

N° 12-001-X au catalogue  
ISSN 1712-5685

## Techniques d'enquête

# Tenir compte des effets de l'intervieweur et du plan de sondage dans la planification des tailles d'échantillon

par Stefan Zins et Jan Pablo Burgard

Date de diffusion : le 30 juin 2020



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](http://www.statcan.gc.ca).

Vous pouvez également communiquer avec nous par :

**Courriel** à [STATCAN.infostats-infostats.STATCAN@canada.ca](mailto:STATCAN.infostats-infostats.STATCAN@canada.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 |

### Programme des services de dépôt

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

## Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site [www.statcan.gc.ca](http://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 la Reine du chef du Canada, représentée par le ministre de l'Industrie 2020

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

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

*This publication is also available in English.*

---

# Tenir compte des effets de l'intervieweur et du plan de sondage dans la planification des tailles d'échantillon

Stefan Zins et Jan Pablo Burgard<sup>1</sup>

## Résumé

La sélection d'une taille d'échantillon adéquate est essentielle à la qualité d'une enquête. Les techniques les plus modernes consistent à tenir compte de plans de sondage complexes en calculant les tailles d'échantillon efficaces. Ces tailles d'échantillon efficaces sont déterminées à l'aide de l'effet de plan des variables d'intérêt centrales. Toutefois, dans les enquêtes en personne, il est souvent suspecté que les estimations empiriques des effets de plan sont confondues avec l'effet des intervieweurs. Parce que cela tend à mener à une surestimation des effets de plan, le risque est souvent d'attribuer incorrectement les ressources à une taille d'échantillon plus élevée au lieu d'utiliser plus d'intervieweurs ou d'améliorer l'exactitude des mesures. C'est pourquoi nous proposons un effet de plan corrigé qui sépare l'effet de l'intervieweur (dit aussi *effet enquêteur*) des effets du plan de sondage sur la variance due à l'échantillonnage. La capacité d'estimer l'effet de plan corrigé est mise à l'essai au moyen d'une étude par simulations. Dans ce contexte, nous abordons la question de la différenciation entre la variance due aux grappes et celle due aux intervieweurs. Les effets corrigés du plan sont estimés pour les données de la 6<sup>e</sup> édition de l'Enquête sociale européenne (ESS) et comparés aux estimations classiques de l'effet de plan. En outre, nous montrons que pour certains pays étudiés dans la 6<sup>e</sup> édition de l'ESS, les estimations de l'effet de plan classique sont effectivement fortement augmentées par les effets de l'intervieweur.

**Mots-clés :** Effet de plan; effet de l'intervieweur; modèle multiniveau; taille d'échantillon; Enquête sociale européenne (ESS).

## 1 Introduction

Il est parfois très difficile de déterminer la taille de l'échantillon d'une enquête. La complexité de la tâche est souvent accrue par le manque d'information et de données permettant de planifier l'enquête. C'est pourquoi ceux qui planifient les enquêtes cherchent à réduire la complexité de ce problème par des simplifications et des modèles statistiques. L'une de ces approches consiste à sélectionner une taille d'échantillon au moyen de ce qu'on appelle l'*effet de plan*. L'effet de plan est ensuite défini comme le ratio entre la variance d'un estimateur selon le plan de sondage de l'enquête planifiée et la variance du même estimateur selon un plan d'échantillonnage aléatoire simple. Par conséquent, l'effet de plan est une propriété d'une stratégie d'estimation – c'est-à-dire un plan de sondage et un estimateur (Chaudhuri et Stenger, 2005, page 4) – et non de l'enquête. Habituellement, on utilise la moyenne pondérée de l'échantillon d'une seule variable comme estimateur de référence. Toutefois, par souci de simplification, dans la suite de l'article, chaque fois que nous parlerons de l'effet de plan d'un plan de sondage, nous le ferons toujours en ce qui concerne la variance due à l'échantillonnage d'une moyenne d'échantillon pondérée.

Pour planifier la taille de l'échantillon, on peut établir une cible de taille d'échantillon efficace, ce qui signifie que la taille d'échantillon planifiée divisée par l'effet de plan planifié doit être supérieure à une certaine valeur. La taille d'échantillon efficace d'un plan de sondage est l'équivalent sous échantillonnage

---

1. Stefan Zins, Institut pour la recherche d'emploi de l'Agence fédérale allemande pour l'emploi (BA), Regensburger Strasse 104, D-90478, Nuremberg. Courriel : st.zins@gmail.com; Jan Pablo Burgard, RIFOSS - Research Institute for Official and Survey Statistics, Université de Trèves, D-54286 Trier.

aléatoire simple de sa taille d'échantillon, en termes d'efficacité, c'est-à-dire que si un plan de sondage a une taille d'échantillon efficace de 1 000, alors sa variance due à l'échantillonnage est égale à celle d'un échantillonnage aléatoire simple de 1 000.

Idéalement, la personne planifiant l'enquête la conçoit en ayant à l'esprit une analyse ou un test d'hypothèse précis et formule son avis sur les niveaux d'erreur d'échantillonnage tolérables ou les probabilités d'erreur de type II. Cet avis doit se fonder sur deux éléments : premièrement, un certain degré d'expérience à propos la question de recherche importante et, deuxièmement, des hypothèses sur les paramètres de la population cible nécessaires à la planification des erreurs d'échantillonnage et aux calculs de puissance. Les hypothèses sur les paramètres de la population cible peuvent provenir des cycles précédents d'une enquête ou être fondées sur les données collectées pendant l'essai sur le terrain de l'enquête. Les calculs de puissance et la planification des erreurs d'échantillonnage sont nettement moins complexes et nécessitent moins d'information sur la population cible si l'on suppose un plan d'échantillonnage aléatoire simple. C'est pourquoi la plupart des méthodes de planification de la taille de l'échantillon proposées dans les manuels conviennent à la détermination d'une taille d'échantillon efficace. On prend ensuite compte de l'effet de l'échantillonnage complexe en multipliant la taille d'échantillon efficace prévue par un effet de plan prévu. La détermination de l'effet de plan peut ainsi être séparée de la sélection d'une taille d'échantillon efficace. Par exemple, si un échantillon aléatoire simple de taille 1 000 garantit les éléments suivants : l'erreur d'échantillonnage d'un estimateur ne dépasse pas une valeur donnée avec une probabilité de 95 %, ou que la puissance d'un test statistique est de 80 %, c'est-à-dire la probabilité de rejeter une hypothèse nulle au cas où l'alternative est vraie devrait être de 80 % (Ellis, 2010, chapitre 3). Alors multiplier 1 000 par l'effet de plan présumé d'une étude donnera au planificateur de l'enquête la taille d'échantillon nette requise pour atteindre les objectifs de précision fixés.

La décision relative à la taille d'échantillon efficace doit également représenter un compromis entre le coût de l'enquête et la précision des estimations de l'enquête. En ce qui concerne ce compromis, la personne planifiant l'enquête devrait, par exemple, tenir compte des conséquences en cas d'erreur de type II, c'est-à-dire si une hypothèse nulle n'est pas rejetée même si l'hypothèse alternative est vraie.

Dans les enquêtes devant principalement servir à des analyses secondaires, c'est-à-dire qu'elles fournissent des données aux chercheurs sans viser une seule application, comme l'Enquête sociale européenne (ESS) ou l'*European Value Study* (EVS, enquête européenne sur les valeurs), on ne peut pas décider d'une taille d'échantillon efficace qui serait prévue pour une seule question de recherche ou un test d'hypothèse. C'est pourquoi l'ESS emploie une taille d'échantillon efficace moyenne. Cela signifie que les plans de sondage de l'ESS sont planifiés de telle sorte que l'effet de plan moyen pour un ensemble d'éléments du questionnaire principal de l'ESS ait une certaine valeur. On calcule la taille d'échantillon nette prévue en multipliant l'effet de plan de sondage moyen prévu par la taille d'échantillon efficace moyenne requise. La taille d'échantillon nette est la taille de l'échantillon après la non-réponse totale, c'est-à-dire le nombre d'interviews terminées. Pour planifier la taille brute de l'échantillon – à savoir la taille de l'échantillon avant la non-réponse totale – on divise la taille d'échantillon nette par le produit du taux de réponse et du taux d'admissibilité supposés. Le taux d'admissibilité est la fraction des personnes

échantillonnées appartenant à la population cible, qui peut être inférieure à 100 % en raison des imperfections de la base de sondage.

Il reste que les effets du plan peuvent tout de même être difficiles à quantifier, en raison de la complexité du plan de sondage. En vue de réduire cette complexité, on utilise des modèles statistiques pour les données d'enquête afin d'obtenir une approximation de l'effet de plan. Ces modèles tentent généralement d'intégrer l'effet de l'échantillonnage en grappes, qui peut avoir un effet important sur la variance due à l'échantillonnage des estimations. Les grappes peuvent être des zones spatiales comme des peuplements, des unités organisationnelles comme des municipalités ou des établissements comme des hôpitaux et des écoles. Elles sont soit utilisées comme *unités primaires d'échantillonnage* (UPE) – qui sont sélectionnées en premier, puis à l'intérieur desquelles un échantillonnage supplémentaire est effectué – soit enquêtées dans leur totalité. À titre d'exemple, le plan de sondage de la 6<sup>e</sup> édition de l'ESS allemande (ESS6) comporte deux étapes d'échantillonnage. En effet, les UPE sont des municipalités et les unités d'échantillonnage secondaires sont les personnes inscrites dans les municipalités. Souvent, on ne peut pas considérer les variables d'intérêt comme étant distribuées identiquement dans toutes les grappes de la population. En fait, on peut supposer que les répondants d'une même grappe sont habituellement plus semblables les uns aux autres que ceux appartenant à une autre grappe. Kish (1965, page 162) donne la formule suivante pour un effet de plan dû à la mise en grappe :

$$deff = 1 + (b - 1) \rho. \quad (1.1)$$

Cet effet de plan  $deff$  se compose de deux paramètres :  $b$  qui est généralement une taille moyenne de grappe en terme de répondants réalisés et  $\rho$ , le coefficient de corrélation intra-grappe, qui est une mesure de l'homogénéité des mesures d'une variable dans une même grappe.  $\rho$  peut être défini au moyen de la décomposition de la variance comme étant la variance entre les grappes divisée par la somme des variances intra-grappe et inter-grappes. Plus la variance entre les grappes est élevée, plus  $\rho$  est élevé.

Pour utiliser l'effet de plan lors de la sélection d'une taille d'échantillon, il faut faire des hypothèses sur le paramètre inconnu  $\rho$ . La taille de la grappe  $b$  ne dépend pas de la variable mesurée et peut être influencée par la personne planifiant l'enquête. Pour  $\rho$ , on peut utiliser les données des enquêtes précédentes pour formuler l'hypothèse nécessaire. Particulièrement dans les enquêtes transversales répétées, les données accumulées sont très utiles pour la planification du plan d'échantillonnage en vue de la prochaine mise en œuvre de l'enquête.

Lynn, Häder, Gabler et Laaksonen (2007) décrivent comment les effets de plan prédits sont utilisés par l'ESS aux fins de planification des tailles d'échantillon de façon à obtenir une certaine taille d'échantillon efficace moyenne selon un plan de sondage donné. Dans les éditions récentes de l'ESS, la prédiction de l'effet de plan et de ses composantes a été éclairée par des estimations de ces statistiques fondées sur les données des éditions précédentes de l'ESS (The ESS Sampling Expert Panel, 2016).

Un autre facteur important peut apporter de l'homogénéité aux mesures dans les enquêtes en personne : l'intervieweur. Dans le cadre de l'erreur d'enquête totale (EET) (Groves, 2009), différents mécanismes ont été décrits pour expliquer la façon dont un intervieweur peut influencer les mesures d'une enquête.

Comme l'échantillonnage en grappes, les intervieweurs sont depuis longtemps considérés comme une source de mesures dépendantes (Kish, 1965, page 522, Kish, 1962). Ils introduisent une homogénéité par des erreurs de mesure et des effets de sélection, plutôt que l'homogénéité des grappes intrinsèque à la population. West et Blom (2017) présentent une vue d'ensemble des recherches sur les effets de l'intervieweur. Ils expliquent en détail comment les tâches de l'intervieweur, comme la production ou l'application de bases de sondage, la prise de contact avec les répondants et l'obtention de leur collaboration et de leur consentement, peuvent avoir un effet de sélection sur le recrutement des répondants. West et Blom (2017) présentent également des preuves selon lesquelles les intervieweurs qui effectuent des mesures, font des observations et, enfin, enregistrent l'information recueillie peuvent introduire des erreurs de mesure et de traitement dans les données utilisées aux fins d'analyse. Pour un aperçu des autres sources de variance dans les enquêtes, nous faisons référence au cadre de l'EET décrit notamment par Groves et Lyberg (2010) et Biemer (2010).

L'analyse des effets de l'intervieweur à l'aide des données de l'ESS issues de différents pays et années a montré que cet effet peut être considérable (Beullens et Loosveldt, 2016). Ces constatations soulèvent la question suivante : dans quelle mesure la valeur  $\rho$  de l'équation (1.1) est-elle attribuable à la corrélation intra-grappe plutôt qu'à la corrélation intra-intervieweur ? Schnell et Kreuter (2005) montrent que l'effet de l'intervieweur peut être plus élevé que l'effet de grappe, y compris pour les variables où une forte corrélation spatiale peut être supposée. Par conséquent, l'effet de plan estimé dans les enquêtes en personne est généralement confondu avec l'effet de l'intervieweur. C'est pourquoi l'effet de plan est systématiquement surestimé dans les enquêtes en personne. Ce risque d'attribution incorrecte des fonds pourrait poser problème dans les enquêtes qui prédisent les effets de plan à partir de données historiques pour planifier la taille des échantillons. La personne planifiant l'enquête pourrait essayer de compenser l'augmentation de l'effet de plan prédit en augmentant la taille de l'échantillon pour maintenir constante la taille d'échantillon efficace. Si le facteur déterminant l'augmentation de l'effet de plan prédit est l'effet de l'intervieweur, on pourrait attribuer les fonds plus efficacement en embauchant d'autres intervieweurs ou en améliorant leur formation de façon à accroître l'exactitude des mesures et réduire les effets de la sélection.

La nouveauté apportée par la méthode proposée ici est qu'elle permet d'estimer un effet de plan corrigé non confondu avec les effets de l'intervieweur. Au moyen de l'effet de plan corrigé proposé, la personne ou l'entité planifiant l'enquête est en mesure de prendre des décisions fondées sur des données probantes à propos des modifications du plan de sondage, comme la taille de l'échantillon et le nombre d'UPE, ou encore le déploiement des intervieweurs.

L'article est organisé comme suit. La section 2 présente le cadre pour décrire les effets du plan de sondage et de l'intervieweur. Le cadre suit la justification fondée sur un modèle de l'effet de plan qu'ont décrite Gabler, Häder et Lahiri (1999) et l'introduction d'un effet de l'intervieweur dans ce cadre étudiée par Gabler et Lahiri (2009). Les modèles de mesure servant à décrire les données observées suivent une structure multiniveau. L'influence de l'échantillonnage à plusieurs degrés ou en grappes et celle des intervieweurs sur les données observées sont modélisées à l'aide d'effets aléatoires qui impliquent une

certaine structure de variance-covariance. Cette méthode permet la factorisation de l'effet global en des effets distincts de plan de sondage et d'intervieweur. Cette séparation est essentielle quand on traite les effets séparément afin de les contrôler.

À la section 3, les effets de l'échantillonnage et de l'intervieweur décrits à la section 2 sont estimés pour les données de l'ESS6 à l'aide de modèles multiniveaux. Tout d'abord, nous présentons les résultats d'une étude par simulations qui visait à évaluer la possibilité de distinguer les variances de grappe et d'intervieweur pour la structure UPE-intervieweur observée dans les données de l'ESS6. Ensuite, nous évaluons l'applicabilité des différents modèles de mesure pour un ensemble sélectionné de variables de l'ESS. Les modèles sélectionnés servent à estimer les variances de différents effets aléatoires dans des modèles multiniveaux, qui sont à leur tour utilisés aux fins d'estimation de la corrélation intra-UPE et intra-intervieweur.

À la section 4, nous présentons nos conclusions et nous formulons des recommandations à l'intention des planificateurs d'enquête à partir du travail théorique de la section 2 et des constatations empiriques de la section 3. Nous indiquons ensuite les recherches futures possibles qui adapteraient nos modèles de mesures relativement simplistes de façon à mieux prendre en compte les plans de sondage complexes et l'hétérogénéité des intervieweurs.

## 2 Effets de l'intervieweur et de plan

Nous définissons un échantillon comme un ensemble de  $n$  répondants distincts, que nous notons  $s = \{1, \dots, n\}$ , avec  $n \in \mathbb{N}$ . Pour le répondant  $k^e$  notre variable d'intérêt  $y$  est une variable à valeur réelle, où  $y_k$  est l'observation de cette variable pour le répondant  $k^e$  dans notre échantillon  $s$ . Les données observées sont données par  $\mathbf{y} = (y_1, \dots, y_n)^\top$ . Nous associons les poids d'enquête à chaque répondant de l'échantillon, donnés par  $\mathbf{w} = (w_1, \dots, w_n)^\top$ , où  $w_k$  est le poids du répondant  $k^e$  et  $w_k > 0$ , pour tous les  $k \in s$ .

Nous considérons la moyenne pondérée de l'échantillon  $\mathbf{y}$  comme notre estimateur, qui s'écrit

$$\bar{y}(\mathbf{w}) = \frac{\mathbf{w}^\top \mathbf{y}}{\mathbf{w}^\top \mathbf{I}_n} = \frac{\sum_{k \in s} w_k y_k}{\sum_{k \in s} w_k}, \quad (2.1)$$

où  $\mathbf{I}_n$  est un vecteur-colonne de un de longueur  $n$ . Nous nous intéressons à un estimateur d'intérêt,  $\bar{y}(\mathbf{w})$ , qui est l'estimateur le plus couramment choisi pour décrire les effets de l'intervieweur et du plan (Kish, 1965, section 8.1, Kish, 1962; Särndal, Swensson et Wretman, 1992, page 53). Ce choix nous permet d'utiliser un cadre établi (Gabler et coll., 1999) et de produire des formules reconnaissables par les lecteurs connaissant quelque peu le sujet. Cependant, les effets de plan d'autres estimateurs ont été étudiés, notamment par Lohr (2014), qui calcule les effets de plan pour les estimateurs des coefficients de régression, et par Fischer, West, Elliott et Kreuter (2018), qui décrivent l'incidence des effets de l'intervieweur sur l'estimation des coefficients de régression.

Dans ce qui suit, la variance de  $\bar{y}(\mathbf{w})$  est dérivée selon différents modèles de mesure pour  $y$ . Les différents modèles servent à faire la distinction entre les plans d'échantillonnage complexes et simples, ainsi qu'à distinguer les situations comportant un effet de l'intervieweur et celles n'en comportant pas. Notons que la variance fondée sur un modèle de l'estimateur  $\bar{y}(\mathbf{w})$ , que nous utilisons, n'est généralement pas identique aux variances fondées sur le plan de sondage, c'est-à-dire la variance de  $\bar{y}(\mathbf{w})$  selon un plan de sondage donné (Särndal et coll., 1992, page 492). Les variances fondées sur un plan peuvent être très complexes et donc difficiles à afficher de façon accessible, surtout en cas d'échantillonnage à plusieurs degrés. L'approche fondée sur un modèle réduit la complexité tout en conservant la propriété essentielle des plans de sondage complexes que nous étudions, l'effet de grappe d'un échantillonnage à plusieurs degrés. Il permet également d'intégrer facilement l'effet de grappe et de l'intervieweur dans un même cadre.

## 2.1 Échantillonnage aléatoire simple sans effet de l'intervieweur

Pour modéliser un échantillonnage aléatoire simple en l'absence d'effet de l'intervieweur, c'est-à-dire sans corrélation intra-UPE et intra-intervieweur, nous supposons le modèle de mesure suivant ( $M_0$ )

$$y_k = \mu_k + e_k, \quad (M_0)$$

où  $\mu_k$  est la valeur de  $y$  pour le répondant  $k^e$  et  $e_k$  est l'erreur de mesure. Les erreurs de mesure  $e_k$  pour tous les  $k \in s$  sont des variables aléatoires indépendantes et identiquement distribuées (iid) avec une structure variance-covariance de

$$\text{Cov}_{M_0}(e_k, e_l) = \begin{cases} \sigma^2, & \text{si } k = l \\ 0, & \text{sinon} \end{cases}, \quad (2.2)$$

où  $\sigma$  est un paramètre de valeur réelle supérieur à zéro. Selon le modèle ( $M_0$ ), la variance de  $\bar{y}(I_n)$  est obtenue par  $V_{M_0}(\bar{y}(I_n)) = \sigma^2/n$ . Cette variance peut être interprétée comme la variance de la moyenne non pondérée de l'échantillon selon un échantillonnage aléatoire simple avec remise (Särndal et coll., 1992, page 73). L'échantillonnage aléatoire simple avec estimateur  $\bar{y}(I_n)$  sert habituellement de stratégie d'estimation de référence, qui est comparée avec des plans de sondage et des estimateurs plus complexes.

## 2.2 Échantillonnage aléatoire simple avec effet de l'intervieweur

Ensuite, nous introduisons la variance de l'intervieweur dans notre modèle de mesure pour  $y$ . Chaque répondant est interviewé par un intervieweur seulement. On a  $R \in \mathbb{N}_{>0}$ , intervieweurs réalisant les interviews de tous les répondants  $n$ . Nous désignons par  $s_i \subset s$  l'ensemble de tous les répondants qui sont interviewés par l'intervieweur  $i^e$  et  $\mathcal{R} = \{1, \dots, R\}$  l'ensemble de tous les intervieweurs. La charge de travail de l'intervieweur  $i^e$  est donnée par  $n_i$ ,  $\mathbf{n}_I = (n_1, \dots, n_R)^\top$  est le vecteur des charges de travail des intervieweurs et  $\sum_{i=1}^R n_i = n$ . Selon le modèle de mesure ( $M_1$ ), qui suit les explications de Särndal et coll. (1992), à la page 623, les valeurs observées de  $y$  pour  $k \in s_i$  sont décrites sous la forme

$$y_{ik} = \mu_k + \mathfrak{F}_i + e_{ik}, \quad (M_1)$$

$\mathfrak{F}_i$  étant l'effet de l'intervieweur associé à toutes les mesures effectuées pour les répondants  $k \in s_i$ .  $e_{ik}$  représentant l'erreur aléatoire attribuable à des sources autres que l'intervieweur. Tous les  $e_{ik}$  pour  $i \in \mathcal{R}$  et  $k \in s$  sont des variables iid avec une moyenne nulle et une variance  $\sigma_e^2$ .  $\mathfrak{F}_1, \dots, \mathfrak{F}_R$  sont des variables aléatoires iid avec moyenne nulle et une variance  $\sigma_i^2$ , que nous appelons variance de l'intervieweur, et elles sont indépendantes de  $e_{ik}$  pour tous les  $i \in \mathcal{R}$  et  $k \in s$ . Särndal et coll. (1992) interprètent le modèle ( $M_1$ ) comme une attribution aléatoire des intervieweurs à une partition prédéfinie de l'échantillon  $s$  dans  $R$  sous-ensembles disjoints  $s_i, i = 1, \dots, R$ . Ces sous-ensembles pourraient correspondre aux différentes régions géographiques où l'enquête est menée et auxquelles les intervieweurs sont alloués aléatoirement. En pratique, dans de nombreuses enquêtes, les organismes travaillant sur le terrain affectent les intervieweurs à des régions géographiques en fonction de leur expérience et de leur proximité. Comme ce processus n'est pas nécessairement observable par le chercheur estimant l'effet de plan, nous supposons une répartition aléatoire des intervieweurs dans les UPE. Cela peut être perçu comme le recrutement d'intervieweurs à partir d'un bassin infini, ou très grand, d'intervieweurs possibles.

Si nous définissons la partie aléatoire dans  $y_{ik}$  comme étant  $\varepsilon_{ik} = \mathfrak{F}_i + e_{ik}$ , alors la structure variance-covariance de  $y_{ik}$  selon le modèle ( $M_1$ ) est donnée par

$$\text{Cov}_{M_1}(\varepsilon_{ik}, \varepsilon_{jl}) = \begin{cases} \sigma^2, & \text{si } i = j, k = l \\ \rho_I \sigma^2, & \text{si } i = j, k \neq l, \\ 0, & \text{sinon} \end{cases} \quad (2.3)$$

où  $\sigma_i^2 + \sigma_e^2 = \sigma^2$  et  $\rho_I = \frac{\sigma_i^2}{\sigma^2}$  est la corrélation entre deux observations différentes de  $y$  faites par un même intervieweur. Pour calculer la variance de  $\bar{y}(\mathbf{w})$  selon le modèle ( $M_1$ ), nous déterminons d'abord la variance de  $\sum_{i \in \mathcal{R}} \sum_{k \in s_i} w_{ik} y_{ik}$ , où  $w_{ik}$  et  $y_{ik}$  sont le poids d'enquête et l'observation pour le répondant  $k \in s_i$ , respectivement. Nous obtenons ainsi

$$\begin{aligned} \text{Var}_{M_1} \left( \sum_{i \in \mathcal{R}} \sum_{k \in s_i} w_{ik} y_{ik} \right) &= \sigma^2 \left( \sum_{i \in \mathcal{R}} \sum_{k \in s_i} w_{ik}^2 + \rho_I \sum_{i \in \mathcal{R}} \sum_{k \in s_i} \sum_{\substack{l \in s_i \\ l \neq k}} w_{ik} w_{il} \right) \\ &= \sigma^2 \left( \sum_{i \in \mathcal{R}} \sum_{k \in s_i} w_{ik}^2 + \rho_I \left[ \sum_{i \in \mathcal{R}} \left( \sum_{k \in s_i} w_{ik} \right)^2 - \sum_{i \in \mathcal{R}} \sum_{k \in s_i} w_{ik}^2 \right] \right) \\ &= \sigma^2 \sum_{i \in \mathcal{R}} \sum_{k \in s_i} w_{ik}^2 \left( 1 + \rho_I \left[ \frac{\sum_{i \in \mathcal{R}} \left( \sum_{k \in s_i} w_{ik} \right)^2}{\sum_{i \in \mathcal{R}} \sum_{k \in s_i} w_{ik}^2} - 1 \right] \right), \end{aligned}$$

et il s'ensuit que

$$\text{Var}_{M_1}(\bar{y}(\mathbf{w})) = \frac{\sigma^2 \sum_{i \in \mathcal{R}} \sum_{k \in s_i} w_{ik}^2}{\left( \sum_{i \in \mathcal{R}} \sum_{k \in s_i} w_{ik} \right)^2} \left( 1 + \rho_I \left[ \frac{\sum_{i \in \mathcal{R}} \left( \sum_{k \in s_i} w_{ik} \right)^2}{\sum_{i \in \mathcal{R}} \sum_{k \in s_i} w_{ik}^2} - 1 \right] \right). \quad (2.4)$$

## 2.3 Échantillonnage à plusieurs degrés avec effet de l'intervieweur

Nous examinons un plan de sondage à deux degrés, où premièrement on sélectionne les UPE, puis deuxièmement les répondants à partir des UPE échantillonnées. Les UPE sont les unités de corrélation intra-grappe et nous traiterons les termes de grappe et d'UPE comme étant interchangeables. L'échantillon des UPE est noté  $\mathcal{K} = \{1, \dots, K\}$ , avec  $K > 1$ . Chaque répondant appartient à une UPE seulement. Soit  $s_q \subset s$  l'ensemble de tous les répondants appartenant à l'UPE  $q^e$ ,  $n_q$  le nombre de répondants observés dans l'UPE  $q^e$ ,  $\mathbf{n}_C = (n_1, \dots, n_K)^\top$  le vecteur des tailles de grappes, et  $\sum_{q \in \mathcal{K}} n_q = n$ . Encore une fois, chaque répondant est interviewé par un seul intervieweur. Les intervieweurs peuvent travailler dans plusieurs UPE et les UPE peuvent recevoir la visite de plusieurs intervieweurs. Bien que les intervieweurs puissent concentrer leur travail dans une région en particulier, généralement, les régions se composent de plusieurs UPE et les intervieweurs ne travaillent pas exclusivement dans une seule UPE. Cette situation se retrouve souvent dans les enquêtes en personne réalisées en Europe, comme dans l'ESS ou l'EVS. Le tableau 3.1 de la section 3.1 donne un aperçu du degré d'interpénétration entre les UPE et les intervieweurs dans les pays qui utilisent un plan de sondage à plusieurs degrés dans l'ESS6. L'interaction entre les UPE et l'intervieweur est observable dans toutes les éditions de l'ESS pour les pays utilisant un plan de sondage à plusieurs degrés.

Nous introduisons maintenant le modèle de mesure ( $M_2$ ), qui comprend à la fois la variance de grappe et de l'intervieweur dans les valeurs observées de  $y$ . Pour  $k \in s_{qi} = s_q \cap s_i$  nous modélisons les observations de  $y$  sous la forme

$$y_{qik} = \mu_k + \mathbb{C}_q + \mathfrak{I}_i + \epsilon_{qik}, \quad (M_2)$$

avec la valeur  $\mathbb{C}_q$  définie comme étant une variable aléatoire de moyenne nulle et de variance  $\sigma_C^2$ , que nous appelons variance de l'UPE, commune à tous les répondants de l'UPE  $q$ .  $\mathbb{C}_1, \dots, \mathbb{C}_K$  sont des variables aléatoires iid et sont indépendantes de  $\epsilon_{qik}$  et  $\mathfrak{I}_i$  pour tous les  $i \in \mathcal{R}$ ,  $q \in \mathcal{K}$ , et  $k \in s_{qi}$ .  $\mathbb{C}_q$  introduit un certain degré de similitude entre les répondants d'une même UPE. Cela permet un effet aléatoire permanent de l'UPE sur la mesure de  $y$  pour le répondant  $k^e$ , ce qui le fait s'écarter de  $\mu_k$  (Chambers et Skinner, 2003, page 201).

Pour établir l'effet de l'échantillonnage et des intervieweurs sur  $\bar{y}(\mathbf{w})$ , nous définissons la partie aléatoire de  $y_{qik}$  comme étant  $\epsilon_{qik} = \mathbb{C}_q + \mathfrak{I}_i + \epsilon_{qik}$ , qui a la structure de variance-covariance suivante

$$\text{Cov}_{M_2}(\epsilon_{qik}, \epsilon_{pjl}) = \begin{cases} \sigma^2, & \text{si } q = p, i = j, k = l \\ \rho_C \sigma^2, & \text{si } q = p, i \neq j, k \neq l \\ \rho_I \sigma^2, & \text{si } q \neq p, i = j, k \neq l \\ (\rho_I + \rho_C) \sigma^2, & \text{si } q = p, i = j, k \neq l \\ 0, & \text{sinon} \end{cases}, \quad (2.5)$$

où  $\sigma_C^2 + \sigma_I^2 + \sigma_e^2 = \sigma^2$  et  $\rho_C = \frac{\sigma_C^2}{\sigma^2}$  est la corrélation entre l'observation de la même UPE. La structure de variance-covariance de  $\epsilon_{qik}$  signifie que les mesures de  $y$  sont corrélées si elles sont effectuées dans la

même UPE ou le même intervieweur. De plus, les mesures de  $y$  sont plus homogènes si elles sont réalisées par un même intervieweur dans une même UPE. Le modèle ( $M_2$ ) représente une généralisation du modèle  $M_4$  de Gabler et Lahiri (2009), en supprimant la restriction selon laquelle aucun intervieweur ne travaille dans plus d'une UPE.

La variance de  $\bar{y}(\mathbf{w})$  selon le modèle ( $M_2$ ) est donnée par

$$\text{Var}_{M_2}(\bar{y}(\mathbf{w})) = \frac{\sigma^2 \sum_{q \in \mathcal{K}} \sum_{i \in \mathcal{R}} \sum_{k \in s_{qi}} w_{qik}^2}{\left( \sum_{q \in \mathcal{K}} \sum_{i \in \mathcal{R}} \sum_{k \in s_{qi}} w_{qik} \right)^2} (1 + \rho_I [\bar{m}_I(\mathbf{w}) - 1] + \rho_C [\bar{m}_C(\mathbf{w}) - 1]), \quad (2.6)$$

où  $w_{qik}$  et  $y_{qik}$  sont le poids d'enquête et l'observation pour le répondant  $k \in s_{qi}$ , respectivement et

$$\bar{m}_I(\mathbf{w}) = \frac{\sum_{i \in \mathcal{R}} \left( \sum_{q \in \mathcal{K}} \sum_{k \in s_{qi}} w_{qik} \right)^2}{\sum_{q \in \mathcal{K}} \sum_{i \in \mathcal{R}} \sum_{k \in s_{qi}} w_{qik}^2} \quad \text{et} \quad \bar{m}_C(\mathbf{w}) = \frac{\sum_{q \in \mathcal{K}} \left( \sum_{i \in \mathcal{R}} \sum_{k \in s_{qi}} w_{qik} \right)^2}{\sum_{q \in \mathcal{K}} \sum_{i \in \mathcal{R}} \sum_{k \in s_{qi}} w_{qik}^2}.$$

Nous pouvons modifier le modèle ( $M_2$ ) pour permettre un effet d'interaction entre l'intervieweur et l'UPE, ce qui signifie que la covariance entre les observations faites par un même intervieweur dans une même UPE n'est pas égale à la somme de la covariance intra-intervieweur et intra-UPE. Nous appelons ce modèle de mesure ( $M_{2*}$ ) et pour  $k \in s_{qi}$  l'observation de  $y$  est modélisée comme suit

$$y_{qik} = \mu_k + \mathbb{C}_q + \mathbb{I}_i + \mathbb{D}_{qi} + \epsilon_{qik}, \quad (M_{2*})$$

avec  $\mathbb{D}_{qi}$  comme variable aléatoire de moyenne nulle et de variance  $\sigma_{IC}^2$  commune à tous les répondants de l'UPE  $q$  qui ont été interviewés par l'intervieweur  $i$ . Toutes les valeurs  $\mathbb{D}_{qi}$  pour  $q \in \mathcal{K}$  et  $i \in \mathcal{R}$  sont des variables aléatoires iid et sont indépendantes de  $\epsilon_{qik}$ ,  $\mathbb{I}_i$ ,  $\mathbb{C}_q$  pour tous les  $q \in \mathcal{K}$ ,  $i \in \mathcal{R}$ , et  $k \in s_{qi}$ . L'effet aléatoire  $\mathbb{D}_{qi}$  introduit une corrélation supplémentaire entre les observations réalisées par un même intervieweur au sein d'une même UPE, ce qui ne peut pas être expliqué par les variances distinctes de l'UPE et de l'intervieweur.

Pour  $k \neq l$  et  $\epsilon_{qik} = \mathbb{C}_q + \mathbb{I}_i + \mathbb{D}_{qi} + \epsilon_{qik}$  nous avons selon le modèle ( $M_{2*}$ )  $\text{Cov}_{M_{2*}}(\epsilon_{qik}, \epsilon_{qil}) = (\rho_I + \rho_C + \rho_{IC}) \sigma^2$ . Par conséquent, nous pouvons écrire la variance de  $\bar{y}(\mathbf{w})$  selon le modèle ( $M_{2*}$ ) comme suit

$$\text{Var}_{M_{2*}}(\bar{y}(\mathbf{w})) = \frac{\sigma^2 \sum_{q \in \mathcal{K}} \sum_{i \in \mathcal{R}} \sum_{k \in s_{qi}} w_{qik}^2}{\left( \sum_{q \in \mathcal{K}} \sum_{i \in \mathcal{R}} \sum_{k \in s_{qi}} w_{qik} \right)^2} (1 + \rho_I [\bar{m}_I(\mathbf{w}) - 1] + \rho_C [\bar{m}_C(\mathbf{w}) - 1] + \rho_{IC} [\bar{m}_{IC}(\mathbf{w}) - 1]), \quad (2.7)$$

où

$$\bar{m}_{IC}(\mathbf{w}) = \frac{\sum_{q \in \mathcal{K}} \sum_{i \in \mathcal{R}} \left( \sum_{k \in s_{qi}} w_{qik} \right)^2}{\sum_{q \in \mathcal{K}} \sum_{i \in \mathcal{R}} \sum_{k \in s_{qi}} w_{qik}^2}.$$

## 2.4 Effet d'enquête

Après avoir établi la variance de  $\bar{y}(\mathbf{w})$  selon différents modèles de mesure, nous pouvons définir l'effet associé à l'échantillonnage complexe et aux intervieweurs. Nous l'appellerons *effet d'enquête*, que nous définissons comme suit :

$$\text{eff}_{ab}(\mathbf{w}) = \frac{\text{Var}_{M_a}(\bar{y}(\mathbf{w}))}{\text{Var}_{M_b}(\bar{y}(\mathbf{w}))}, \quad (2.8)$$

où  $M_a$  est le modèle de mesure supposé pour notre enquête d'intérêt et  $M_b$  est le modèle de référence. Nous parlerons d'*effet d'enquête* pour distinguer  $\text{eff}_{ab}(\mathbf{w})$  de l'effet de plan et de l'intervieweur, car  $\text{eff}_{ab}(\mathbf{w})$  englobe les deux effets. Les autres sources de variance, décrites dans le cadre de l'EET, ne sont pas prises en compte. Par conséquent, nous désignerons par *plan d'enquête* la combinaison d'un plan de sondage et d'un plan de travail d'intervieweur.

L'effet d'enquête associé au modèle de mesure ( $M_2$ ) est donné par

$$\begin{aligned} \text{eff}_{20}(\mathbf{w}) &= \frac{\text{Var}_{M_2}(\bar{y}_w)}{\text{Var}_{M_0}(\bar{y})} \\ &= \text{eff}_w(\mathbf{w}) (1 + \rho_I [\bar{m}_I(\mathbf{w}) - 1] + \rho_C [\bar{m}_C(\mathbf{w}) - 1]), \end{aligned} \quad (2.9)$$

où

$$\text{eff}_w(\mathbf{w}) = \frac{n \sum_{k \in S} w_k^2}{(\sum_{k \in S} w_k)^2} \geq 1.$$

Le facteur  $\text{eff}_w(\mathbf{w})$  ne dépend pas du modèle de mesure et peut être interprété comme une mesure de la variance des poids  $\mathbf{w}$ . Si nous écrivons la variance des poids comme étant  $\sigma_w^2 = 1/n \sum_{k \in S} w_k^2 - \bar{w}^2$ , avec  $\bar{w} = 1/n \sum_{k \in S} w_k$ , cette relation devient plus claire, puisque  $\text{eff}_w(\mathbf{w}) = \text{CV}_w^2 + 1$ , avec  $\text{CV}_w = \sigma_w / \bar{w}$  comme coefficient de variation des poids d'enquête. Si les poids sont tous égaux, alors  $\text{CV}_w = 0$  et  $\text{eff}_w(\mathbf{w})$  devient 1. Les termes  $\bar{m}_I(\mathbf{w})$  et  $\bar{m}_C(\mathbf{w})$  peuvent être considérés comme des mesures de la charge de travail moyenne des intervieweurs et de la taille de l'UPE, respectivement. Si tous les poids sont égaux,  $\bar{m}_I(\mathbf{w})$  a la valeur  $\bar{m}_I(I_n) = \sum_{i \in \mathcal{R}} n_i^2 / n$ . De plus, si tous les intervieweurs ont exactement la même charge de travail, c'est-à-dire  $n_i = n/R$  pour  $i = 1, \dots, R$ , nous avons  $\bar{m}_I(I_n) = n/R$ .  $\bar{m}_C(\mathbf{w})$  a des propriétés semblables.

À la suite de Gabler et coll. (1999) et de Gabler et Lahiri (2009) nous pouvons donner la borne supérieure suivante pour l'effet d'enquête.

### Résultat 1.

$$\text{eff}_{20}^*(\mathbf{w}) \leq \text{eff}_w(\mathbf{w}) \text{eff}_{20}^*(I_n),$$

où  $\text{eff}_{20}^*$  est l'effet d'enquête sous la condition que  $n_i = n/R$  pour toutes les valeurs  $i \in \mathcal{R}$  et  $n_q = n/K$  pour toutes les valeurs  $q \in \mathcal{K}$ . La borne supérieure de  $\text{eff}_{20}^*(\mathbf{w})$ , donnée dans le résultat 1, découle de

$\bar{m}_I(\mathbf{w}) \leq n/R$  si  $n_i = n/R$  pour tous les  $i \in \mathcal{R}$  (Gabler et coll., 1999). La démonstration se trouve dans l'annexe. Pour  $\bar{m}_C(\mathbf{w})$  un résultat analogue se vérifie. Notons qu'en général, nous n'avons pas

$$\text{eff}_{20}(\mathbf{w}) \geq \text{eff}_{20}(\mathbf{I}_n) = 1 + \rho_I \left[ \sum_{i \in \mathcal{R}} \frac{n_i^2}{n} - 1 \right] + \rho_C \left[ \sum_{q \in \mathcal{K}} \frac{n_q^2}{n} - 1 \right]. \quad (2.10)$$

Autrement dit, nous ne pouvons pas dire que l'effet d'enquête est supérieur ou égal à l'effet d'enquête d'un plan de sondage de pondération égale. Si les poids ont la même distribution de fréquence relative entre tous les ensembles  $s_{q_i}$  l'inégalité (2.10) se vérifie (Gabler et Lahiri, 2009), c'est-à-dire que nous obtenons

$$n_{qig} = \frac{n_{qi}}{n} n_g, \quad g = 1, \dots, G, \quad (2.11)$$

où  $G$  est le nombre de valeurs uniques dans  $\mathbf{w}$ ,  $n_g$  la fréquence de la valeur de pondération  $g^e$  et  $n_{qig}$  la fréquence de la valeur de pondération  $g^e$  pour les répondants interviewés par l'intervieweur  $i^e$  dans l'UPE  $q^e$ .

Nous pouvons toutefois donner une borne inférieure à  $\text{eff}_{20}(\mathbf{w})$ . En utilisant le même argument que Gabler et Lahiri (2009) dans la démonstration de leur résultat 6, nous obtenons

$$\text{eff}_{20}(\mathbf{w}) \geq \left( 1 + \rho_I \left[ \frac{n}{R} - 1 \right] + \rho_C \left[ \frac{n}{K} - 1 \right] \right). \quad (2.12)$$

Avec le deuxième membre de l'inégalité (2.12), nous avons un minimum facilement calculable de  $\text{eff}_{20}(\mathbf{w})$  qui ne dépend ni des poids, ni de la répartition des charges de travail des intervieweurs, ni des tailles d'UPE. Cela donne de précieux éléments d'orientation à l'étape de la planification d'un plan d'enquête, car l'effet d'enquête prévu devrait être au moins aussi élevé que  $\text{eff}_{20}^*(\mathbf{I}_n)$ . L'utilité pratique de la borne supérieure dans le résultat 1 est quelque peu limitée par des hypothèses fortes sur  $\mathbf{n}_I$  et  $\mathbf{n}_C$ . Plus les valeurs de  $\mathbf{n}_I$  et  $\mathbf{n}_C$  s'écartent de la distribution en un point des charges de travail d'intervieweur et de tailles d'UPE, moins cette borne devrait servir d'indication. Pour donner aux personnes planifiant une enquête une statistique moins complexe aux fins de prévision de la valeur de  $\bar{m}_I(\mathbf{w})$ , Lynn et Gabler (2004) ont proposé d'utiliser

$$\bar{m}'_I(\mathbf{w}) = \frac{H_{\mathbf{n}_I}}{H_{\mathbf{w}}}, \quad (2.13)$$

comme prédicteur de  $\bar{m}_I(\mathbf{w})$ , où  $H_{\mathbf{n}_I} = \sum_{i \in \mathcal{R}} (n_i/n)^2$  est l'indice de Herfindahl pour la charge de travail des intervieweurs, une mesure de la concentration, avec  $1/R \leq H_{\mathbf{n}_I} \leq 1$  (Fahrmeir, Heumann, Künstler, Pigeot et Tutz, 1997, page 83).  $H_{\mathbf{n}_I} = 1$  correspond à  $R = 1$  et  $H_{\mathbf{n}_I} = 1/R$  correspond à  $n_i = n/R$  pour toutes les valeurs  $i \in \mathcal{R}$ .  $H_{\mathbf{w}} = \sum_{k \in \mathcal{S}} (w_k / \sum_{k \in \mathcal{S}} w_k)^2$  est l'indice de Herfindahl pour les poids. Si l'équation (2.11) se vérifie, nous avons  $\bar{m}_I(\mathbf{w}) = \bar{m}'_I(\mathbf{w})$ , mais, la plupart des enquêtes, cela ne sera pas vrai. C'est pourquoi Lynn et Gabler (2004) proposent d'examiner  $\text{Cov}(w_{qik}, n_i)$ , la covariance entre les poids et les charges de travail des intervieweurs. Plus  $\text{Cov}(w_{qik}, n_i)$  est proche de zéro, plus la distance entre  $\bar{m}_I(\mathbf{w})$  et  $\bar{m}'_I(\mathbf{w})$  est petite. Il devrait être plus facile de planifier une enquête avec des

valeurs supposées pour  $H_{n_I}$  et  $H_w$  qu'avec des valeurs exactes de  $\mathbf{n}_I$  et  $\mathbf{w}$ . Pour trouver des valeurs raisonnables de  $H_{n_I}$  et  $H_w$ , on pourrait s'appuyer sur la comparaison de ces valeurs provenant d'enquêtes ayant des plans d'enquête semblables. Dans l'équation (2.11), les résultats sont analogues pour  $\bar{m}_C(\mathbf{w})$ .

Notons que nous pouvons aussi écrire  $\text{eff}_w(\mathbf{w})$  comme suit

$$\text{eff}_w(\mathbf{w}) = H_w n. \quad (2.14)$$

L'expression de  $\text{eff}_w(\mathbf{w})$  dans l'équation (2.14) pourrait également être utile à l'étape de planification d'une enquête, car elle montre qu'on peut la planifier avec une certaine concentration des poids, plutôt que des valeurs spécifiques de  $\mathbf{w}$ .

Il est difficile d'établir une borne supérieure proche générale pour  $\text{eff}_{20}(\mathbf{w})$  en l'absence de restrictions sur les valeurs de  $\mathbf{n}_I$ ,  $\mathbf{n}_C$  et  $\mathbf{w}$ . Toutefois, les poids d'enquête sont habituellement mis à l'échelle de l'échantillon ou de la taille de la population, et il n'est pas rare qu'ils soient bornés. Par exemple, l'ESS fournit à ses utilisateurs des poids supérieurs à zéro et plus petits ou égaux à 4 et les met à l'échelle de la taille de l'échantillon (ESS, 2014c, 2014b). Si  $a \leq w_k \leq b$  pour tous les  $k \in s$  avec  $b < \infty$  et  $a > 0$ , alors avec une valeur donnée de  $\mathbf{n}_I$  (ou  $\mathbf{n}_C$ ) on peut trouver les bornes supérieures de  $\bar{m}_I(\mathbf{w})$  (ou  $\bar{m}_C(\mathbf{w})$ ) en résolvant un problème d'optimisation linéaire. On peut déduire une borne supérieure pour  $\text{eff}_w(\mathbf{w})$  à partir des valeurs données de  $a$  et  $b$ , comme le montre l'équation (A.5) en annexe.

La borne supérieure obtenue de  $\text{eff}_{20}(\mathbf{w})$  correspondra aux distributions de poids avec une concentration très élevée, c'est-à-dire un nombre maximal des poids les plus élevés possible. Cependant, on pourrait trouver des bornes plus pertinentes en pratique en ajustant les contraintes du problème d'optimisation linéaire, à partir de la distribution des poids des enquêtes ayant des plans de sondage comparables. (Voir la formulation de ce programme linéaire en annexe.)

## 2.5 Effet de plan corrigé

Maintenant que nous avons établi l'effet d'enquête d'un plan d'enquête, nous proposons un nouveau type d'effet d'enquête que nous nommons *effet de plan corrigé*. Cette statistique vise à quantifier l'effet marginal d'un plan d'enquête complexe en présence d'un effet de l'intervieweur. Pour cela, nous définissons l'effet suivant

$$\begin{aligned} \text{eff}_{21}(\mathbf{w}) &= \frac{\text{Var}_{M_2}(\bar{y}_w)}{\text{Var}_{M_1}(\bar{y})} \\ &= \text{eff}_w(\mathbf{w}) \text{eff}_I (1 + \rho_I [\bar{m}_I(\mathbf{w}) - 1] + \rho_C [\bar{m}_C(\mathbf{w}) - 1]), \end{aligned} \quad (2.15)$$

où

$$\text{eff}_I = \frac{n}{n + \rho_I (\sum_{i \in R} n_i^2 - n)}.$$

Le modèle de référence ( $M_1$ ) dans  $\text{eff}_{21}(\mathbf{w})$  modélise un échantillon aléatoire simple avec un effet de l'intervieweur. Le facteur  $\text{eff}_I$  indique à quel point l'effet de plan corrigé est proche de l'effet d'enquête. Pour  $\text{eff}_I = 1$  l'effet de plan corrigé et l'effet d'enquête sont égaux et plus  $\text{eff}_I$  est proche de zéro, plus les deux effets sont éloignés. Par conséquent, nous pouvons utiliser  $\text{eff}_I$  pour construire une mesure de la contribution de l'effet de l'intervieweur à l'effet d'enquête  $\text{eff}_{20}$ . Pour cela, nous établissons d'abord les bornes suivantes pour  $\text{eff}_I$ , qui sont données dans le résultat 2.

### Résultat 2.

$$\frac{1}{n} \leq \frac{n}{\rho_I(n-R)(n-R+1)+n} \leq \text{eff}_I \leq \frac{R}{R+(n-R)\rho_I} \leq 1.$$

La démonstration du résultat 2 se trouve en annexe.

Nous définissons maintenant une mesure de la contribution de l'effet de l'intervieweur à l'effet d'enquête  $\text{inv}_I$  comme suit :

$$\begin{aligned} \text{inv}_I : \left[ \frac{1}{n}, 1 \right] &\mapsto [0, 1], \\ \text{inv}_I(a) &:= \frac{n(1-a)}{n-1} \text{ pour } \left[ \frac{1}{n} \leq a \leq 1 \right]. \end{aligned} \quad (2.16)$$

Pour toute valeur donnée de la charge de travail des intervieweurs  $\mathbf{n}_I$ , la mesure  $\text{inv}_I(\text{eff}_I)$  augmente strictement quand  $\text{eff}_I$  diminue. Le maximum de  $\text{inv}_I(\text{eff}_I)$  se produit quand  $R = 1$  et  $\rho_I = 1$ , qui se produit si l'on a un seul intervieweur produisant toujours la même mesure. Le minimum de  $\text{inv}_I(\text{eff}_I)$  se produit quand  $\rho_I = 0$  pour toute valeur donnée de  $\mathbf{n}_I$ . Si la concentration de la distribution de la charge de travail entre intervieweurs augmente et que  $\rho_I$  reste fixe,  $\text{inv}_I(\text{eff}_I)$  augmente également. Cette relation apparaît plus clairement si nous écrivons

$$\text{eff}_I = \frac{1}{1 + \rho_I(H_{\mathbf{n}_I}n - 1)}. \quad (2.17)$$

Par ailleurs, on pourrait aussi se servir du coefficient de variation pour les charges de travail des intervieweurs  $\text{CV}_{\mathbf{n}_I} = R\sigma_{\mathbf{n}_I}/n$ , avec  $\sigma_{\mathbf{n}_I}^2 = 1/R \sum_{i \in R} n_i^2 - (n/R)^2$ , pour décrire  $\text{eff}_I$ , étant donné que  $H_{\mathbf{n}_I} = (1 + \text{CV}_{\mathbf{n}_I}^2)/R$  (Lynn et Gabler, 2004). Notons que pour  $\sigma_{\mathbf{n}_I}^2 = 0$  nous avons  $\text{eff}_I = R/(R + (n-R)\rho_I)$ .

Au moyen des résultats 1 et 2, ainsi que de l'inégalité (2.12), nous pouvons donner les bornes suivantes à l'effet de plan corrigé.

### Résultat 3.

$$\frac{n}{\rho_I(n-R)(n-R+1)+n} \text{eff}_{20}^*(I_n) \leq \text{eff}_{21}^*(\mathbf{w}) \leq \text{eff}_{\mathbf{w}}(\mathbf{w}) \frac{R}{R+(n-R)\rho_I} \text{eff}_{20}^*(I_n),$$

où  $\text{eff}_{21}^*$  est l'effet de plan corrigé en cas de charges de travail d'intervieweurs égales et de tailles d'UPE égales. Les bornes de  $\text{eff}_{21}^*$ , données dans le résultat 3, ne dépendent pas de  $\mathbf{n}_I$ , mais il faut noter que dans la borne inférieure de  $\text{eff}_{21}^*(\mathbf{w})$ ,  $\text{eff}_I$  prend sa valeur de concentration maximale dans  $\mathbf{n}_I$ , tandis que  $\text{eff}_{20}^*(I_n)$  correspond à la concentration minimale de  $\mathbf{n}_I$ . Comme  $\text{eff}_I$  ne dépend pas de  $\mathbf{w}$ , on peut trouver une borne supérieure (ou inférieure) de  $\text{eff}_{21}(\mathbf{w})$  en obtenant les bornes supérieures (ou inférieures) de  $\bar{m}_I(\mathbf{w})$ ,  $\bar{m}_C(\mathbf{w})$  et  $\text{eff}_w(\mathbf{w})$ , comme cela est décrit en annexe.

Enfin, nous introduisons un effet de plan corrigé qui suppose le modèle de mesure ( $M_{2*}$ ), donné par

$$\begin{aligned} \text{eff}_{2*1}(\mathbf{w}) &= \frac{\text{Var}_{M_{2*}}(\bar{y}_w)}{\text{Var}_{M_1}(\bar{y})} \\ &= \text{eff}_w(\mathbf{w}) \text{eff}_I (1 + \rho_I [\bar{m}_I(\mathbf{w}) - 1] + \rho_C [\bar{m}_C(\mathbf{w}) - 1] + \rho_{IC} [\bar{m}_{IC}(\mathbf{w}) - 1]). \end{aligned} \quad (2.18)$$

Comme dans le résultat 3, nous pouvons établir les bornes suivantes de  $\text{eff}_{2*1}(\mathbf{w})$ .

#### Résultat 4.

$$\frac{n}{\rho_I(n-R)(n-R+1) + n} \text{eff}_{2*0}^*(I_n) \leq \text{eff}_{2*1}^*(\mathbf{w}) \leq \text{eff}_w(\mathbf{w}) \frac{R}{R + (n-R)\rho_I} \text{eff}_{2*0}^*(I_n).$$

Ici  $\text{eff}_{2*1}^*$  correspond au cas où  $n_{qi}$ , le nombre de répondants qui appartiennent à l'UPE  $q^c$  et qui sont interviewés par l'intervieweur  $i^c$ , est une constante, c'est-à-dire  $n_{qi} = n/(RK)$  pour toutes les valeurs  $i \in \mathcal{R}$  et  $q \in \mathcal{K}$ . Cela implique également que pour  $\text{eff}_{2*1}$  nous ayons  $n_i = n/R$  et  $n_q = n/K$ . La démonstration du résultat 4 se trouve en annexe. En utilisant le modèle ( $M_{2*}$ ) au lieu de ( $M_2$ ), on obtient plus de souplesse pour ajuster le modèle de mesure aux données observées. La question de savoir si cela est nécessaire est traitée dans la section 3.2, où les différents modèles de mesure sont mis à l'essai l'un par rapport à l'autre pour les données de l'ESS6.

### 3 Résultats empiriques tirés de l'ESS

Après avoir établi les effets associés aux intervieweurs et à l'échantillonnage en grappes ou à plusieurs degrés, nous estimons l'effet d'enquête et l'effet de plan corrigé proposé pour les données de l'ESS6 (ESS, 2016).

Vingt-neuf pays ont participé à l'ESS6 (ESS, 2018a), mais nous ne les avons pas tous pris en compte dans notre analyse. Nous avons exclu tous les pays dotés d'un plan à un degré (il n'y avait pas de plan d'échantillonnage en grappes à un degré dans l'ESS6). De plus, nous avons exclu les pays qui utilisaient un plan de sondage multidomaine. Ces pays ont utilisé différents plans d'échantillonnage dans différentes régions du pays, mais ils font tous référence à un niveau de la Nomenclature des unités territoriales statistiques (NUTS), établie par Eurostat (ESS, 2013, pages 21-22). Par exemple, la Norvège a utilisé un échantillonnage à un degré pour ses régions les plus densément peuplées, qui à elles seules contenaient près de 75 % de la population cible, et un plan d'échantillonnage à deux degrés pour le reste du pays.

Tout d'abord, dans la section 3.1, nous évaluons la faisabilité générale de l'estimation des modèles de mesure décrits à la section 2, compte tenu de la structure l'UPE-intervieweur qui se trouve dans l'ESS6. À cette fin, nous utilisons une étude par simulations fondée sur un modèle. À la section 3.2, nous testons les différents modèles de mesure les uns par rapport aux autres afin d'utiliser les modèles les plus adéquats dans l'estimation de l'effet d'enquête et de l'effet de plan corrigé. Ensuite, nous comparons nos résultats avec l'effet de plan utilisé par l'ESS dans la planification de la taille de l'échantillon.

Les variables d'identification de l'intervieweur et de l'UPE nécessaires dans notre étude par simulations et l'estimation des effets ont été obtenues à partir des fichiers de données des plans de sondage (FDPS) et du questionnaire de l'intervieweur, respectivement (ESS, 2014a). Les FDPS contiennent des renseignements sur le plan de sondage, y compris un identificateur d'UPE. Dans le cas de l'ESS, il faut télécharger les FDPS par pays (ESS, 2018b).

### 3.1 Simulation de l'évaluation de la stabilité des estimations des effets

Depuis longtemps, on sait que les intervieweurs et l'échantillonnage sont les principales sources des erreurs d'enquête. En raison des modalités de déploiement des intervieweurs sur le terrain, il est difficile de séparer la variance de l'intervieweur de celle de l'UPE. Pour accroître l'efficacité de la collecte de données, les intervieweurs sont habituellement affectés exclusivement à certaines régions (Von Sanden, 2004, section 1.3). Ainsi, les intervieweurs de l'ESS6 travaillent rarement dans plusieurs régions. Pour ce qui est de l'ESS6, nous observons la situation suivante : en général, les intervieweurs travaillent dans plusieurs UPE d'une même région, mais jamais dans toutes les UPE. Les UPE peuvent recevoir la visite de plusieurs intervieweurs, mais jamais de tous. Dans 25 % de tous les pays considérés, le nombre moyen de régions (variable *région*, ESS (2013), pages 21-22) dans lesquelles un intervieweur travaillait était inférieur ou égal à 1,017. Dans 75 % de tous les pays, le nombre moyen de régions par intervieweur était inférieur ou égal à 1,256.

La structure non hiérarchique des UPE et des intervieweurs peut être considérée comme typique des enquêtes sociales à grande échelle comme l'ESS. Un plan d'enquête dit *entièrement interpénétré*, dans lequel tous les intervieweurs travaillent dans toutes les UPE, est généralement impossible dans les enquêtes nationales. Il est donc difficile de déterminer dans quelle mesure la similitude constatée entre les observations faites par un intervieweur est due à la corrélation intra-intervieweur ou à la corrélation intra-UPE. De nombreuses études essaient de résoudre ce problème, par exemple en utilisant un plan d'enquête entièrement emboîté, où plusieurs intervieweurs travaillent dans une même UPE, mais pas dans plusieurs UPE (Schnell et Kreuter, 2005). Ont également été analysées des enquêtes dites *partiellement interpénétrées*, où différents intervieweurs travaillent dans plusieurs UPE et où les UPE reçoivent la visite de plusieurs intervieweurs (Davis et Scott, 1995; O'Muircheartaigh et Campanelli, 1998). Ces enquêtes partiellement interpénétrées ressemblent davantage à la situation que nous observons dans l'ESS6.

Pour tester nos modèles de mesure et distinguer les différentes composantes de la variance, nous ajustons un modèle multiniveau avec des effets aléatoires croisés. Dans un autre contexte, Raudenbush

(1993) propose de permettre des effets dits *croisés* dans la structure des effets aléatoires. Ces effets croisés permettent la situation de facteurs partiellement interpénétrés et peuvent estimer les trois composantes de la variance du modèle de mesure ( $M_{2*}$ ), à savoir  $\sigma_I^2$ ,  $\sigma_C^2$  et  $\sigma_{IC}^2$ .

Vassallo, Durrant et Smith (2017) montrent, à l'aide de simulations sur des données synthétiques, dans quelle mesure un modèle multiniveau avec des effets aléatoires croisés pour la grappe et l'intervieweur peut estimer la structure de variance-covariance du modèle de données selon différentes configurations d'interpénétration entre la grappe et l'intervieweur. Ils constatent que la taille de l'échantillon, le nombre d'intervieweurs et d'UPE et le degré d'interpénétration sont les facteurs déterminant la qualité des estimations des composantes de la variance. Le degré d'interpénétration joue un rôle décisif dans la qualité des estimations des composantes de la variance. Vassallo et coll. (2017) ont constaté que deux ou trois intervieweurs par UPE donnent déjà des estimations relativement stables des composantes de la variance. Toutefois, tous leurs plans d'enquête étaient équilibrés et symétriques, ce qui signifie que l'interpénétration des UPE par les intervieweurs était constante pour toutes les UPE et vice versa. Or ce n'est pas le cas pour les pays de l'ESS6. Par conséquent, nous effectuons une simulation pour vérifier si selon les plans d'enquête partiellement interpénétrés de l'ESS6, il est possible ou non d'estimer les composantes de la variance de notre modèle de mesure ( $M_2$ ).

Pour la simulation, nous générons des échantillons à partir d'une distribution normale à plusieurs variables à  $n$  dimensions. Le vecteur des moyennes  $\mu$  contient la même valeur pour chaque dimension. La matrice de covariance  $\Sigma$  suit la structure de variance-covariance du modèle de mesure ( $M_2$ ) et a été construite pour chaque pays en fonction de la structure UPE-intervieweur observée. Les composantes de la variance ont été établies à  $\sigma_I^2 = 0,2$ ;  $\sigma_C^2 = 0,08$ ;  $\sigma^2 = 2$ . Nous avons produit 1 000 échantillons à partir du modèle de superpopulation NPV ( $\mu, \Sigma$ ) pour chaque pays et estimé le modèle de mesure ( $M_2$ ) pour chacun de ces échantillons. La simulation a été mise en œuvre dans R (R Core Team, 2019). Les échantillons pour la simulation ont été générés au moyen du module *mvtnorm* (Genz, Bretz, Miwa, Mi et Hothorn, 2019) et l'estimation du modèle a été réalisée au moyen du module *lme4* (Bates, Mächler, Bolker et Walker, 2015, 2019).

Le tableau 3.1 illustre le biais de Monte-Carlo relatif des estimateurs pour les composantes de la variance du modèle ( $M_2$ ). Pour un estimateur  $\hat{\theta}$  de  $\theta$  nous définissons cette mesure comme étant

$$\text{biaisR-MCs } \hat{\theta} = \frac{\bar{\hat{\theta}}_{MC}}{\theta} - 1,$$

où  $\bar{\hat{\theta}}_{MC} = \sum_{d=1}^D \hat{\theta}_d / D$ ,  $\theta$  est la valeur vraie,  $\hat{\theta}_d$  la valeur de  $\hat{\theta}$  pour l'échantillon  $d^e$  de la simulation et  $D$  est le nombre total d'échantillons générés, c'est-à-dire  $D = 1\,000$ , dans notre simulation. Nous constatons que  $\sigma_I^2$  et  $\sigma_C^2$  sont estimés avec un biais relativement faible pour tous les pays pris en considération dans l'ESS6. En plus du biais relatif de Monte-Carlo, nous avons ajouté au tableau 3.1 le nombre d'UPE  $K$ , le nombre d'intervieweurs  $R$ , la taille de l'échantillon  $n$ , le nombre moyen d'UPE dans lesquelles un intervieweur travaille  $\bar{K}_I$  et le nombre moyen d'intervieweurs qui travaillent dans une

UPE  $\bar{R}_C$ . Les valeurs  $\bar{K}_I$  et  $\bar{R}_C$  sont respectivement utilisées comme mesures du degré d'interpénétration des UPE par les intervieweurs et des intervieweurs par les UPE. Pour tous les pays autres que l'Allemagne, il y a plus d'UPE que d'intervieweurs et  $\bar{K}_I$  est supérieur à  $\bar{R}_C$ .  $\bar{K}_I$  va de 1,423 en Allemagne à 17,396 en Albanie. Le niveau de  $\bar{K}_I$  observé pour tous les pays semble suffisamment élevé pour permettre de distinguer les composantes de variance du modèle ( $M_2$ ). Nous pouvons observer une relation négative entre  $\bar{K}_I$  le biaisR-MC  $\hat{\sigma}_I^2$ , dont la médiation peut être faite par  $n$  et  $K$ . Des valeurs plus élevées de  $n$  et  $K$  correspondent à une plus grande exactitude de  $\hat{\sigma}_I^2$ . On peut faire une observation analogue pour le biaisR-MC  $\hat{\sigma}_C^2$ . Une variable  $\bar{R}_C$  plus élevée améliore également la précision des estimations et peut compenser une valeur basse de  $\bar{K}_I$ . Une interpénétration unilatérale suffisamment élevée, que ce soit des UPE par les intervieweurs ou l'inverse, suffit pour estimer avec précision  $\sigma_I^2$  et  $\sigma_C^2$  pour le modèle ( $M_2$ ). Par exemple, la République tchèque, qui a la variable  $\bar{R}_C$  la plus basse, mais une variable  $\bar{K}_I$  d'environ 1,848 permet des estimations relativement précises des composantes de la variance.

Notons que pour le modèle de mesure ( $M_{2*}$ ),  $\bar{K}_I$  et  $\bar{R}_C$  sont importants. Par exemple, on ne peut pas estimer précisément  $\sigma_C^2$  et  $\sigma_{IC}^2$  si  $\bar{R}_C$  est trop faible. Par exemple, dans une simulation similaire pour le modèle ( $M_{2*}$ ), il n'a pas été possible d'obtenir des estimations exactes de  $\sigma_C^2$  et  $\sigma_{IC}^2$  pour la République tchèque, bien que le biais relatif de  $\hat{\sigma}_I^2$  était d'environ 1 %.

Pour la Bulgarie et la République tchèque  $\bar{R}_C = 1$ , c'est-à-dire que leurs UPE sont emboîtées dans les intervieweurs. Dans ce cas, nous n'avons pas d'effets aléatoires croisés, mais des effets aléatoires emboîtés, car nous n'avons jamais de cas où les répondants se trouvent dans une même UPE, mais ne sont pas interviewés par le même intervieweur. Pour ce cas particulier, à strictement parler,  $\sigma_C^2$  devrait être étiqueté  $\sigma_{IC}^2$ . Cependant, par souci de simplification, nous utilisons dans les deux cas  $\sigma_C^2$  comme étiquette des variances de l'effet aléatoire de l'UPE. Cela n'est pas entièrement injustifié, car  $\sigma_{IC}^2$  définit la corrélation supplémentaire entre les répondants qui sont dans la même UPE, comparativement aux répondants qui sont interviewés par le même intervieweur, mais qui sont dans des UPE différentes.

**Tableau 3.1**  
**Biais relatif des estimations de la variance de l'effet aléatoire**

	biaisR-MC $\hat{\sigma}_I^2$	biais-MC $\hat{\sigma}_C^2$	$K$	$R$	$n$	$\bar{K}_I$	$\bar{R}_C$
Albanie	0,00	-0,02	264	53	1 201	17,40	3,49
Belgique	0,00	-0,02	363	155	1 869	3,00	1,28
Bulgarie	-0,01	0,04	400	247	2 260	1,63	1,00
République tchèque	0,01	-0,01	426	231	2 009	1,85	1,00
France	0,01	0,01	267	165	1 968	1,99	1,23
Allemagne	0,01	-0,00	156	194	2 958	1,42	1,77
Irlande	-0,01	0,01	212	116	2 628	2,15	1,17
Israël	-0,00	0,01	190	114	2 508	3,00	1,80
Italie	-0,02	0,05	129	117	960	1,49	1,35
Kosovo	0,01	-0,02	160	72	1 295	2,29	1,03
Slovaquie	-0,02	0,04	249	132	1 847	1,93	1,02
Slovénie	-0,01	0,00	150	50	1 257	3,30	1,10
Espagne	-0,01	0,03	422	74	1 889	8,20	1,44
Ukraine	0,00	0,00	306	237	2 178	1,44	1,11
Royaume-Uni	-0,01	0,00	226	150	2 286	2,36	1,57

Notre étude par simulations confirme et enrichit les conclusions de Vassallo et coll. (2017) pour la situation déséquilibrée de l'ESS6. Nous avons également constaté que la structure UPE-intervieweur observée pour l'ESS6 n'empêche pas de distinguer  $\sigma_C^2$  et  $\sigma_I^2$  pour le modèle de mesure ( $M_2$ ).

### 3.2 Les effets d'enquête dans la 6<sup>e</sup> édition de l'Enquête sociale européenne

Comme le montre notre étude par simulations, il est possible d'estimer la variance de l'intervieweur et de la grappe dans l'ESS6. Nous testons maintenant, pour un ensemble de variables sélectionnées du questionnaire principal de l'ESS (ESS, 2013), chaque composante de variance du modèle ( $M_{2*}$ ) sur sa signification. Toutes les variables utilisées, sauf l'âge et le genre, ont une échelle ordinale, mais sont traitées comme des variables métriques aux fins de la présente analyse. La liste complète des variables utilisées se trouve en annexe.

Comme le minimum d'une composante de variance est à zéro, l'essai est effectué à la limite de l'espace de paramètre, ce qui impose des problèmes classiques de la théorie des tests. Scheipl, Greven et Kuechenhoff (2008) ont proposé un test de rapport de vraisemblance restreint, conçu pour tester une variance nulle des effets aléatoires. Nous utilisons leur mise en œuvre de ce test dans le module R *RLRsim* et nous faisons trois tests de décision.

Premièrement, nous testons la signification de la variance d'interaction des intervieweurs et des UPE, en supposant des variances des intervieweurs et des UPE pertinentes. Notre hypothèse nulle est  $H_0 : \sigma_{IC}^2 = 0$  tandis que l'hypothèse alternative est  $H_A : \sigma_{IC}^2 > 0$ . La moyenne par pays de l'hypothèse nulle rejetée pour les différentes variables est présentée dans le tableau 3.2. Les deux premières colonnes correspondent à deux niveaux d'erreur de type I différents pour le test de  $H_0 : \sigma_{IC}^2 = 0$ , indiqués par  $\alpha = 0,01$  et  $0,05$ . Israël est le pays où la variance d'interaction significative  $\sigma_{IC}^2$  est la plus élevée pour tous les niveaux d'erreur de type I. Pour tous les autres pays, l'hypothèse nulle n'est pas rejetée pour toutes les variables à un niveau de signification de 1 %. Bien qu'on ne le voie pas dans le tableau 3.2, il faut noter qu'à un niveau de signification de 10 %, les deux tiers des pays ont au moins quelques variables ayant une variance d'interaction significative. Par conséquent, il faut tenir compte de la possibilité d'un effet d'interaction au moment d'estimer les effets d'enquête.

Dans notre deuxième test de décision, nous supposons une variance d'intervieweur, mais aucune variance d'interaction. L'hypothèse nulle est que la variance de l'UPE n'est pas pertinente, c'est-à-dire  $H_0 : \sigma_C^2 = 0$  par opposition à l'hypothèse alternative  $H_A : \sigma_C^2 > 0$ . Les colonnes 3 à 4 du tableau 3.2 présentent les résultats moyens des tests pour les différents niveaux d'erreur de type I. Pour certaines variables, les variances d'UPE ne sont pas significatives comme ajout à la variance d'intervieweur. Ce résultat est particulièrement marqué en Belgique, où 3 % des variables seulement semblent avoir une variance d'UPE. On constate aussi que dans les cas de la France et de la Slovaquie, la variance d'UPE est significative seulement à un niveau de 1 % pour un nombre relativement petit de variables et dans le cas de l'Albanie pour aucune des variables. En revanche, la Bulgarie, l'Irlande, Israël et la Slovaquie présentent une variance d'UPE significative pour la majorité des variables. Dans l'ensemble, la variance

d'UPE semble pertinente dans la plupart des pays et devrait donc être prise en compte dans l'estimation des effets d'enquête.

Pour le troisième test de décision que nous faisons, nous supposons une variance d'UPE, mais aucune variance d'interaction. L'hypothèse nulle est que l'effet de l'intervieweur n'est pas pertinent  $H_0 : \sigma_I^2 = 0$  par rapport à l'hypothèse alternative  $H_A : \sigma_I^2 > 0$ . Les colonnes 7 à 9 du tableau 3.2 présentent les résultats moyens des tests. Les taux de rejet les plus faibles se trouvent en Allemagne et en France, bien que 19 % des variables de l'Allemagne et 23 % des variables de la France présentent toujours une variance d'intervieweur significative à un niveau de signification de 1 %. Les autres pays affichent une proportion beaucoup plus élevée de variables avec une variance d'intervieweur importante. Pour ce qui est du niveau de signification de 1 % et de 5 %, la variance d'intervieweur a un taux de rejet plus élevé que la variance d'UPE pour 13 des 15 pays. Ainsi, il semble que la variance des intervieweurs soit pertinente pour tous les pays de l'ESS6, ce qui indique que les effets possibles des intervieweurs doivent être pris en compte dans l'évaluation de l'efficacité des plans d'enquête.

**Tableau 3.2**  
**Taux de rejet de l'existence des composantes de la variance**

$H_0 :$ $\alpha$	$\sigma_{\alpha}^2 = 0$		$\sigma_c^2 = 0$		$\sigma_I^2 = 0$	
	0,01	0,05	0,01	0,05	0,01	0,05
Albanie	0,00	0,03	0,00	0,16	0,55	0,77
Belgique	0,00	0,00	0,03	0,03	0,77	0,90
Bulgarie	0,00	0,00	0,81	0,90	0,90	1,00
République tchèque	0,00	0,00	0,52	0,58	1,00	1,00
France	0,00	0,00	0,10	0,23	0,23	0,45
Allemagne	0,00	0,00	0,26	0,61	0,19	0,42
Irlande	0,00	0,06	0,77	0,81	0,94	0,97
Israël	0,13	0,32	0,94	1,00	0,84	0,94
Italie	0,00	0,03	0,10	0,32	0,42	0,65
Kosovo	0,00	0,00	0,45	0,58	0,94	0,97
Slovaquie	0,00	0,00	0,77	0,90	0,97	0,97
Slovénie	0,00	0,00	0,03	0,16	0,74	0,84
Espagne	0,00	0,00	0,13	0,23	0,74	0,84
Ukraine	0,00	0,00	0,55	0,74	0,90	0,94
Royaume-Uni	0,00	0,03	0,19	0,35	0,71	0,87

On estime les effets d'enquête définis dans l'équation (2.8) à partir des modèles sélectionnés pour les différentes variables. Le tableau 3.3 présente la moyenne par pays des effets d'enquête estimés pour toutes les variables considérées. Le tableau 3.3 contient également la moyenne de l'effet de plan deff, utilisé par l'ESS dans la planification de la taille des échantillons. Nous notons cet effet de plan sous la forme

$$\text{deff} = \text{eff}_w (1 + \rho_c (\bar{m}_c(\mathbf{w}) - 1)).$$

Pour estimer  $\rho_c$  dans deff, nous avons utilisé un estimateur d'analyse de la variance (Panel d'experts sur l'échantillonnage de l'ESS, 2016; Ganninger, 2010, page 45) et nous ne testons pas la signification de la variance d'UPE. Le modèle de mesure  $a$  utilisé dans  $\text{eff}_{a0}$  peut inclure l'intervieweur, l'UPE et la variance d'interaction, si la sélection du modèle la juge significative à un niveau de 0,05. Il en va de même pour le modèle de mesure  $a$  utilisé dans  $\text{eff}_{a1}$ , c'est-à-dire l'effet de plan corrigé. Si la variance

d'intervieweur est jugée comme non significative pour une variable, alors  $eff_{a1}$  devient  $eff_{a0}$ . Pour mesurer l'influence de l'intervieweur sur l'effet d'enquête,  $inv_I$  est également montré.

En comparant  $deff$  et  $eff_{a1}$  dans le tableau 3.3, on peut faire une observation intéressante : dans le cas de l'Allemagne,  $deff$  est nettement inférieur à celui de l'Irlande et de la République tchèque. Nous pouvons en déduire que l'Allemagne aurait besoin d'une taille d'échantillon beaucoup plus petite pour obtenir la même taille d'échantillon efficace moyenne que l'Irlande et la République tchèque. Toutefois, si nous observons  $eff_{a1}$ , cette relation change. Le tableau 3.3 montre que l'effet de grappe du plan de sondage complexe est plus élevé en Allemagne qu'en Irlande ou en République tchèque. Cela signifie que si nous souhaitons un échantillon efficace moyen égal pour tous les pays, l'Allemagne aurait besoin d'une taille d'échantillon plus grande que l'Irlande ou la République tchèque. Par exemple, pour que la République tchèque atteigne une taille d'échantillon efficace de 1 500 avec l'effet de plan  $deff$  type du tableau 3.3, nous prévoirions un échantillon net d'environ 3 925 pour ce pays et de 3 115 pour l'Allemagne. En revanche, si nous utilisons l'effet de plan corrigé  $eff_{a1}$  pour fonder la planification de la taille nette de l'échantillon uniquement sur l'effet du plan de sondage, nous choisirions une taille d'échantillon nette d'environ 1 707 pour la République tchèque et d'environ 2 598 pour l'Allemagne. Cette constatation se reflète également dans les valeurs de  $inv_I$ , ce qui indique qu'une grande partie de  $eff_{a0}$  pour l'Irlande et la République tchèque peut être attribuée à un effet de l'intervieweur, tandis qu'en Allemagne, l'effet de l'intervieweur est plus petit et  $eff_{a0}$  semble dominé par l'effet de grappe. Mis à part Israël, la Slovaquie et la Slovénie, tous les pays ont des rangs différents pour  $deff$  et  $eff_{a1}$ , ce qui indique que la répartition de la taille de l'échantillon dans tous les pays serait très différente, si le plan corrigé avait été utilisé dans la planification des tailles d'échantillons efficaces, plutôt que l'effet de plan classique  $deff$ .

**Tableau 3.3**  
**Tailles de l'effet moyenne l'ESS6**

	<b>deff</b>	<b>eff<sub>a0</sub></b>	<b>eff<sub>a1</sub></b>	<b>inv<sub>I</sub></b>
Albanie	2,07	2,87	1,68	0,35
Belgique	1,18	1,75	1,01	0,37
Bulgarie	2,32	3,88	1,21	0,65
République tchèque	2,62	6,58	1,14	0,78
France	1,69	1,80	1,46	0,16
Allemagne	2,08	2,28	1,73	0,19
Irlande	3,32	5,42	1,26	0,73
Israël	2,41	4,67	1,42	0,61
Italie	1,76	2,20	1,32	0,34
Kosovo	4,01	10,97	1,51	0,80
Slovaquie	5,02	20,28	2,27	0,85
Slovénie	1,59	3,03	1,06	0,55
Espagne	1,16	2,01	1,05	0,42
Ukraine	2,97	5,61	1,18	0,73
Royaume-Uni	1,76	2,24	1,32	0,38

$deff$  : effets de plan moyens définis dans l'équation (1.1).

$eff_{a0}$  : effet d'enquête moyen avec modèle de mesure d'intérêt ( $M_a$ ) et ( $M_0$ ) comme référence.

$eff_{a1}$  : effets de plan corrigés moyens avec modèle de mesure d'intérêt ( $M_a$ ) et ( $M_1$ ) comme référence.

$inv_I$  : contribution moyenne des effets de l'intervieweur à l'effet de plan selon sa définition dans l'équation (2.16).

$eff_{a1}$  est plus petit que  $deff$  pour tous les pays, et leur distance  $|deff - eff_{a1}|$  a une relation positive mais non linéaire avec  $inv_j$ . Les valeurs les plus basses de  $|deff - eff_{a1}|$  sont observées dans les cas de l'Espagne, la Belgique, la France et l'Allemagne, qui sont tous des pays dont la valeur  $inv_j$  est inférieure à la médiane de  $inv_j$ . L'inverse est observé pour la Slovaquie, le Kosovo, l'Irlande et l'Ukraine, soit les pays qui présentent la plus grande distance entre  $deff$  et  $eff_{a1}$ . Ces pays ont tous une valeur de  $inv_j$  supérieure à la valeur médiane de  $inv_j$ . Ces tendances observées pour les pays ayant une distance relativement élevée entre  $deff$  et  $eff_{a1}$  correspondent à ce à quoi nous nous attendrions en cas d'effet de l'intervieweur élevé dans les données. On peut dire le contraire pour les pays dans lesquels on observe une distance relativement petite entre  $deff$  et  $eff_{a1}$ .

Les effets de l'intervieweur dépendent de nombreux facteurs différents (West et Blom, 2017), y compris le type de question posée, or les données de l'ESS6 utilisées sont principalement tirées de questions sur les attitudes. Par conséquent, on ne peut pas extrapoler les résultats présentés dans la section à d'autres types d'enquêtes dans les mêmes pays.

## 4 Conclusions

L'utilisation d'un effet de plan aux fins de sélection d'une taille d'échantillon est une méthode dont on se sert couramment pour tenir compte de la perte d'efficacité qu'un plan de sondage complexe peut entraîner. Toutefois, un effet de l'intervieweur dans les enquêtes en personne peut augmenter l'effet de plan. Cela peut conduire à des conclusions erronées au sujet de l'effet que l'échantillonnage complexe a sur l'efficacité d'une stratégie d'échantillonnage. Ces conclusions erronées peuvent entraîner à leur tour une mauvaise attribution des ressources. La taille d'échantillon prévue pourrait être trop élevée, si elle est fondée sur un effet de plan surestimé. C'est pourquoi nous proposons de tenir compte de l'effet de plan et de l'effet de l'intervieweur simultanément au moment de planifier la taille de l'échantillon. L'effet d'enquête, que nous élaborons à la section 2, tient compte à la fois de la variance des intervieweurs et des UPE pour évaluer l'efficacité d'un plan d'enquête. À partir de l'effet d'enquête, nous introduisons un effet de plan corrigé, qui utilise comme plan de référence un échantillon aléatoire simple avec effet de l'intervieweur. Il en résulte que l'effet de plan corrigé n'est plus confondu avec l'effet de l'intervieweur et qu'il peut donc servir à mieux fonder la décision pour la taille des échantillons en fonction de l'effet du plan de sondage sur la précision des estimations d'enquête.

Dans le cas de l'ESS6, nos constatations empiriques de la section 3.2 montrent que les effets de plan élevés sont liés à des effets de l'intervieweur élevés. Les effets de plan corrigés moyens que nous observons donnent à penser que le plan de sondage influence la variance d'un estimateur à un degré moindre que les intervieweurs dans de nombreux pays de l'ESS6. La capacité d'estimer l'effet de plan corrigé, par exemple à partir des données historiques qui serviraient d'indication dans la planification d'enquête, dépend principalement de la structure UPE-intervieweur et de la répartition des charges de travail des intervieweurs et de la taille des grappes. Nous constatons qu'un plan d'enquête partiellement interpénétré, c'est-à-dire à un niveau régional, peut suffire pour distinguer la variance de l'UPE et celle de l'intervieweur. Dans notre étude par simulations, un nombre moyen de 1,5 UPE par intervieweur ou

d'intervieweur par UPE était suffisant pour l'estimation des composantes de la variance du modèle de mesure ( $M_2$ ). Pour les données d'enquête réelles, c'est-à-dire catégoriques, ce niveau d'interpénétration pourrait ne pas être assez élevé, mais un nombre élevé d'UPE, d'intervieweurs et un échantillon de grande taille pourraient contrebalancer la faible interpénétration. Concernant les applications pratiques, nous recommandons de tester par simulations si le modèle de mesure supposé peut être estimé avec la structure UPE-intervieweur donnée, comme nous l'avons fait à la section 3.1.

Si l'on utilise l'effet d'enquête et l'effet de plan corrigé aux fins de planification d'une taille d'échantillon, il peut être utile de travailler avec les bornes supérieures et inférieures de ces statistiques. À la section 2, nous calculons ces bornes, mais en nous fondant sur des hypothèses quelque peu irréalistes sur la distribution des poids d'enquête, des charges de travail des intervieweurs et des tailles des UPE. Toutefois, s'il est possible de formuler des hypothèses réalistes sur la concentration des poids d'enquête, les charges de travail des intervieweurs et la taille des UPE, nous proposons alors d'utiliser une optimisation linéaire – comme cela est montré en annexe – pour calculer des bornes beaucoup plus pertinentes en pratique, susceptibles de donner de précieuses orientations aux personnes planifiant des enquêtes. En général, nous recommandons d'avoir des distributions peu concentrées des charges de travail des intervieweurs et des tailles de grappes d'UPE afin d'accroître la précision des estimations d'enquête. Par conséquent, les charges de travail des intervieweurs et les tailles des grappes d'UPE devraient être aussi égales que possible pour tout nombre donné d'intervieweurs et d'UPE.

Les modèles de mesure présentés à la section 2 sont sans doute simplistes. Cela les rend applicables à la plupart des plans d'enquête. Les seuls renseignements, à part les données d'enquête, utilisés dans le calcul des estimations du tableau 3.3 étaient les indicateurs d'UPE et d'intervieweurs. On pourrait toutefois intégrer certains aspects des mesures d'enquête à un modèle de mesure pratique, comme la stratification qui, en général, accroît l'efficacité d'une stratégie d'estimation (Särndal et coll., 1992, section 3.7). Notre analyse a négligé cette possibilité, bien que de nombreux pays de l'ESS6 aient utilisé un plan stratifié pour leur échantillon d'UPE. Gabler, Häder et Lynn (2006) ont proposé un effet de plan pour les stratégies d'estimation qui combine différents plans de sondage pour les domaines d'échantillonnage. On pourrait adapter cette méthode pour ajouter un effet de stratification à la variance de l'UPE. En outre, il serait plausible de supposer que les intervieweurs diffèrent quant au degré d'homogénéité qu'ils ajoutent à leurs mesures. On pourrait alors inclure cette hétérogénéité des intervieweurs dans un modèle de mesure en permettant à des groupes d'intervieweurs d'avoir des distributions différentes de  $\mathcal{I}_i$ , c'est-à-dire des valeurs de  $\sigma_i^2$  (West et Elliott, 2014). Il faudrait néanmoins une procédure de classification des intervieweurs. Il serait alors préférable de la fonder principalement sur les données de l'enquête plutôt que sur des renseignements concernant les intervieweurs, qui peuvent différer d'une enquête à l'autre.

Une application future du cadre de l'effet d'enquête présenté ici consisterait à trouver une affectation budgétaire optimale pour ce qui est du nombre d'UPE et d'intervieweurs, pour une taille d'échantillon efficace donnée. Cette optimisation nécessite un modèle de coût pour le déploiement des intervieweurs dans un ensemble possible d'UPE. Les instituts chargés du travail sur le terrain pourraient éventuellement fournir les renseignements nécessaires au calcul d'un tel modèle pour un pays donné. Cette méthode

pourrait aider les responsables de la planification d'enquête à accroître l'efficacité des enquêtes en personne, ce qui est de plus en plus important, car les enquêtes fondées sur des échantillons probabilistes sont soumises à des pressions en raison des solutions de rechange peu coûteuses consistant à recruter des répondants à partir de panels élargis en ligne.

De futures recherches pourraient s'intéresser aussi à l'élaboration d'un effet d'enquête pour d'autres estimateurs que la moyenne pondérée de l'échantillon, notamment pour des estimateurs qu'on peut décrire comme des fonctions de totaux estimés, qui comprennent l'estimateur par les moindres carrés ordinaires pour les coefficients de régression (Särndal et coll., 1992, section 5.10). On devrait pouvoir calculer des effets d'enquête, selon le cadre montré à la section 2, qui permettent une factorisation semblable à l'effet d'enquête présenté dans l'article.

## Annexe

En annexe, nous présenterons une courte notation de plusieurs sommes, où, par exemple,  $\sum_{qik} y_{qik}$  sera l'abréviation de

$$\sum_{q \in \mathcal{K}} \sum_{i \in \mathcal{R}} \sum_{k \in s_{qi}} y_{qik}.$$

## Résultats 1

$$\text{eff}_{20}^*(\mathbf{w}) \leq \frac{n \sum_{qik} w_{qik}^2}{\left( \sum_{qik} w_{qik} \right)^2} \left( 1 + \rho_I \left[ \frac{n}{R} - 1 \right] + \rho_C \left[ \frac{n}{K} - 1 \right] \right).$$

*Démonstration* : Nous devons montrer que

$$\frac{\sum_i \left( \sum_{qk} w_{qik} \right)^2}{\sum_{qik} w_{qik}^2} \leq \frac{n}{R} \quad (\text{A.1})$$

et

$$\frac{\sum_i \left( \sum_{qk} w_{qik} \right)^2}{\sum_{qik} w_{qik}^2} \leq \frac{n}{K} \quad (\text{A.2})$$

se vérifie, si  $n_i = \frac{n}{R}$  et  $n_q = \frac{n}{K}$ , pour tous les  $i = 1, \dots, R$  et  $q = 1, \dots, K$ .

Comme le montrent Gabler et coll. (1999), si  $a_{qik} = 1$  pour tous les  $q \in \mathcal{K}$ ,  $i \in \mathcal{R}$ ,  $k \in s_{qi}$ , au moyen de l'inégalité de Cauchy-Schwarz, nous savons que

$$\begin{aligned} \left( \sum_{qk} w_{qik} a_{qik} \right)^2 &= \left( \sum_{qk} w_{qik} \right)^2 \leq n_i \sum_{qk} w_{qik}^2 = \sum_{qk} a_{qik}^2 \sum_{qk} w_{qik}^2 \\ \frac{\sum_i \left( \sum_{qk} w_{qik} \right)^2}{\sum_i \sum_{qk} w_{qik}^2} &\leq \frac{\sum_i n_i \sum_{qk} w_{qik}^2}{\sum_i \sum_{qk} w_{qik}^2}. \end{aligned}$$

Si nous avons  $n_i = \frac{n}{R}$  pour tous les  $i = 1, \dots, R$ , alors

$$\frac{\sum_i \left( \sum_{qk} w_{qik} \right)^2}{\sum_{qik} w_{qik}^2} \leq \frac{n}{R}.$$

La démonstration de l'inégalité (A.2) est analogue à celle qui précède, ce qui complète la démonstration du résultat 1.

### Bornes supérieures pour $\bar{m}_I(\mathbf{w})$ , $\bar{m}_C(\mathbf{w})$ et $\text{eff}_w(\mathbf{w})$

Pour des valeurs données  $\mathbf{n}_I^\top$  et  $\mathbf{n}_C^\top$  et  $w_k \in [a, b]$  avec  $a, b \in \mathbb{R}_+$  pour tous les  $k \in s$ , et  $\sum_k w_k = n$  nous pouvons construire une borne supérieure pour  $\bar{m}_I(\mathbf{w})$  et  $\bar{m}_C(\mathbf{w})$ .

Nous savons que

$$\frac{\sum_i \left( \sum_{qk} w_{qik} \right)^2}{\sum_i \sum_{qk} w_{qik}^2} \leq \frac{\sum_i n_i \sum_{qk} w_{qik}^2}{\sum_i \sum_{qk} w_{qik}^2} \leq \frac{\sum_i n_i \sum_{qk} w_{qik}^2}{n}. \tag{A.3}$$

Nous devons maintenant trouver une valeur suffisamment élevée de  $\sum_i n_i \sum_{qk} w_{qik}^2$ . Pour cela, nous définissons  $x_i = \sum_{qk} w_{qik}^2$  et  $\mathbf{x} = (x_1, \dots, x_I)^\top$ . Par conséquent, nous devons régler le problème suivant :

$$\begin{aligned} & \max_{\mathbf{x} \in \mathbb{R}^I} \mathbf{n}_I^\top \mathbf{x} \\ & \text{s.c.} \\ & x_i \geq a^2 n_i \quad \forall i \in \mathcal{R} \\ & x_i \leq b^2 n_i \quad \forall i \in \mathcal{R} \\ & \sum_i x_i \geq n \\ & \sum_i x_i \leq f_{sqm}(a, b, n), \end{aligned} \tag{A.4}$$

où

$$f_{sqm}(a, b, n) = b^2 \left\lfloor \frac{n - nb}{a - b} \right\rfloor + (n - nb) - \left\lfloor \frac{n - nb}{a - b} \right\rfloor (a - b) + b + a^2 \left( n - \left\lfloor \frac{n - nb}{a - b} \right\rfloor - 1 \right),$$

où  $\lfloor \cdot \rfloor$  signifie arrondi au nombre entier inférieur le plus proche. On peut résoudre le problème formulé dans l'équation (A.4) au moyen d'un solveur pour programmes linéaires, par exemple la fonction *solveLP* du module R (Henningsen, 2012). La fonction  $f_{sqm}$  donne un maximum de  $\sum_k w_k^2$  étant donné que les bornes supérieure et inférieure des poids sont  $a$  et  $b$  et que les poids sont mis à l'échelle à  $n$ , c'est-à-dire  $\sum_k w_k = n$ . On maximise la somme des carrés en donnant aux plus de poids possible leur valeur la plus élevée  $b$  sous la condition que chaque poids ait au moins une valeur de  $a$  et que  $\sum_k w_k = n$ . On peut ensuite résoudre le problème au moyen d'un algorithme du simplexe. On peut déterminer une borne supérieure pour  $\bar{m}_C$  de la même façon. Si l'on change le problème pour une minimisation, on peut trouver une borne inférieure pour  $\text{eff}_{20}$ . Il n'est toutefois pas garanti que l'optimisation séparée de  $\bar{m}_C$  et  $\bar{m}_I$  produise des valeurs de  $\mathbf{x}$  permettant qu'une valeur de  $\mathbf{w}$  maximise

(ou minimise) conjointement  $\bar{m}_c$  et  $\bar{m}_l$ . Toutefois, si  $\mathbf{x}_c$  et  $\mathbf{x}_l$  sont les vecteurs optimisant  $\bar{m}_c$  et  $\bar{m}_l$  respectivement, il devrait être possible de trouver une valeur possible pour  $\mathbf{w}$ , par exemple au moyen d'un ajustement proportionnel itératif.

Pour  $\text{eff}_w(\mathbf{w})$  nous obtenons selon les mêmes hypothèses que précédemment

$$1 \leq \text{eff}_w(\mathbf{w}) = \frac{\sum_{k \in S} w_k^2}{n} \leq \frac{f_{sqx}(a, b, n)}{n}. \quad (\text{A.5})$$

## Résultat 2

$$\frac{n}{\rho_l(n-R)(n-R+1)+n} \leq \text{eff}_l \leq \frac{R}{R+(n-R)\rho_l}.$$

*Démonstration* : On peut montrer la borne supérieure du résultat 2 en utilisant l'inégalité de Cauchy-Schwarz, qui nous donne

$$\begin{aligned} R \sum_i n_i^2 &\geq \left( \sum_i n_i \right)^2 \\ \sum_i n_i^2 &\geq \frac{n^2}{R}. \end{aligned} \quad (\text{A.6})$$

Avec un peu d'algèbre, nous pouvons formuler la borne supérieure de  $\text{eff}_l$ .

Pour démontrer la borne inférieure du résultat 2, nous résolvons le problème suivant :

$$\begin{aligned} &\max_{\mathbf{n}_l \in \mathbb{N}_{\geq 0}^R} \mathbf{n}_l^\top \mathbf{n}_l \\ &\text{s.c.} \\ &\sum_i n_i = n. \end{aligned} \quad (\text{A.7})$$

On peut trouver une solution au problème formulé dans (A.7) en considérant que si nous avons  $n_i - 1 \geq 1$  et  $n_i \leq n_j$ , alors par conséquent, pour  $n_j = \max_{i \in \mathcal{R}} n_i$  nous pouvons augmenter  $\sum_i n_i^2$  si nous réduisons tout  $n_i > 1$   $i \neq j$  de un et ajoutons un à  $n_j$ . Ainsi, si  $n_i = 1$  pour tous les  $i \neq j \in \mathcal{R}$  et  $n_j = n - R + 1$  alors  $\sum_i n_i^2$  est à son maximum, avec  $\sum_i n_i^2 = (R-1) + (n-R+1)^2$ .

## Résultat 4

*Démonstration* : Compte tenu du résultat 2, pour prouver le deuxième membre du résultat 4, nous devons montrer que

$$\text{eff}_{2^*0}^*(\mathbf{w}) \leq \frac{n \sum_{qik} w_{qik}^2}{\left( \sum_{qik} w_{qik} \right)^2} \left( 1 + \rho_l \left[ \frac{n}{R} - 1 \right] + \rho_c \left[ \frac{n}{K} - 1 \right] + \rho_{lc} \left[ \frac{n}{RK} - 1 \right] \right). \quad (\text{A.8})$$

Pour prouver l'inégalité (A.8), il nous suffit de montrer que

$$\frac{\sum_{qi} \left( \sum_k w_{qik} \right)^2}{\sum_{qik} w_{qik}^2} \leq \frac{n}{RK}.$$

Le reste découle des preuves d'inégalités (A.1) et (A.2). Par conséquent, il suffit de montrer que

$$\left( \sum_k w_{qik} \right)^2 = \left( \sum_k w_{qik} a_{qik} \right)^2 \leq n_{qi} \sum_k w_{qik}^2 = \sum_k a_{qik}^2 \sum_k w_{qik}^2,$$

si  $a_{qik} = 1$  pour toutes les valeurs  $q \in \mathcal{K}$ ,  $i \in \mathcal{R}$ ,  $k \in S_{qi}$ , ce qui découle également de l'inégalité de Cauchy-Schwarz. L'inégalité (A.8) suit alors si  $n_{qi} = \frac{n}{RK}$  pour  $i = 1, \dots, R$  et  $q = 1, \dots, K$ .

Le premier membre du résultat 4 suit la démonstration du résultat 6 dans Gabler et Lahiri (2009) et du résultat 2.

## Variables de l'ESS6 utilisées aux fins d'évaluation empirique

**Tableau A.1**  
Variables de l'ESS6 utilisées aux fins d'évaluation empirique

pplfair	trstprt	stfdem	imueclt	iorgact
pplhlp	trstep	stfedu	imwbcnt	agea
polintr	trstun	stfhlth	happy	gndr
trstprl	lrscale	gincdif	aesfdrk	
trstlgl	stflife	freehms	health	
trstple	stfec0	eufff	rlgdgr	
trstplt	stfgov	imbgeco	wkdcorga	

La définition de ces variables, y compris le texte de la question, se trouve dans l'ESS (2013).

## Bibliographie

- Bates, D.M., Mächler, M., Bolker, B.M. et Walker, S.C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1, 1-48. <https://doi.org/10.18637/jss.v067.i01>.
- Bates, D.M., Mächler, M., Bolker, B.M. et Walker, S.C. (2019). *Lme4: Linear Mixed-Effects Models Using 'Eigen' and S4*. <https://CRAN.R-project.org/package=lme4>.
- Beullens, K., et Loosveldt, G. (2016). Interviewer effects in the European social survey. *Survey Research Methods*, 10, 2, 103-118.
- Biemer, P.P. (2010). Total survey error: Design, implementation, and evaluation. *Public Opinion Quarterly*, Oxford University Press, 74, 5, 817-848.
- Chambers, R.L., et Skinner, C.J. (2003). *Analysis of Survey Data*. New York: John Wiley & Sons, Inc.

- Chaudhuri, A., et Stenger, H. (2005). *Survey Sampling: Theory and Methods*. CRC Press.
- Davis, P., et Scott, A. (1995). La variance de l'intervieweur et ses effets sur les comparaisons de domaines. *Techniques d'enquête*, 21, 2, 111-118. Article accessible à l'adresse <https://www150.statcan.gc.ca/n1/pub/12-001-x/1995002/article/14405-fra.pdf>.
- Ellis, P.D. (2010). *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*. Cambridge University Press.
- European Social Survey (ESS) (2013). *ESS6 Data Protocol. 1.4*. Londres: ESS ERIC. <http://www.europeansocialsurvey.org/data/download.html?r=6>.
- European Social Survey (ESS) (2014a). *European Social Survey Round 6 Interviewer Questionnaire*. Dataset edition: 2.1. Londres: ESS ERIC.
- European Social Survey (ESS) (2014b). *Weighting European Social Survey Data*. Londres: ESS ERIC. [www.europeansocialsurvey.org/docs/methodology/ESS\\_weighting\\_data\\_1.pdf](http://www.europeansocialsurvey.org/docs/methodology/ESS_weighting_data_1.pdf).
- European Social Survey (ESS) (2014c). *ESS6 – 2012 Documentation Report*. Edition: 2.3. Londres: ESS ERIC. [http://www.europeansocialsurvey.org/docs/round6/survey/ESS6\\_data\\_documentation\\_report\\_e02\\_3.pdf](http://www.europeansocialsurvey.org/docs/round6/survey/ESS6_data_documentation_report_e02_3.pdf).
- European Social Survey (ESS) (2016). *European Social Survey Round 6 Data*. Dataset edition: 2.2. Londres: ESS ERIC.
- European Social Survey (ESS) (2018a). *Countries by Round (Year)*. Londres: ESS ERIC. [http://www.europeansocialsurvey.org/data/country\\_index.html](http://www.europeansocialsurvey.org/data/country_index.html).
- European Social Survey (ESS) (2018b). *Data and Documentation by Round European Social Survey (ESS)*. Londres: ESS ERIC. <http://www.europeansocialsurvey.org/data/download.html?r=6>.
- Fahrmeir, L., Heumann, C., Künstler, R., Pigeot, I. et Tutz, G. (1997). *Statistik: Der Weg Zur Datenanalyse*. 1<sup>ière</sup> éd. Berlin: Springer-Verlag.
- Fischer, M., West, B.T., Elliott, M.R. et Kreuter, F. (2018). The impact of interviewer effects on regression coefficients. *Journal of Survey Statistics and Methodology*, mai. <https://doi.org/10.1093/jssam/smy007>.
- Gabler, S., Häder, S. et Lahiri, P. (1999). Justification à base de modèle de la formule de Kish pour les effets de plan de sondage liés à la pondération et à l'effet de grappe. *Techniques d'enquête*, 25, 1, 119-120. Article accessible à l'adresse <https://www150.statcan.gc.ca/n1/pub/12-001-x/1999001/article/4718-fra.pdf>.
- Gabler, S., Häder, S. et Lynn, P. (2006). Effets de plan pour les échantillons à plans de sondage multiples. *Techniques d'enquête*, 32, 1, 127-133. Article accessible à l'adresse <https://www150.statcan.gc.ca/n1/fr/pub/12-001-x/2006001/article/9256-fra.pdf>.
- Gabler, S., et Lahiri, P. (2009). De la définition et de l'interprétation de la variabilité d'intervieweur pour un plan d'échantillonnage complexe. *Techniques d'enquête*, 35, 1, 91-106. Article accessible à l'adresse <https://www150.statcan.gc.ca/n1/fr/pub/12-001-x/2009001/article/10886-fra.pdf>.

- Ganninger, M. (2010). *Design Effects: Model-Based Versus Design-Based Approach*. Édité par GESIS - Leibniz-Institut für Sozialwissenschaften. Array 3.
- Genz, A., Bretz, F., Miwa, T., Mi, X. et Hothorn, T. (2019). *Mvtnorm: Multivariate Normal and T Distributions*. <https://CRAN.R-project.org/package=mvtnorm>.
- Groves, R.M. (2009). *Survey Methodology*. 2<sup>ième</sup>éd. Wiley Series in Survey Methodology. Hoboken, New York: John Wiley & Sons, Inc.
- Groves, R.M., et Lyberg, L. (2010). Total survey error: Past, present, and future. *Public Opinion Quarterly*, Oxford University Press, 74, 5, 849-879.
- Henningsen, A. (2012). *Linprog: Linear Programming/Optimization*. <https://CRAN.R-project.org/package=linprog>.
- Kish, L. (1962). Studies of interviewer variance for attitudinal variables. *Journal of the American Statistical Association*, 57, 297, 92-115. <https://doi.org/10.1080/01621459.1962.10482153>.
- Kish, L. (1965). *Survey Sampling*. New York: John Wiley & Sons, Inc.
- Lohr, S.L. (2014). Design effects for a regression slope in a cluster sample. *Journal of Survey Statistics and Methodology*, 2, 2, 97-125. <https://doi.org/10.1093/jssam/smu003>.
- Lynn, P., et Gabler, S. (2004). *Approximations to B\* in the Prediction of Design Effects Due to Clustering*. Série de documents de travail de l'ISER.
- Lynn, P., Häder, S., Gabler, S. et Laaksonen, S. (2007). Methods for achieving equivalence of samples in cross-national surveys: The European social survey experience. *Journal of Official Statistics*, 23, 1, 107.
- O'Muircheartaigh, C., et Campanelli, P. (1998). The relative impact of interviewer effects and sample design effects on survey precision. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 161, 1, 63-77.
- Raudenbush, S.W. (1993). A crossed random effects model for unbalanced data with applications in cross-sectional and longitudinal research. *Journal of Educational Statistics*, 18, 4, 321-349. <https://doi.org/10.2307/1165158>.
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienne, Autriche: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Särndal, C.-E., Swensson, B. et Wretman, J. (1992). *Model Assisted Survey Sampling*. New York: Springer-Verlag.
- Scheipl, F., Greven, S. et Kuechenhoff, H. (2008). Size and power of tests for a zero random effect variance or polynomial regression in additive and linear mixed models. *Computational Statistics & Data Analysis*, 52, 7, 3283-3299.
- Schnell, R., et Kreuter, F. (2005). Separating interviewer and sampling-point effects. *Journal of Official Statistics*, 21, 3, 389-410.

- The ESS Sampling Expert Panel (2016). *Sampling Guidelines: Principles and Implementation for the European Social Survey*. Londres: ESS ERIC Headquarters. [http://www.europeansocialsurvey.org/docs/round8/methods/ESS8\\_sampling\\_guidelines.pdf](http://www.europeansocialsurvey.org/docs/round8/methods/ESS8_sampling_guidelines.pdf).
- Vassallo, R., Durrant, G. et Smith, P. (2017). Separating interviewer and area effects by using a cross-classified multilevel logistic model: Simulation findings and implications for survey designs. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 180, 2, 531-550.
- Von Sanden, N.D. (2004). *Interviewer Effects in Household Surveys: Estimation and Design*. Thèse de doctorat, Wollongong: University of Wollongong. <http://ro.uow.edu.au/theses/312>.
- West, B.T., et Blom, A.G. (2017). Explaining interviewer effects: A research synthesis. *Journal of Survey Statistics and Methodology*, 5, 2, 175-211. <https://doi.org/10.1093/jssam/smw024>.
- West, B.T., et Elliott, M.R. (2014). Approches fréquentiste et bayésienne pour comparer les composantes de l'écart intervieweurs dans deux groupes d'intervieweurs d'enquête. *Techniques d'enquête*, 40, 2, 183-210. Article accessible à l'adresse <https://www150.statcan.gc.ca/n1/fr/pub/12-001-x/2014002/article/14092-fra.pdf>.