquête, comme des solutions entraînant un taux de réponse global supérieur à 0,5, ou sur le coût de l'enquête, comme le coût global des intervieweurs s'adressant aux non-répondants demeurant inférieur à un montant donné. Dans la présente étude, nous nous concentrons sur des contraintes de coût quant à la charge de travail relative à la prise de contact avec les candidats non répondants au moyen d'un mode de suivi. Dans les ententes sur les niveaux de service avec des commanditaires d'enquête, une probabilité maximale de dépassement budgétaire est souvent spécifiée, disons de 10 %. α désigne cette proportion de dépassement de budget. Si le budget est strictement limité, alors $\alpha = 0.$ Soit h le niveau budgétaire précisé. La contrainte de coût est désormais :

$$p\left(\sum_g s_{g,t}\,(n_{g,t} - r_{g,t}) \geq \sum_g h(n_{g,t} - r_{g,t})\right) \leq \alpha, \tag{3.4}$$

où p est la probabilité que la charge de travail adaptative dépasse la charge de travail non adaptative. Lorsque les valeurs de $s_{g,t}$ respectent la contrainte (3.4), la solution correspondante de la variable de décision est dite acceptable; sinon, la solution contredit la règle. Il est naturel de préciser des limites inférieure et supérieure de la variable de décision $\mathbf{s} = \{s_{g,t} \mid \forall g, t\},$ appelées « contraintes de boîtes »,

$$0 \leq \mathbf{s} \leq 1. \tag{3.5}$$

Ainsi, le problème d'optimisation dans une application de répartition d'échantillon dans le cadre non bayésien est formulé afin de détecter un vecteur $\mathbf{s}$ minimisant l'objectif (3.1) sujet aux contraintes (3.4) et (3.5) en fonction des paramètres $(n, r).$ Comme nous l'avons énoncé auparavant, (3.1) et la charge de travail figurant dans (3.4) sont des variables aléatoires dans le cadre de l'approche bayésienne, de sorte que nous tirons des espérances des répartitions *a posteriori*.

Puisqu'aucune expression explicite n'existe pour les répartitions *a posteriori*, nous les estimons empiriquement. Nous utilisons les répliques de l'échantillonneur de Gibbs des répartitions *a posteriori* et calculons, par réplique, le c.v. et les charges de travail. Nous obtenons :

$$\hat{E}\left(\text{c.v.}(s,t)\right) = \frac{1}{K}\sum_k \frac{\sqrt{\sum_g d_{g,t}\left(\rho_{g,t}^{(k)} - \bar{\rho}_t^{(k)}\right)^2}}{\bar{\rho}_t^{(k)}}, \tag{3.6}$$

où $\hat{E}\left(\text{c.v.}(s,t)\right)$ est l'espérance estimée *a posteriori* au moment t, $\rho_{g,t}^{(k)}$ est la k^{e} estimation itérée de la fonction prédictive *a posteriori* de $\rho_{g,t},$ et l'indice k désigne les passages sur les tirages MCMC. La probabilité d'un dépassement de budget est estimée empiriquement par la fréquence à laquelle le budget nécessaire pour assurer le plan adaptatif dépasse le budget établi :

$$\frac{\sum_k \mathbf{1}_{\sum_g s_{g,t}\left(n_{g,t}-r_{g,t}^{(k)}\right) \geq \sum_g h\left(n_{g,t}-r_{g,t}^{(k)}\right)}}{K} \leq \alpha. \tag{3.7}$$

$\mathbf{1}_{\sum_g s_{g,t}(n_{g,t}-r_{g,t}^{(k)}) \geq \sum_g h(n_{g,t}-r_{g,t}^{(k)})}$ est une fonction indicatrice qui prend la valeur de un lorsque l'inégalité dans son indice est satisfaite pour la k^{e} itération et de zéro autrement. Ainsi, l'optimisation bayésienne vise à réduire au minimum l'objectif de l'équation (3.6) en fonction des contraintes des équations (3.5) et (3.7).

Étalonnage : dans le problème d'optimisation bayésien, nous établissons un étalonnage afin d'évaluer le rendement de PEA de deux points de vue : améliorer la qualité et réaliser des économies. En particulier, promouvoir la représentativité de l'échantillon par recrutement peut améliorer la qualité de la collecte de données, alors que répartir les ressources coûteuses là où elles sont les plus nécessaires peut permettre de réaliser des économies. Cet objectif peut être atteint en passant, par exemple, d'un mode simple à des modes mixtes, mais optimalement répartis de nouveau ou en passant de modes mixtes complets à des modes mixtes partiels. En définissant les variables de décision en $\mathbf{s}=0$ ou $\mathbf{s}=1$, le problème d'optimisation proposé ci-dessus peut établir ces nouvelles répartitions. Pour ce faire, les rendements du plan d'enquête à mode unique et du plan multimodal complet sont des normes de rendement des PEA et peuvent être comparés.

3.2 Optimisation de plans d'enquête adaptatifs statiques

Par analogie avec les méthodes de traitement adaptatif, nous disons que les PEA sont *statiques* lorsqu'ils reposent uniquement sur les renseignements disponibles dans les données de registre et de la base de sondage avant le début de la collecte de données et qu'ils sont *dynamiques* lorsqu'ils reposent (également) sur des paradonnées (données recueillies au cours de la collecte de données). Les PEA *dynamiques* reflètent la nature dynamique de l'optimisation, puisque celle-ci a lieu à chaque étape de la collecte de données, c'est-à-dire à l'issue de chaque mode.

Pour ce qui est des PEA *dynamiques* dans le présent contexte, on prend des décisions d'affectation d'intervieweurs aux strates en fonction des résultats d'enquête intermédiaires provenant des modes précédents. La corrélation entre les propensions à répondre pour le mode sans intervieweur et celles pour le mode avec intervieweur mène à une mise à jour intermédiaire des répartitions *a priori* du mode avec intervieweur. Théoriquement, l'évaluation peut déterminer les priorités des personnes qui refusent de répondre dans les strates contenant des répondants à interviewer et organiser la charge de travail des intervieweurs. En réalité, le temps peut manquer pour calculer à temps la charge de travail des intervieweurs faisant l'objet d'une nouvelle affectation du fait du regroupement géographique. De plus, la réaffectation nécessite une logistique complexe de gestion des cas; nous reviendrons sur ce point dans la section sur l'analyse.

Par conséquent, dans la présente étude, nous nous concentrons sur l'environnement des PEA statiques; autrement dit, la décision d'effectuer un suivi est prise au début de la collecte de données. Nous nous limitons à cette intervention, mais la méthodologie permettrait de multiples interventions, par exemple après plusieurs appels ou visites.

Dans la section 3.3, nous élaborons une stratégie pour tenir compte de l'incertitude de la prise de décision et préciser la routine d'optimisation afin de déterminer les affectations optimales du problème d'optimisation bayésien présenté à la section 3.1.

3.3 Stratégie d'optimisation

Pour résoudre le problème d'optimisation formulé dans la section 3.1, nous proposons une stratégie à deux étapes au moment t :

1. *Élaboration de la répartition a posteriori des nombres de réponses* $r_{g,t}$. Posons que les ensembles de données de série chronologique historiques jusqu'au moment $t-1$ sont des données utilisées pour l'entraînement du modèle et que les ensembles de données au moment t sont des données d'essai pour la prédiction. Tous les coefficients et hyperparamètres de modèle de l'équation (2.8) peuvent être estimés par la taille d'un échantillon $\mathbf{n}_{1:t-1} = \{\mathbf{n}_{g,1:t-1} \mid \forall g\}$ et les nombres de réponses pour tous les modes $\mathbf{r}_{1:t-1} = \{\mathbf{r}_{g,1:t-1,j} \mid \forall g, j\}$. Selon le modèle estimé, des prédictions peuvent être obtenues sur des propensions dépendantes $\tilde{\boldsymbol{\rho}}_t = \{\tilde{\rho}_{g,t,j} \mid \forall g, j\}$ et $\mathbf{r}_t = \{\mathbf{r}_{g,tj} \mid \forall g, j\}$, en fonction des données $\mathbf{n}_t = \{n_{g,t} \mid \forall g\}$.
2. *Détermination des répartitions optimales*. Précisons le niveau budgétaire h et le seuil de dépassement α. Définissons plusieurs vecteurs de départ de répartitions de strates $\mathbf{s}$, chaque vecteur étant considéré comme un état initial et ayant chacun un nombre fini d'états successifs bien définis. Pour toute strate g, supposons K itérations des estimations de $\mathbf{r}_{g,t} = \{r_{g,t,m} \mid \forall j\}$ et $\tilde{\boldsymbol{\rho}}_{g,t} = \left\{\tilde{\rho}_{g,t,j} \mid \forall j\right\}$ générées à partir des répartitions *a posteriori* en 1. Ces estimations *a posteriori* et paramètres donnés h et α sont substitués séparément dans les équations (3.6) et (3.7) pour calculer l'espérance *a posteriori* $\hat{E}\left(\text{c.v.}(s,t)\right)$ et la probabilité *a posteriori* de l'excès de charge de travail. Pour détecter les optima, en commençant à chaque état initial, un tel calcul se déroule en passant par ses états successifs, produit un extrant et finalement atteint un état final. Éliminons les états ne respectant pas les contraintes, et leurs extrants, et conservons les états respectant les contraintes et leurs extrants. Dans ces résultats, additionnons le minimum de $\hat{E}\left(\text{c.v.}(s,t)\right)$ et ses optima de répartition correspondants.

Résoudre ce programme mathématique constitue une tâche qui nécessite beaucoup de calculs. Par conséquent, les méthodes de la première étape sont mises en œuvre dans R à l'aide du paquet *mcmcsae* (Boonstra, 2022), alors que les méthodes de la deuxième étape sont mises en œuvre dans R à l'aide de la fonction *auglag* (algorithme de minimisation lagrangien augmenté) du paquet *Alabama* (Varadhan, 2022) pour une optimisation non linéaire limitée.

3.4 Évaluation du rendement

La présente section introduit un critère d'évaluation afin d'évaluer l'exactitude de la prédiction. Ce critère peut mettre en lumière l'amélioration en matière de réduction du risque de non-réponse des différents

modèles ou des différents plans d'enquête. Cette amélioration est quantifiée par la racine carrée de l'erreur quadratique moyenne (REQM) de la répartition *a posteriori* d'un paramètre τ (par exemple la propension à répondre ou le c.v.), par rapport à la « réalité », cette dernière étant estimée au moyen d'observations.

Nous considérons un rendement tout au long d'une fenêtre dynamique de trois mois. Ce choix est motivé par la durée de trois mois du travail sur le terrain de l'application de la présente étude, mais peut être modifié pour refléter n'importe quelle durée. Dans la fenêtre temporelle $q = \{t, t+1, t+2 \mid \forall t\}$, la REQM de la g^{e} strate est alors exprimée sous la forme

$$\text{REQM}(\tau, q) = \sqrt{B^2(\tau, q) + \text{SD}^2(\tau q)}, \tag{3.8}$$

où le premier terme s'appelle le terme de biais, représenté comme la différence quadratique entre la moyenne *a posteriori* du paramètre τ et le τ observé,

$$B^2(q) = \sum_g d_{g,q} \left(E_{\pi_q} \text{c.v.}(s_q, q) - \widehat{\text{c.v.}}(s_q, q) \right)^2 \tag{3.9}$$

et le second terme est la variance *a posteriori* du c.v., qui est la forme quadratique de l'écart-type (é.-t.),

$$\text{é.-t.}^2(q) = \sum_g d_{g,q} \, \text{Var}_{\pi_q} \text{c.v.}(s_q, q). \tag{3.10}$$

Le poids $d_{g,q} = n_{g,q} \Big/ \sum_g n_{g,q}$ est le ratio de la taille de la strate g sur la taille de l'échantillon dans la fenêtre q. Les répartitions *a posteriori* π_q du c.v. et la répartition $\mathbf{s}_q$ sont dérivées de la stratégie de calcul présentée à la section 3.3.

Ces critères dépendent fortement de la variation d'échantillonnage de la taille de l'échantillon, en particulier pour les enquêtes présentant une petite taille d'échantillon. Des données empiriques assujetties à la variation d'échantillonnage sont utilisées pour évaluer le rendement. Alors que les enquêtes présentant d'importantes tailles d'échantillon fournissent des renseignements riches et, donc que leur rendement peut être évalué avec précision, pour les petites enquêtes, une contradiction par rapport à la variation temporelle devient forte, c'est-à-dire qu'une évaluation précise prend davantage de temps. Il est plus difficile de tirer une conclusion solide sur la mise en pratique de l'adaptation en cas de rendement basé sur des critères instables.

4. Étude de cas de l'Enquête sur la santé aux Pays-Bas

Dans la présente section, nous étudions et exploitons l'application des modèles de série chronologique multinomiaux exposés dans la deuxième section et l'approche d'optimisation dans un cadre bayésien présentée dans la troisième section à l'Enquête sur la santé aux Pays-Bas (enquête GEZO). La section 4.1 présente brièvement le contexte de l'enquête GEZO. Nous montrons séparément dans les sections 4.2 à 4.4 la façon dont les variations temporelles dans des propensions séquentielles peuvent être modélisées, dans quelle mesure le rendement de répartitions optimales dépend du niveau budgétaire et dans quelle mesure des décisions optimales dépendent de la durée des données historiques applicables.

4.1 Enquête sur la santé aux Pays-Bas

L'enquête GEZO est menée chaque année par le bureau central de la statistique des Pays-Bas, pour fournir un aperçu méthodique de l'évolution des contacts médicaux, du mode de vie, de la santé et des comportements préventifs de la population néerlandaise, notamment toutes les personnes vivant dans des ménages privés. Un plan de sondage autopondéré à deux degrés est utilisé; il produit d'abord un échantillon de municipalités, puis un échantillon de personnes vivant dans certaines municipalités. Cette enquête est passée à un plan de sondage multimodal après 2014. La méthode d'observation fait intervenir des interviews en ligne et sur place. Tout d'abord, des IWAO permettent de demander la participation d'unités d'échantillonnage de la population. Ensuite, des non-répondants sont recrutés pour participer à des IPAO. À compter de 2018, cependant, une adaptation a été mise en œuvre pour stabiliser la charge de travail des intervieweurs. Seule une portion des non-répondants aux IWAO est abordée de nouveau pour participer à une IPAO, afin de réduire les coûts de l'enquête et d'améliorer la représentativité. Les taux de réponse plus élevés des unités d'échantillonnage des IWAO signifient une possibilité moindre d'être abordés de nouveau.

Dans la présente étude, nous nous concentrons sur une série chronologique de données recueillies de 2014 à 2017, comprenant 48 mois. Il convient de mentionner que les données recueillies au début de 2017 étaient « anormales », du fait de problèmes techniques relatifs au serveur Web, ce qui a entraîné une interruption de la collecte de données. Cela se justifie par des raisons pratiques : le bureau central de la statistique des Pays-Bas a mis en œuvre des PEA statiques depuis 2018; cette adaptation peut gâcher la valeur potentielle des données historiques utilisées pour améliorer l'exactitude de la prédiction (Wu et coll., 2023). De plus, les unités d'échantillonnage sont stratifiées en 13 strates disjointes par deux variables auxiliaires provenant de la base de sondage ou de registres administratifs : l'âge et l'ethnicité. Voir la stratification à l'annexe A. Il convient de souligner que cette stratification est fixe tout au long de la présente étude et que nos strates de série chronologique sont différentes des strates de PEA (Van Berkel et coll., 2020). L'ensemble de variables auxiliaires disponibles avant le début du travail sur le terrain est bien plus grand. Il comprend plusieurs données démographiques, comme le genre, le pays de naissance, la composition du ménage, des caractéristiques socioéconomiques, comme le revenu personnel et le revenu du ménage enregistrés, le niveau de scolarité, le type de profession, ainsi que des caractéristiques relatives au logement et à la région, comme le type de logement, la valeur du logement et l'urbanisation. Une étude portant sur la sélection efficiente et efficace des strates est une étape importante à venir.

4.2 Comment construire un modèle de série chronologique sous un plan séquentiel multimodal

La présente section fournit des précisions sur l'approche permettant de créer des modèles de série chronologique hiérarchiques. Nous observons deux décisions potentiellement influentes. La première décision consiste à inclure une composante saisonnière. L'ajout d'une saison améliorera probablement l'exactitude, mais comporte un coût. Inclure une saison signifie qu'une série plus longue de données historiques est nécessaire; il faut observer au moins deux années de données pour évaluer une composante

saisonnière, mais préférablement davantage. La deuxième décision est l'inclusion d'associations ou de corrélations explicites entre les paramètres du modèle. Encore une fois, l'exactitude sera probablement légèrement plus élevée, mais davantage de données sont nécessaires. Nous pouvons encore différencier les décisions à trois autres niveaux : la strate, le temps et le mode. Nous pouvons ainsi rendre les paramètres des composantes saisonnières et les corrélations dépendants du mode et de la strate, et même du temps. Nous explorons quatre scénarios : composante saisonnière présente ou absente multipliée par corrélation présente ou absente.

Notre objectif est de trouver le scénario le plus avantageux parmi ces quatre scénarios. Puisqu'il existe de nombreux modèles possibles, par analogie avec Wu et coll. (2023), nous adoptons une stratégie pas à pas pour passer de modèles simples à des modèles plus avancés. En évaluant le rendement, nous tenons compte de deux critères d'information : le critère d'information de déviance (DIC, voir Spiegelhalter, Best, Carlin et van der Linde, 2002) et le critère d'information de Watanabe-Akaike (WAIC, voir Watanabe, 2010 et 2013). En offrant un compromis raisonnable entre la qualité de l'ajustement du modèle et la complexité du modèle, ces mesures devraient permettre de sélectionner des modèles adéquats pour la tâche de prédiction qui nous intéresse. Nous pouvons les interpréter comme étant des approximations de mesures de validation croisée avec retrait d'un élément, relativement faciles et peu onéreuses à calculer à partir de l'extrant de la simulation de la méthode MCMC (Vehtari, Gelman et Gabry, 2017). Lors de la construction des modèles, nous pouvons choisir parmi divers effets aléatoires : le bruit blanc, la tendance temporelle globale, les ordonnées à l'origine aléatoires pour chaque strate et les tendances temporelles propres à la strate. De plus, nous pouvons faire cela avec ou sans saisonnalité et disposons de paramètres de modèles dépendants ou indépendants du mode. Nous suivons les étapes ci-dessous :

1. Modèle de base : définir deux modèles de base; l'un sans composantes saisonnières et l'autre ayant des composantes saisonnières. Les deux modèles présentent des effets de mode fixes et des effets fixes de variables auxiliaires propres au mode β. La différence entre les deux modèles est l'inclusion ou non de la saisonnalité δ.
2. Tendance de base : ajouter un seul effet aléatoire, c'est-à-dire {tendance temporelle globale u_t}, {ordonnées à l'origine aléatoires pour les strates v_g} ou {bruit blanc e_{gt}} aux modèles de 1. Chaque effet aléatoire est corrélé avec des modes ou rendu indépendant des modes. Examiner si le modèle avec ou sans saisonnalité de 1 est amélioré par l'un des trois effets aléatoires.
3. Certaine dépendance à la strate : ajouter une combinaison d'effets aléatoires, c'est-à-dire, {tendance temporelle globale u_t, ordonnées à l'origine aléatoires pour les strates v_g} ou {ordonnées à l'origine aléatoires pour les strates v_g, tendance temporelle propre à la strate z_{gt}}, aux modèles de 1. Chaque effet aléatoire est corrélé avec des modes ou est indépendant des modes. Examiner si les modèles mis à jour fournissent de meilleurs résultats que les modèles exposés au point 2.
4. Dépendance modérée à la strate : ajouter une combinaison de trois effets aléatoires, c'est-à-dire $\{u_t, v_g, e_{gt}\}$ ou $\{u_t, v_g, z_{gt}\}$, aux modèles de 1. Chaque effet aléatoire est corrélé avec des modes ou est indépendant des modes. Examiner si les modèles mis à jour fournissent de meilleurs résultats que les modèles exposés au point 3.

5. Dépendance complète à la strate : ajouter tous les effets aléatoires aux modèles de 1. Chaque effet aléatoire est corrélé avec des modes ou est indépendant des modes. Examiner si la combinaison complète améliore le rendement du modèle.

Le tableau 4.1 présente les résultats de ces cinq étapes. Comme l'indique chaque ligne du tableau 4.1, les modèles avec et sans saisonnalité obtiennent une correspondance égale en matière d'ajustement et de prédiction. Les résultats des critères d'information (CI) des deux modèles de base (modèle 1) indiquent que le rendement du modèle avec saisonnalité est légèrement supérieur à celui du modèle sans saisonnalité. Cet avantage se maintient lorsque certains effets aléatoires s'ajoutent (voir les modèles 1, 3 et 6), alors que l'inclusion de la saisonnalité δ produit des CI inférieurs. D'autre part, les modèles avec saisonnalité présentent un rendement légèrement plus faible, comme on le voit pour les modèles 2, 4, 7, 8 et 9. En particulier, les résultats du modèle 5 présentent des résultats mitigés. Les résultats du DIC appuient la modélisation des effets saisonniers pour des prédictions de propension exactes, mais le WAIC ne peut pas s'y conformer.

En matière d'équilibre entre la complexité et la qualité de l'ajustement du modèle, le rendement des modèles indépendants des modes est pratiquement équivalent aux modèles corrélés aux modes, malgré un rendement légèrement plus élevé (modèle 5 pour le modèle sans saison et le modèle 6).

Tableau 4.1
Critères d'information et chiffres réels lors de l'évaluation de la qualité de l'ajustement et de la complexité des modèles.

Modèle	Effet fixe	Effet aléatoire	DIC		p_{DIC}		WAIC		p_{WAIC}	
			IND	COR	IND	COR	IND	COR	IND	COR
1	β	-	6 808		18		6 817		28	
	β, δ		6 800		25		6 813		38	
2	β	u_t	6 504	6 504	67	67	6 509	6 509	72	72
	β, δ		6 506	6 505	70	70	6 510	6 510	74	74
3	β	v_g	6 775	6 774	26	26	6 786	6 785	36	36
	β, δ		6 768	6 767	32	32	6 781	6 780	46	45
4	β	e_{gt}	6 602	6 601	322	321	6 479	6 479	200	200
	β, δ		6 609	6 608	322	321	6 488	6 487	201	200
5	β	u_t, v_g	6 488	6 490	147	146	6 472	6 475	131	130
	β, δ		6 474	6 472	77	77	6 477	6 476	80	81
6	β, δ	v_g, z_{gt}	6 494	6 495	143	144	6 481	6 482	130	129
	β, δ		6 488	6 490	147	146	6 472	6 475	131	130
7	β	u_t, v_g, e_{gt}	6 449	6 448	195	194	6 398	6 396	143	142
	β, δ		6 453	6 452	191	191	6 404	6 402	142	142
8	β	u_t, v_g, z_{gt}	6 396	6 395	109	108	6 381	6 380	94	94
	β, δ		6 398	6 397	112	111	6 382	6 381	96	95
9	β	u_t, v_g, z_{gt}, e_{gt}	6 398	6 398	137	137	6 371	6 371	110	110
	β, δ		6 400	6 397	132	134	6 375	6 372	108	108

Note : La série chronologique de 2014 à 2017 est ajustée à un modèle indépendant du mode (« IND ») et un modèle est corrélé au mode (« COR »). Chaque modèle est simplifié en utilisant uniquement les composantes d'effets fixes, puis en ajustant la corrélation au fil du temps ou entre les modes en tenant compte de plusieurs effets aléatoires.
DIC = Deviance Information Criterion;
WAIC = Widely Applicable Information Criterion.

En tenant compte des effets aléatoires, les modèles mixtes sont meilleurs, car ils entraînent une diminution des CI, contrairement aux modèles qui comportent uniquement des effets fixes (modèle 1). Comparer les modèles 2 à 4 au modèle 1 signifie que le modèle avec ou sans saisonnalité est amélioré en introduisant un seul effet aléatoire, où la tendance temporelle globale u_t entraîne la plus grande diminution des CI, suivie d'un bruit blanc e_{gt} et d'une ordonnée à l'origine aléatoire pour la strate. Une telle amélioration persiste pour les CI en appliquant la combinaison de deux effets aléatoires, comme le montrent les comparaisons du modèle 5 avec les modèles 2 et 3, et du modèle 6 avec le modèle 3. Apparemment, le modèle 5 présente la diminution la plus forte des CI jusqu'à présent. Les modèles 7 et 8 indiquent que les modèles peuvent être encore améliorés en ajoutant le bruit blanc e_{gt} et une tendance temporelle propre à la strate z_{gt} au modèle 5, et dans le modèle 8, les CI diminuent davantage que dans le modèle 7. Inclure le bruit blanc e_{gt} est utile pour obtenir un meilleur rendement, car cela ajoute peu à la diminution du WAIC du modèle 9 malgré la faible contribution apportée au DIC.

Comme le montre le modèle 9, les modèles corrélés au mode et indépendants du mode (respectivement les colonnes COR et IND) présentent un rendement semblable en matière de CI en ne tenant pas compte de la saisonnalité; toutefois, pour les modèles avec saisonnalité, la modélisation des corrélations (colonne COR) est meilleure en ce qui a trait aux résultats de CI par rapport à la colonne IND. Il est cependant difficile de conclure que le modèle avec saisonnalité présente un avantage absolu par rapport au modèle sans saisonnalité en termes de qualité de l'ajustement et de complexité du modèle. Pour déterminer si les effets saisonniers jouent un rôle essentiel dans des répartitions adaptatives, nous considérons les modèles sans saisonnalité et avec saisonnalité (indiqués en rouge) dans la section 4.4.

4.3 Dans quelle mesure le rendement des plans d'enquête adaptatifs est-il sensible au niveau budgétaire donné ?

Cette question de recherche porte sur la façon d'adapter les répartitions des non-réponses aux IWAO pour toutes les strates, afin de réduire au maximum le risque de non-réponse, en fonction d'un niveau budgétaire donné. Elle soulève également la question de savoir si une telle réduction peut se maintenir pour différents niveaux budgétaires.

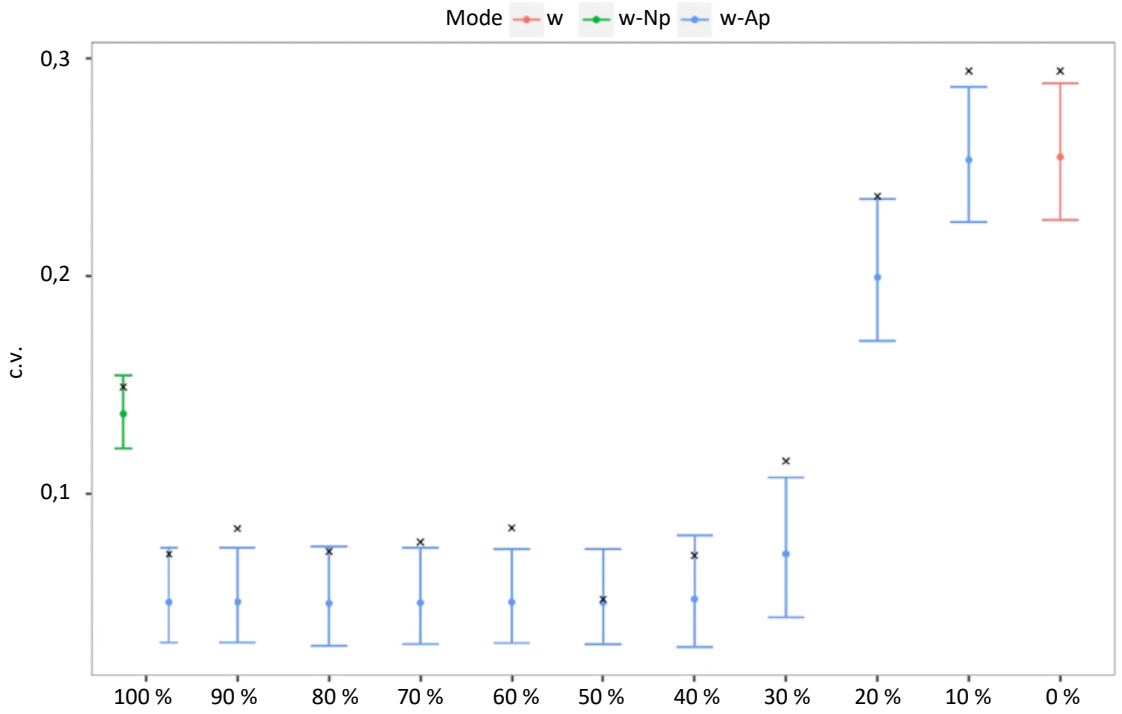
Nous répondons à cette question en réduisant d'abord au minimum (3.6) en fonction de (3.5) et (3.7) pour le trimestre de collecte de données suivant, lorsque le niveau budgétaire est précisé, puis en comparant la valeur optimale de (3.5) pour le c.v. réalisé pour le même niveau budgétaire, et enfin, en comparant la valeur optimale de (3.5) pour différents niveaux budgétaires. Nous nous concentrons sur le trimestre suivant, car, dans le cas statique, le nombre de répondants aux IWAO est inconnu tant que les données ne sont pas recueillies et car l'échantillon suffisant d'un trimestre peut assurer la précision de la prédiction. En se reportant à la stratégie d'optimisation de la section 3.3, la procédure d'évaluation du trimestre q est la suivante :

1. Soit le niveau budgétaire h commençant à 100 % puis diminuant successivement par paliers de 10 %, c'est-à-dire $h \in \{1, 0{,}9, 0{,}8, 0{,}7, 0{,}6, 0{,}5, 0{,}4, 0{,}3, 0{,}2, 0{,}1\}$.
2. Considérons le trimestre à venir q et l'ensemble de données dans q comme l'ensemble de données d'essai.
3. Posons les données de la série chronologique jusqu'au trimestre $q-1$ comme l'ensemble de données d'essai pour estimer les modèles sélectionnés. Les modèles comprennent les composantes du modèle 9 considérées comme la « meilleure » représentation.
4. Utilisons la taille d'échantillon dans q pour simuler les réponses aux IWAO. Pour chaque strate et chaque mois dans q, 3 000 tirages sont générés à partir des répartitions prédictives *a posteriori*.
5. Selon le modèle simulé à l'étape 3, des prédictions individuelles *a posteriori* des IWAO et des IPAO conditionnelles sont générées séparément 3 000 fois pour chaque strate et chaque mois dans q.
6. Substituons le niveau budgétaire précisé h et les réponses aux IWAO simulées à l'étape 4 dans la contrainte de coût (3.7).
7. Calculons les propensions mixtes en substituant le niveau h et les prédictions individuelles de l'étape 5 dans (3.3).
8. Initialisons trois solutions de départ de probabilités de répartition, $s \in \{0, 0.5, 1\}$, chacune s'appliquant à 13 strates simultanément.
9. Commençons à chaque point initial de l'étape 8 pour trouver les solutions optimales pour chaque strate par le résolveur *auglag* en fonction des étapes 6 et 7.
10. Faisons le lien entre les solutions relevées et l'échantillon réel pour le calcul des prédictions de c.v. *a posteriori* et les réalisations de c.v.
11. Effectuons des comparaisons en répétant les étapes 2 à 10 pour chaque niveau budgétaire exposé à l'étape 1.

$s = 0$ indique aucun suivi par IPAO, $s =$ 0,5 signifie que la moitié des non-réponses aux IWAO a été affectée aux IPAO, et $s = 1$ représente un suivi complet par IPAO. Pour distinguer les stratégies des différents modes et simplifier la notation, les plans IWAO uniquement, les plans non adaptatifs et les plans adaptatifs sont désignés par w, w-Ap et w-Np dans l'ensemble du document.

Dans la figure 4.1, les prédictions de c.v. *a posteriori* pour le premier trimestre (T1) 2017 sont résumées pour chaque niveau budgétaire; voir le tableau B.1 de l'annexe B pour connaître les résultats de c.v. tenant compte du biais. Nous étalonnons le rendement w-Ap en tant que fonction du niveau budgétaire h par rapport au rendement de w et de w-Np. Par souci de concision, les c.v. pour les plans par IWAO uniquement, les plans non adaptatifs et les plans adaptatifs, sont simplifiés sous la forme c.v.(w), c.v.(w-Np) et c.v.(w-Ap).

Figure 4.1 Comparaison des coefficients de variation (c.v.) des prédictions de propension à répondre fondées sur un modèle comportant des observations de c.v. et tenant compte du biais au premier trimestre 2017.



Note : Les estimations de c.v. sont effectuées séparément pour les plans par IWAO uniquement (« w »), les plans non adaptatifs (« w-Np ») et les plans adaptatifs (« w-Ap »). Les prédictions de c.v. *a posteriori* sont résumées par la région crédible à 95 %, alors que les observations sont indiquées par un nuage de points (« x »).

La comparaison du c.v.(w-Np) avec le c.v.(w) indique que le recrutement des non-répondants aux IWAO par IPAO peut donner lieu à une diminution de la non-réponse, puisque la région crédible à 95 % (intervalle de confiance [IC]) du c.v.(w-Np) *a posteriori* est bien plus étroite que celle du c.v.(w) et le quantile à 97,5 % du c.v.(w-Np) *a posteriori* est bien inférieur au quantile à 2,5 % du c.v.(w). Lorsque $h = 100$ %, une autre diminution de la variation globale est obtenue par les répartitions optimisées de l'enquête adaptative. Les prédictions *a posteriori* et les observations du c.v.(w-Ap) divergent de 0,1 et se rapprochent de 0 par rapport au quantile à 2,5 % du c.v.(w-Np); cependant, l'IC de l'approche adaptative indique que l'exactitude de la prédiction est relativement compromise. Puisque les IC sont à peine modifiés lorsque le budget est réduit pour passer de 90 % à 50 %, la réduction de l'incertitude associée au c.v.(w-Ap) n'augmenterait probablement pas de plus de 100 %. Cela signifie que dans l'intervalle des niveaux de 100 % à 50 %, l'effet du faible budget sur le risque de non-réponse estimée est équivalent à celui du budget élevé. Les limites supérieures des IC semblent s'approcher des c.v. observés, voire les dépasser; par exemple, au niveau de 50 %, l'observation chevauche la moyenne *a posteriori*.

Le risque de non-réponse augmente avec la réduction continue du budget, puisque les estimations du c.v.(w-Ap) augmentent et indiquent un risque accru de biais de non-réponse. Pour les niveaux budgétaires inférieurs à 50 %, la méthode de répartition indiquée accroît l'incertitude des estimations *a posteriori* de la variation globale. De plus, les limites inférieures se rapprochent de 0,1, et même le dépassent largement, lorsque le niveau est de 20 % ou de 10 %, auxquels cas le résolveur finit par obtenir un « optimum » local

faux, du fait du non-respect du critère de convergence. Pour le niveau de 10 %, l'intérêt de la méthode de répartition est particulièrement réduit et elle perd son avantage, comme le montre le c.v.(w-Ap) exactement identique au c.v.(w). Pour déterminer le niveau budgétaire préférable, nous adoptons un critère : le coût relatif défini comme le coût global de la taille adaptative pour les IPAO par rapport à la taille non adaptative limitée par le niveau budgétaire. Voir le tableau B.3 à l'annexe B pour connaître les résultats du coût relatif pour différents niveaux.

Des répartitions optimisées rendent le rendement de l'adaptation constant pour des niveaux budgétaires relativement élevés (100 % à 50 %). De plus, même si cela contribue légèrement à des pertes de précision, l'adaptation est néanmoins efficace pour le risque de non-réponse estimé plus faible par rapport aux plans w et w-Np (barres d'erreur rouges et vertes). Jusqu'à un niveau budgétaire de 40 %, le rendement s'inverse et évolue dans la direction opposée, ce qui signifie que le risque de non-réponse augmente nettement. On s'attend à ce que ce comportement général soit relativement robuste pour des variations raisonnables dans le choix de modèle, car les deux extrêmes de budget ne laissent aucune marge dans la répartition des PEA pour compenser des différences de propensions à répondre.

4.4 Dans quelle mesure le rendement des plans adaptatifs dépend-il des données historiques disponibles ?

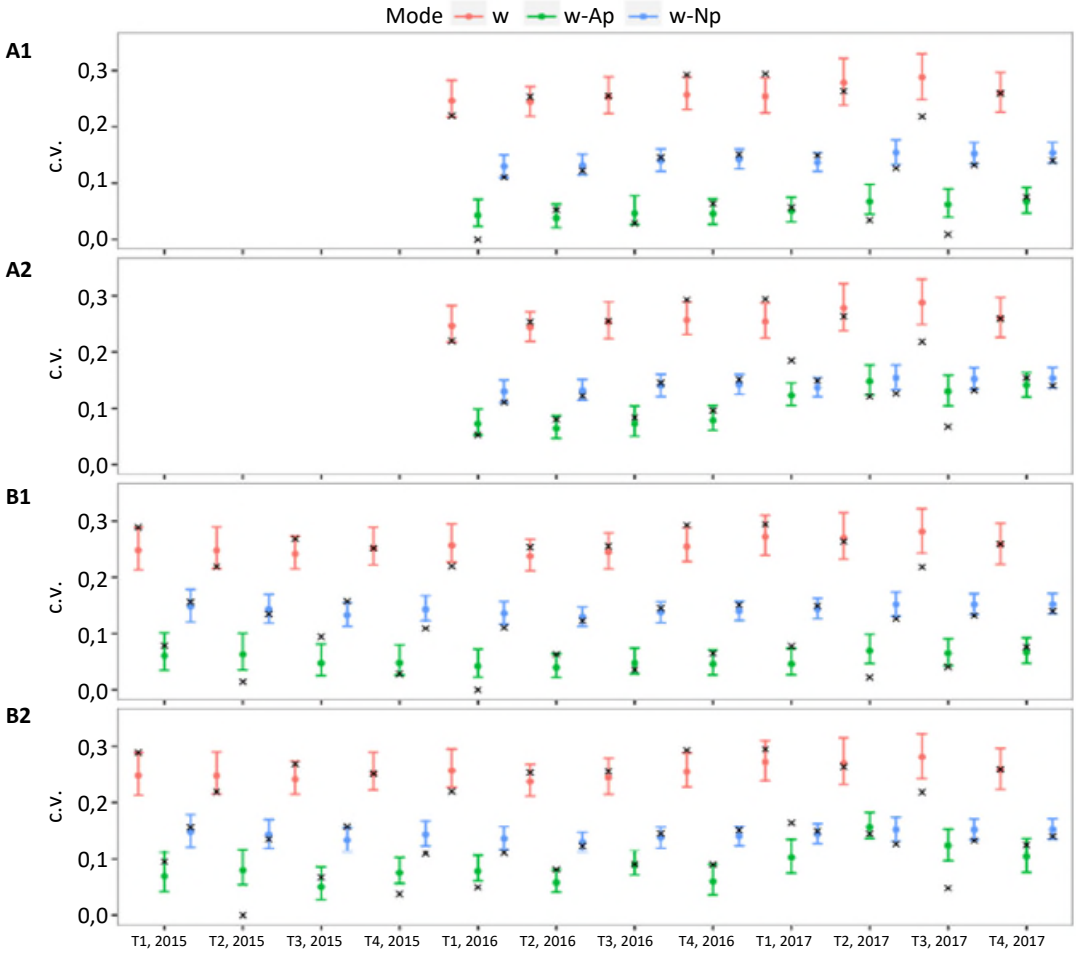
Cette question nécessite d'examiner dans quelle mesure la série chronologique historique accumulée influence le rendement des plans adaptatifs, c'est-à-dire le risque de non-réponse mesuré par les c.v. et l'équilibre entre le biais et la variance mesuré par la REQM. Pour répondre à cette question, nous étudions le rendement des plans w, w-Np et w-Ap au niveau du trimestre civil. De plus, nous étalonnons le rendement des plans adaptatifs par rapport au rendement des plans w et w-Np.

Nous comparons et évaluons les modèles avec et sans l'inclusion de la saisonnalité, et ce, pour des niveaux budgétaires de 50 % et de 30 %. La section 4.2 indique que la saisonnalité est un facteur pouvant être ignoré, puisque les modèles avec et sans ce résultat présentent une qualité de l'ajustement et une complexité semblables. La section 4.3 donne à penser que pour une fenêtre temporelle particulière, le niveau budgétaire de 50 % favorise le rendement de PEA le plus rentable et que le PEA perd son avantage absolu pour des valeurs plus petites. La sensibilité du rendement à la durée de la série chronologique est moins évidente lorsque les modèles tiennent compte de la saisonnalité ou lorsque le niveau budgétaire est inférieur à 50 %; il est donc prématuré de les ignorer dans l'analyse. En croisant les deux conditions, il est possible de procéder à des comparaisons simultanées des quatre scénarios de modèles : 1) l'inclusion de la saisonnalité et un niveau budgétaire de 50 %; 2) l'inclusion de la saisonnalité et un niveau budgétaire de 30 %; 3) la non-inclusion de la saisonnalité et un niveau budgétaire de 50 %; 4) la non-inclusion de la saisonnalité et un niveau budgétaire de 30 %.

Pour étudier la sensibilité à la durée de la série chronologique historique, l'analyse est effectuée sur une base dynamique en ajoutant un mois à chaque fois. Il convient de se rappeler que la durée de la série chronologique historique initiale devrait être d'au moins un an pour les modèles n'incluant pas la saisonnalité (scénarios 3 et 4), mais d'au moins deux ans pour les modèles comprenant la saisonnalité (scénarios 1 et 2). Pour chacun, le processus d'entraînement se termine au troisième trimestre de 2017), car un trimestre doit demeurer pour la prédiction.

Dans la figure 4.2, l'incertitude relative aux c.v. estimés évaluée par la région crédible à 95 % et les moyennes *a posteriori* sont, ensemble, comparées avec les observations des c.v. tout au long des trimestres et selon différents plans. Dans les plans w et w-Np, les c.v. observés se situent dans les intervalles ou sont très proches des limites des intervalles pour la plupart des trimestres, à l'exception du c.v.(w) du troisième trimestre de 2017. Les résultats des PEA des sections A1, A2 et B2 soutiennent cette constatation. De plus, des observations très éloignées des IC apparaissent au deuxième trimestre de 2015 dans la section B2 et au premier trimestre de 2017 dans les sections A2 et B2. Cette exception laisse entendre que pour des trimestres correspondants, on est moins convaincu que le rendement adaptatif évalué reproduit le rendement en pratique.

Figure 4.2 Pour un niveau budgétaire donné, coefficient de variation (c.v.) *a posteriori* pour les plans adaptatifs, les plans non adaptatifs et les plans par IWAO uniquement, par rapport aux observations tout au long des trimestres.



Note : Les régions crédibles à 95 % avec les espérances *a posteriori* sont résumées pour les plans par IWAO uniquement (« w »), les plans non adaptatifs (« w- p), et les plans adaptatifs (« w-Ap »). On désigne les observations par des croix « x » noires. Les sections « A » présentent les résultats des modèles avec saisons et les sections « B », les résultats des modèles sans saisons. Les sections « A1 » et « B1 » correspondent au niveau budgétaire de 50 %, alors que les sections « A2 » et « B2 » correspondent au niveau budgétaire de 30 %. Le trimestre sur l'axe des x désigne le trimestre existant à des fins de prédiction.

Comme cela a été mentionné à la section 4.3, le rendement des plans d'enquête adaptatifs au niveau budgétaire de 50 % est constamment supérieur au rendement des plans au niveau budgétaire de 30 % pour tous les trimestres, comme l'indiquent les comparaisons de A1 à A2 ou de B1 à B2. Au niveau de 50 %, le c.v.(w-Ap) estimé est plus précis du fait de régions crédibles plus étroites, ce qui est implicitement visible. De plus, nous pouvons observer l'avantage absolu de plans adaptatifs dont le rendement est supérieur à celui de plans non adaptatifs, puisque les limites supérieures du c.v.(w-Ap) s'écartent de façon substantielle des limites inférieures du c.v.(w-Np); cependant, au niveau de 30 %, ils sont concurrentiels, par exemple pour les premier et deuxième trimestres de 2017 (voir les sections A2 et B2).

À mesure que les données historiques s'accumulent, on suppose que les modèles peuvent être encore optimisés; l'exactitude de la prédiction des c.v. peut présenter une augmentation systématique et le rendement obtenu peut être amélioré. Il est clair que cela est le cas dans la section B1 jusqu'au quatrième trimestre 2016; toutefois, à partir de cette date, il semble ne plus être possible d'améliorer le rendement, et les estimations *a posteriori* du c.v.(w-Ap) se stabilisent pour se situer à environ 0,1. Les résultats pour les sections A1 et B1 sont similaires; on doit donc conclure que la modélisation de la saisonnalité contribue peu à l'exactitude de la prédiction.

Tirer des conclusions sur la robustesse du rendement de PEA est dogmatique pour deux raisons. Tout d'abord, l'avantage peut être plus grand pour certaines strates. Leurs estimations individuelles de c.v. peuvent être moins biaisées et varier pour tendre vers l'option qui inclut la saisonnalité malgré peu de différence dans la variation globale. Deuxièmement, les tailles d'échantillons de certaines strates sont relativement petites. Tôt au cours des étapes de collecte de données, elles peuvent présenter un comportement volatile quant à l'équilibre entre l'erreur et la variation. La durée de la série chronologique historique déterminée en fonction de ces résultats n'est pas une garantie de robustesse.

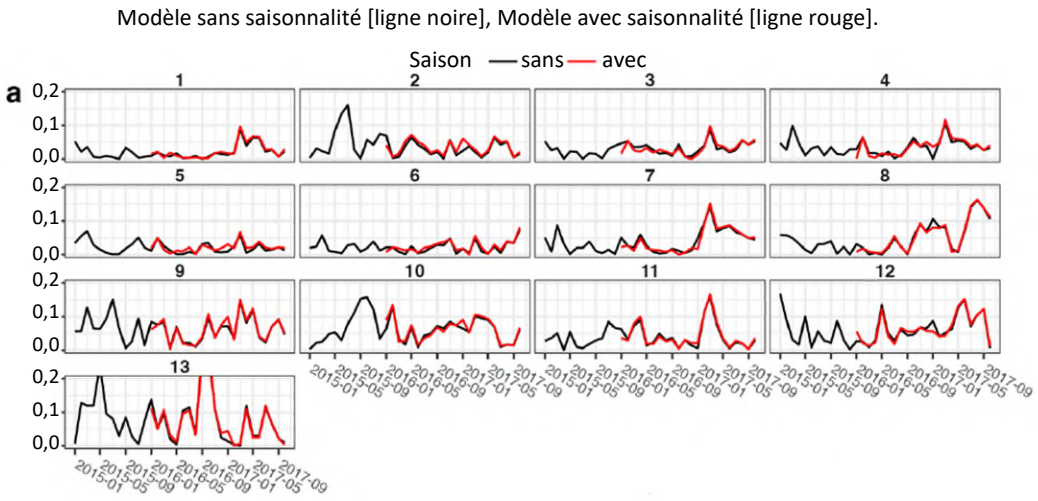
Par conséquent, nous évaluons le rendement de strates individuelles mesuré selon le critère mentionné à la section 3.4. Comme plus haut, nous appliquons l'approche de la fenêtre dynamique tout au long d'une série chronologique. À titre d'illustration, cela est utilisé dans les plans non adaptatifs. Dans le cadre d'une application aux PEA, les répartitions doivent être optimisées de nouveau pour la fenêtre temporelle suivante à l'aide de la stratégique de la section 3.3.

Cette fenêtre temporelle se déplace à mesure que la largeur augmente pour inclure la nouvelle période temporelle suivante. Au trimestre q, nous pouvons évaluer le rendement de la prédiction pour chaque strate en substituant des estimations individuelles *a posteriori* des propensions à répondre et des réalisations individuelles dans (3.8) à (3.10), c'est-à-dire $\text{REQM}(g,q)$, $B(g,q)$ et é.-t.(g,q). Puisque cette analyse est itérée de manière continue, une série chronologique suffisamment longue permet de bien comprendre la variation de chaque rendement de prédiction de strate au fil du temps.

La figure 4.3 montre que lors d'une comparaison des modèles sans saisonnalité (courbes rouges) et des modèles avec saisons (courbes noires), l'introduction de la saisonnalité ne déclenche probablement pas de

réduction efficace du biais et de la variance. Cela s'observe clairement pour pratiquement tous les trimestres, à l'exception du trimestre comprenant les mois de janvier à mars 2017 dans certaines strates (comme la strate 8) pour les estimations des biais et de la REQM. Selon les observations des sections a et b, la variation estimée de la propension à répondre diminue globalement en douceur, ce qui contraste vivement avec le niveau estimé de propension à répondre présentant un comportement volatile. Cette volatilité diffère selon les strates. Les résultats estimés des biais de certaines strates (strates 1 à 8) fluctuent d'environ 0,05 pour tous les trimestres jusqu'à celui commençant en janvier 2017. À partir de cette date, ils présentent une hausse temporaire en raison d'un problème technique à ce moment (voir Wu et coll., 2023 pour une analyse et une éventuelle solution). Lorsque les données d'entraînement sont étendues pour inclure les données « normales », les biais peuvent rapidement diminuer pour atteindre 0,05. Il convient de mentionner que la strate 8 se comporte de manière opposée. Les strates 9 à 13, qui portent sur des tailles d'échantillons relativement petites, obtiennent quant à elles des estimations de l'espérance de la propension à répondre relativement plus biaisées pour la plupart des trimestres.

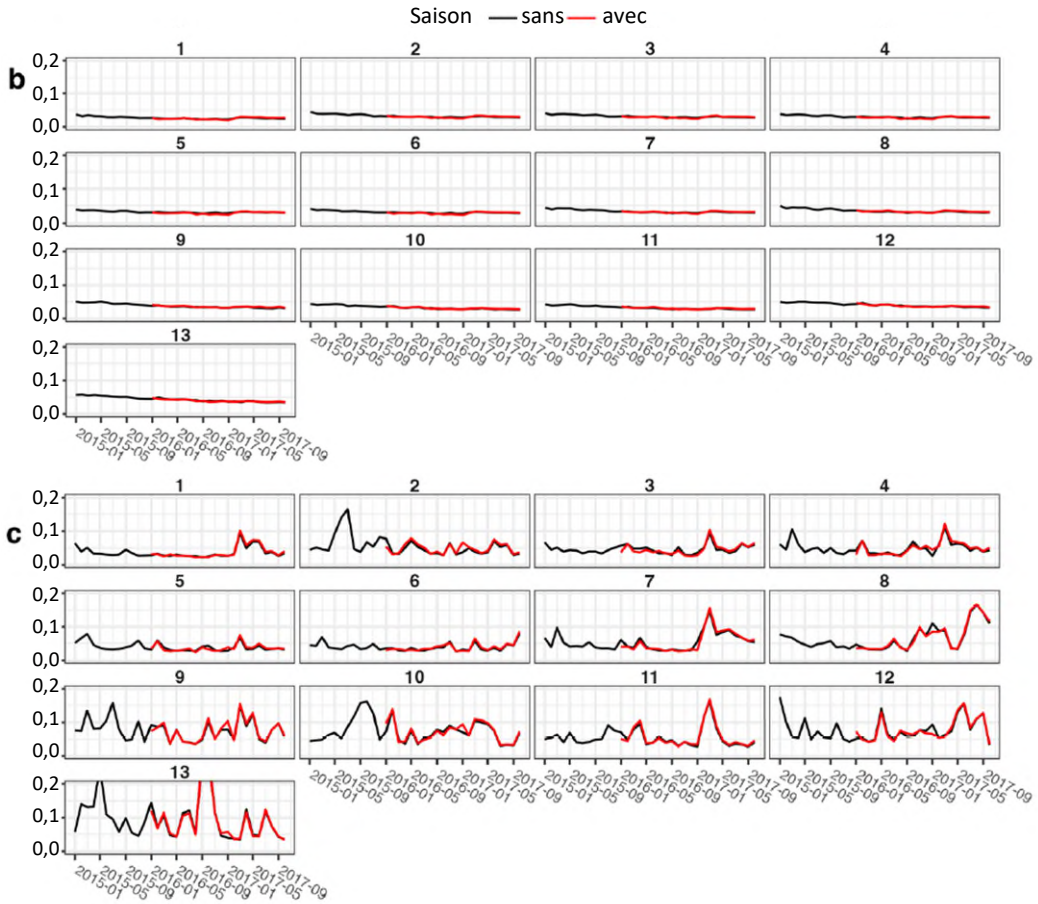
Figure 4.3 Moyennes mobiles prévisionnelles à une étape du biais estimé (section a), de l'écart-type (section b) et de la racine carrée de l'erreur quadratique moyenne (section c) de la propension à répondre au niveau de la strate.



Note : La courbe « noire » représente le modèle sans saisonnalité, alors que la courbe « rouge » fait référence au modèle avec saison. Les deux modèles incluent les corrélations entre les IWAO et les IPAO concernant les prédictions de propensions. L'axe horizontal représente le moment où une décision de PEA est prise.

Figure 4.3(suite) Moyennes mobiles prévisionnelles à une étape du biais estimé (section a), de l'écart-type (section b) et de la racine carrée de l'erreur quadratique moyenne (section c) de la propension à répondre au niveau de la strate.

Modèle sans saisonnalité [ligne noire], Modèle avec saisonnalité [ligne rouge].



Note : La courbe « noire » représente le modèle sans saisonnalité, alors que la courbe « rouge » fait référence au modèle avec saison. Les deux modèles incluent les corrélations entre les IWAO et les IPAO concernant les prédictions de propensions. L'axe horizontal représente le moment où une décision de PEA est prise.

Finalement, les résultats de l'analyse indiquent que lors de la modélisation d'une série chronologique courte, les effets saisonniers, lorsque l'on suppose qu'ils sont semblables pour différents modes, peuvent être moins importants pour l'amélioration du rendement d'un PEA. Lorsque davantage de données d'entraînement sont disponibles, le rendement du PEA peut être constamment amélioré jusqu'à un moment donné; cela sous-entend qu'une règle pour arrêter la collecte de données peut être mise en place et qu'une stratégie fondée sur l'effort peut être adoptée pour les strates présentant de petites tailles d'échantillon.

5. Analyse

En fonction du budget d'une enquête, un PEA cherche la correspondance optimale entre le comportement des répondants et les caractéristiques du plan, c'est-à-dire un ensemble de règles de décision qui peut être

déterminé par des approches d'optimisation (voir Schouten et coll., 2017 pour connaître les avantages et les inconvénients de diverses démarches). Servant d'intrants principaux pour l'optimisation de PEA, des estimations exactes de paramètres de plans d'enquête, comme des propensions à répondre, sont nécessaires pour toute stratégie fiable. À proprement parler, l'inexactitude compromet le rendement et la conception d'un PEA du fait des décisions sous-optimales et inefficaces prises dans le cadre de la démarche d'optimisation. Une incidence négative est évidente lorsque les propensions à répondre varient progressivement au fil du temps.

Dans la présente étude, nous avons discuté d'une méthodologie permettant d'évaluer l'incidence de facteurs temporels (par exemple la saisonnalité) sur l'exactitude de prédictions de propensions séquentielles à répondre dans une enquête multimodale à répliques, et exploré la manière d'appliquer des méthodes de répartition optimale à des strates de population et l'échéancier correspondant. Nous avons introduit un modèle chronologique multinomial bayésien pour les propensions séquentielles à répondre et un modèle d'optimisation de PEA. La forme générale du modèle relatif à la propension décrivait de multiples facteurs temporels et liés aux strates et tenait compte de la dépendance des propensions à répondre du mode actuel envers les comportements de réponse des modes précédents. Le modèle d'optimisation, en revanche, permettait l'inclusion de l'incertitude relative à la charge de travail de suivi et décrivait la façon d'affecter des réviseurs à chaque strate pour favoriser la plus forte diminution du risque de non-réponse. Ce cadre correspond à la plupart des enquêtes transversales multimodales ou à mode unique menées sur de nombreuses années. De plus, nous avons élaboré une analyse pour l'enquête GEZO, afin d'examiner le rendement le plus élevé du modèle de propension. Du fait de diverses compositions de modèle, des critères d'information mesurant la qualité de l'ajustement et la complexité du modèle de propension ont été adoptés pour comparer le rendement de différents modèles. Nous avons ainsi pu atteindre le premier objectif de cette étude visant à sélectionner et à élaborer le modèle de série chronologique « favori » (le modèle 9 enregistrant les résultats de critères d'information les plus bas) contribuant le plus à l'exactitude des prédictions pour une enquête séquentielle multimodale.

Les deuxième et troisième objectifs ont été d'examiner la sensibilité du rendement de PEA au niveau budgétaire donné et au volume de données historiques incluses, respectivement. Dans l'évaluation, le rendement des PEA a dû être optimisé à nouveau, lors de la mise à jour du niveau budgétaire ou de la durée de la série chronologique de données historiques applicables. Ainsi, nous avons étalonné le rendement des PEA par rapport au rendement des plans par IWAO uniquement et des plans non adaptatifs. Cette analyse est essentiellement une comparaison de la réduction du risque de non-réponse, lors de l'affectation à des intervieweurs d'une fraction de non-réponses aux IWAO (sans aucun suivi et avec suivi complet comme cas particulier). Pour que cela soit comparable pour un éventail de scénarios, nous avons utilisé les propriétés des répartitions *a posteriori*, c'est-à-dire la région crédible et l'espérance. L'évaluation a permis d'examiner, pour une fenêtre de temps précise, l'amélioration du rendement pour différents niveaux budgétaires. De

plus, elle a permis d'examiner, pour un niveau budgétaire particulier, l'amélioration par rapport à des fenêtres temporelles dynamiques. L'évaluation a révélé que le rendement d'un PEA était relativement robuste pour des niveaux budgétaires supérieurs à 50 %, mais était moins bon pour des niveaux budgétaires inférieurs à 50 %. Cette évaluation a en outre montré que lorsque nous ne tenions pas compte de la saisonnalité, le rendement d'un PEA était clairement amélioré au début de l'accumulation des données. Ensuite, cette tendance ralentissait, voire disparaissait, malgré l'augmentation des données; le comportement étant pratiquement semblable au modèle avec saison, et laissait entendre, par conséquent, que la saisonnalité n'améliorait pas davantage l'exactitude des prédictions et le rendement du PEA.

Notre étude comporte des limites exigeant de plus amples recherches et la reproduction dans d'autres enquêtes multimodales.

Ignorer les effets saisonniers propres au mode a été notre première simplification. Cela a simplifié la complexité de la spécification des modèles, mais a mené à apparemment compenser les effets saisonniers sur les prédictions de propensions. Nous pensons, toutefois, qu'il est possible d'adapter de manière pratique au modèle ajusté les effets saisonniers propres à chaque modèle, si la saisonnalité s'avère un prédicteur puissant des prédictions de propensions.

Pour assurer l'exactitude et la fiabilité des prédictions, nous avons uniquement envisagé deux stratégies de collecte de données candidates (uniquement en ligne et en ligne suivi d'interviews en personne) comme deuxième simplification. Le nombre de visites d'IPAO aux unités d'échantillonnage peut être encore personnalisé et l'optimisation peut comprendre le nombre réel de visites. Les propensions à répondre après chaque visite peuvent être modélisées et estimées simultanément, et l'on pourrait supposer que les prédictions d'un mode de suivi sont corrélées uniquement à son prédécesseur immédiat. Une telle application est simple, mais fait intervenir une vérification attentive de la fiabilité des prédictions.

La troisième limite a été la sensibilité de notre modèle de propension aux variations structurelles du plan (voir Wu et coll., 2023 pour une analyse plus approfondie). Par conséquent, dans l'étude de cas de la présente étude, nous avons noté une erreur de spécification temporaire dans les répartitions *a priori* des propensions à répondre. Dans un effort de caractérisation d'une variation imprévue, il est important de déterminer les paramètres (et strates) touchés. La robustesse peut alors être améliorée par des paramètres de modèle hiérarchiques supplémentaires. Nous reportons cette extension à des études ultérieures.

Notre dernière limite était que, dans nos modèles hiérarchiques de série chronologique, nous avons supposé que les strates du PEA étaient précisées d'avance. Nous n'avons pas intégré d'étape de sélection de variables dans l'ajustement et l'optimisation du modèle. Lorsque le nombre de variables auxiliaires est important, un compromis intervient entre le temps d'entraînement de l'analyse bayésienne des propensions à répondre et l'utilité de l'optimisation; il s'agit d'un sujet important pour des études ultérieures.

L'inclusion dans les modèles hiérarchiques de série chronologique de paradonnées et d'autres covariables variant dans le temps serait une autre extension pertinente. Cela permettrait l'optimisation dans un

cadre dynamique, c'est-à-dire au cours du travail sur le terrain. Dans le cadre d'une interview sur place, une telle approche dynamique n'est pas simple du point de vue opérationnel. Les charges de travail des intervieweurs sont uniquement connues à un moment proche du travail sur le terrain et à des moments définis, souvent mensuels. Ce serait également le cas, dans une moindre mesure, pour un suivi téléphonique. Une solution pratique, appliquée par le bureau central de la statistique des Pays-Bas, est d'ajouter un sous-échantillonnage aléatoire de non-répondants. Les probabilités de sous-échantillonnage dépendent des charges de travail prédéterminées et fixes dans les régions des intervieweurs. On peut également envisager une étape supplémentaire permettant à un plan dynamique d'intervenir lors du suivi en personne. Afin de laisser une certaine liberté aux intervieweurs, cela peut être mis en œuvre sous la forme d'un ou de plusieurs points de liaisons prédéfinis d'arrêt ou non des unités d'échantillonnage. En considérant en plus les étapes de collecte de données générales (séquentielles), il serait utile d'étendre la méthodologie de la présente étude. Une telle extension fait face à trois défis statistiques. Le premier est l'inclusion de nouvelles données auxiliaires d'intrants (par exemple les paradonnées) dans le modèle et l'optimisation. La décision et la façon d'inclure ces données seraient un compromis entre l'efficacité et l'efficience. Le deuxième est qu'il est nécessaire de mettre davantage l'accent sur la spécification correcte des covariances des propensions à répondre des différentes étapes de collecte de données. Les erreurs de spécifications prolifèrent dans les étapes suivantes. Le troisième est que l'optimisation doit être effectuée pendant la collecte de données, ce qui exige de limiter ses ambitions.

Annexe A

Tableau A.1
Les variables auxiliaires forment 13 strates et la saisonnalité est jugée être un facteur influant sur la prédiction des propensions à répondre.

Strate	Âge	Ethnicité
1	0 à 17 ans	Occidentaux
2	18 à 24 ans	Occidentaux
3	25 à 34 ans	Occidentaux
4	35 à 54 ans	Occidentaux
5	55 à 64 ans	Occidentaux
6	65 à 74 ans	Occidentaux
7	75 ans et plus	Occidentaux
8	0 à 17 ans	Non occidentaux
9	18 à 24 ans	Non occidentaux
10	25 à 34 ans	Non occidentaux
11	35 à 54 ans	Non occidentaux
12	55 à 64 ans	Non occidentaux
13	65 ans et plus	Non occidentaux

Annexe B

Tableau B.1
Observations des coefficients de variation tenant compte du biais (corrigés) ou non (non corrigés) et des erreurs-types (ET) lors de la prise en compte du biais au premier trimestre 2017 pour différents niveaux budgétaires.

	W			w-Np			w-Ap		
	Non corrigés	Corrigés	ET	Non corrigés	Corrigés	ET	Non corrigés	Corrigés	ET
0 %*	0,305	0,294	0,023	-	-	-	-	-	-
100 %	-	-	-	0,156	0,149	0,013	0,102	0,072	0,021
90 %	-	-	-	-	-	-	0,093	0,084	0,021
80 %	-	-	-	-	-	-	0,104	0,074	0,020
70 %	-	-	-	-	-	-	0,092	0,078	0,021
60 %	-	-	-	-	-	-	0,093	0,084	0,021
50 %	-	-	-	-	-	-	0,106	0,052	0,020
40 %	-	-	-	-	-	-	0,114	0,072	0,021
30 %	-	-	-	-	-	-	0,136	0,115	0,022
20 %	-	-	-	-	-	-	0,259	0,237	0,022
10 %	-	-	-	-	-	-	0,305	0,294	0,023

**Aucun budget indique seulement que le mode IWAO est utilisé. « - » désigne aucun résultat*. w désigne les plans par IWAO uniquement, w-Ap désigne les plans non adaptatifs, w-Np désigne les plans adaptatifs.

Tableau B.2
Observations de c.v. tenant compte du biais (corrigés) ou non (non corrigés) et erreurs-types (ET) lors de la prise en compte du biais pour le niveau budgétaire de 100 % pour chaque trimestre.

Année	Trimestre	w			w-Np			w-Ap		
		Non corrigés	Corrigés	ET	Non corrigés	Corrigés	ET	Non corrigés	Corrigés	ET
2016	Q1				-	-	-	-	-	-
	Q2	-	-	-						
	Q3	-	-	-	-	-	-	0,093	0,084	0,021
	Q4	-	-	-	-	-	-	0,104	0,074	0,020
2017	Q1	-	-	-	-	-	-	0,092	0,078	0,021
	Q2	-	-	-	-	-	-	0,093	0,084	0,021
	Q3	-	-	-	-	-	-	0,106	0,052	0,020
	Q4	-	-	-	-	-	-	0,114	0,072	0,021

Note : w désigne les plans par IWAO uniquement, w-Ap désigne les plans non adaptatifs, w-Np désigne les plans adaptatifs.

Tableau B.3
Coût relatif (c) du premier trimestre 2017 pour différents niveaux budgétaires pour des enquêtes adaptatives.

	100 %	90 %	80 %	70 %	60 %	50 %	40 %	30 %	20 %	10 %
c	0,425	0,473	0,532	0,608	0,709	0,851	0,983	1,311	1,966	0,977
C	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	VRAI	FAUX	FAUX	FAUX

Note : Les répartitions dans chaque cas sont déterminées par le résolveur d'optimisation « auglag » en commençant au point initial 0 défini pour les 13 strates. Si la convergence (C) indique VRAI, il est possible de trouver l'optimum local et de produire les répartitions correspondantes; autrement, le processus mène à un faux « optimum » local.

Bibliographie

Boonstra, H.-J. (2022). *mcmcsae*: Markov Chain Monte Carlo small area estimation. R package version 0.7.2.

Boonstra, H.-J., et van den Brakel, J.A. (2019). Estimation du niveau et de la variation du chômage au moyen de modèles de séries chronologiques structurels. *Techniques d'enquête*, 45, 3, 421-454. Article accessible à l'adresse https://www150.statcan.gc.ca/n1/pub/12-001-x/2019003/article/00005-fra.pdf.

Calinescu, M., Bhulai, S. et Schouten, B. (2013). Optimal resource allocation in survey designs. *European Journal of Operational Research*, 226(1), 115-121. DOI: https://doi.org/10.1016/j.ejor.2012.10.046.

Chen, Q., Elliott, M.R. et Little, R.J.A. (2010). Inférence basée sur un modèle bayésien avec splines pénalisées pour les proportions de population finie dans l'échantillonnage avec probabilités inégales. *Techniques d'enquête*, 36, 1, 25-37. Article accessible à l'adresse https://www150.statcan.gc.ca/n1/fr/pub/12-001-x/2010001/article/11250-fra.pdf.

Chun, A.Y., Heeringa, S.G. et Schouten, B. (2018). Responsive and adaptive design for survey optimization. *Journal of Official Statistics*, 34(3), 581-597. DOI: https://doi.org/10.2478/jos-2018-0028.

Coffey, S., West, B.T., Wagner, J. et Elliott, M.R. (2020). What do you think? Using expert opinion to improve predictions of response propensity under a Bayesian framework. *Methods, Data, Analyses*. DOI: https://doi.org/10.12758/mda.2020.05.

Gelman, A., et Hill, J. (2007). Data analysis using regression and multilevel/hierarchical models. Cambridge University Press.

Gelman, A., et Rubin, D.B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457-472.

Linderman, S., Johnson, M.J. et Adams, R.P. (2015). Dependent multinomial models made easy: Stick-breaking with the polya-gamma augmentation. *Advances in Neural Information Processing Systems*, 28. DOI: https://doi.org/10.48550/arXiv.1506.05843.

Ma, Y. (2021). *Optimal Stratification in Bayesian Adaptive Survey Designs*. Thèse de doctorat, University Utrecht, Les Pays-Bas.

Moore, J.C., Durrant, G.B. et Smith, P.W.F. (2018). Data set representativeness during data collection in three UK social surveys: Generalizability and the effects of auxiliary covariate choice. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 181(1), 229-248. DOI: https://doi.org/10.1111/rssa.12256.

Nishimura, R., Wagner, J. et Elliott, M. (2016). Alternative indicators for the risk of non-response bias: A simulation study. *Revue Internationale de Statistique*, 84(1), 43-62. DOI: https://doi.org/10.1111/INSR.12100.

O'Malley, A.J., et Zaslavsky, A.M. (2008). Domain-level covariance analysis for multilevel survey data with structured nonresponse. *Journal of the American Statistical Association*, 103(484), 1405-1418. DOI: https://doi.org/10.1198/016214508000000724.

Polson, N.G., Scott, J.G. et Windle, J. (2013). Bayesian inference for logistic models using Pólya-gamma latent variables. *Journal of the American Statistical Association*, 108, 1339-1349.

Peytchev, A., Riley, S., Rosen, J., Murphy, J. et Lindblad, M. (2010). Reduction of nonresponse bias through case prioritization. *Survey Research Methods*, 4(1), 21-29. DOI: https://doi.org/10.18148/SRM/2010.V4I1.3037.

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. Article accessible à l'adresse https://www150.statcan.gc.ca/n1/pub/12-001-x/2010002/article/11376-fra.pdf.

Schouten, B., Calinescu, M. et Luiten, A. (2013). Optimiser la qualité de la réponse au moyen de plans de collecte adaptatifs. *Techniques d'enquête*, 39, 1, 33-66. Article accessible à l'adresse https://www150.statcan.gc.ca/n1/pub/12-001-x/2013001/article/11824-fra.pdf.

Schouten, B., Cobben, F. et Bethlehem, J. (2009). Indicateurs de la représentativité de la réponse aux enquêtes. *Techniques d'enquête*, 35, 1, 107-121. Article accessible à l'adresse https://www150.statcan.gc.ca/n1/fr/pub/12-001-x/2009001/article/10887-fra.pdf.

Schouten, B., Mushkudiani, N., Shlomo, N., Durrant, G., Lundquist, P. et Wagner, J. (2018). A Bayesian analysis of design parameters in survey data collection. *Journal of Survey Statistics and Methodology*. DOI: https://doi.org/10.1093/jssam/smy012.

Schouten, B., Peytchev, A. et Wagner, J. (2017). *Adaptive Survey Design*. Chapman and Hall/CRC.

Scott, A., et Smith, T.M. (1969). Estimation in multi-stage surveys. *Journal of the American Statistical Association*, 64(327), 830-840.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P. et van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), 583-639. DOI: https://doi.org/10.1111/1467-9868.00353.

van Berkel, K., van der Doef, S. et Schouten, B. (2020). Implementing adaptive survey design with an application to the Dutch Health Survey. *Journal of Official Statistics*, 36(3), 609-629. DOI: https://doi.org/10.2478/jos-2020-0031.

Varadhan, R. (2022). *alabama: Constrained Nonlinear Optimization*. R package version 2022.4-1.

Vehtari, A., Gelman, A. et Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27, 1413-1432. DOI: https://doi.org/10.1007/s11222-016-9696-4.

Wagner, J. (2013). Adaptive contact strategies in telephone and face-to-face surveys. *Survey Research Methods*, 7(1), 45-55. DOI: https://doi.org/10.18148/SRM/2013.V7I1.5037.

Wagner, J., et Hubbard, F. (2013). Using propensity models during data collection for responsive designs: Issues with estimation. Lors de la 68e conference de l'AAPOR, mai (pp. 16-19).

Wagner, J.R. (2008). *Adaptive Survey Design to Reduce Nonresponse Bias*. (Thèse de doctorat, University of Michigan).

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11, 3571-3594. https://arxiv.org/abs/1004.2316.

Watanabe, S. (2013). A widely applicable Bayesian information criterion. *Journal of Machine Learning Research*, 14, 867-897. https://arxiv.org/abs/1208.6338.

West, B.T., Wagner, J., Coffey, S. et Elliott, M.R. (2023). Deriving priors for Bayesian prediction of daily response propensity in responsive survey design: Historical data analysis vs. literature review. *Journal of Survey Statistics and Methodology,* 11(2), 367-392.

Wu, S., Boonstra, H.-J., Moerbeek, M. et Schouten, B. (2023). Modélisation de la variation temporelle des taux de réponse aux enquêtes : approche bayésienne s'appliquant à l'enquête sur la santé réalisée aux Pays-Bas. *Techniques d'enquête*, 49, 1, 177-208. Article accessible à l'adresse https://www150.statcan.gc.ca/n1/fr/pub/12-001-x/2023001/article/00010-fra.pdf.

Wu, S., Schouten, B., Meijers, R. et Moerbeek, M. (2022). Data collection expert prior elicitation in survey design: Two case studies. *Journal of Official Statistics*, 38(2), 637-662. DOI: https://doi.org/10.2478/JOS-2022-0028.