Rapport de méthodes

Estimation de la précision des évolutions dans l'enquête sur la valeur ajoutée



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Département fédéral de l'intérieur DFI
Office fédéral de la statistique OFS

Neuchâtel, 2009

Statistik der Schweiz

Die vom Bundesamt für Statistik (BFS) herausgegebene Reihe «Statistik der Schweiz» gliedert sich in folgende Fachbereiche:

0	Statistische Grundlagen und Übersichten
1	Bevölkerung
2	Raum und Umwelt
3	Arbeit und Erwerb
4	Volkswirtschaft
5	Preise
6	Industrie und Dienstleistungen
7	Land- und Forstwirtschaft
8	Energie
9	Bau- und Wohnungswesen
10	Tourismus
11	Mobilität und Verkehr
12	Geld, Banken, Versicherungen
13	Soziale Sicherheit
14	Gesundheit
15	Bildung und Wissenschaft
16	Kultur, Medien, Informationsgesellschaft, Sport
17	Politik
18	Öffentliche Verwaltung und Finanzen
19	Kriminalität und Strafrecht
20	Wirtschaftliche und soziale Situation der Bevölkerung
21	Nachhaltige Entwicklung und Disparitäten auf regionaler und internationaler Ebene

Statistique de la Suisse

La série «Statistique de la Suisse» publiée par l'Office fédéral de la statistique (OFS) couvre les domaines suivants:

0	Bases statistiques et produits généraux
1	Population
2	Espace et environnement
3	Vie active et rémunération du travail
4	Economie nationale
5	Prix
6	Industrie et services
7	Agriculture et sylviculture
8	Energie
9	Construction et logement
10	Tourisme
11	Mobilité et transports
12	Monnaie, banques, assurances
13	Protection sociale
14	Santé
15	Education et science
16	Culture, Médias, société de l'information, sport
17	Politique
18	Administration et finances publiques
19	Criminalité et droit pénal
20	Situation économique et sociale de la population
21	Développement durable et disparités régionales et internationales

Statistik der Schweiz

Methodenbericht

Statistique de la Suisse

Rapport de méthodes

Estimation de la précision des évolutions dans l'enquête sur la valeur ajoutée

Auteurs **Lionel Qualité, Yves Tillé**
Université de Neuchâtel

Editeur **Office fédéral de la statistique**

Neuchâtel, 2009

Préambule

L'enquête sur la valeur ajoutée (WertschöpfungsStatistik) est une enquête réalisée chaque année auprès des entreprises installées en Suisse. Elle permet de calculer des données structurelles et des ratios micro-économiques au niveau des branches, mais également d'estimer les principaux agrégats macro-économiques du compte de production, du compte d'exploitation et, en partie, du compte de revenu pour les comptes nationaux.

L'Institut de Statistique de l'Université de Neuchâtel a étudié la précision des estimations d'évolution obtenues à partir des différentes années d'enquête. La méthode utilisée et des résultats détaillés au niveau des secteurs d'activités économiques sont donnés dans ce rapport pour l'évolution de quatre variables d'intérêt entre 1999 et 2001: la valeur de production brute des entreprises, leur consommation intermédiaire, la création de valeur ajoutée, et le volume des salaires.

Le travail présenté ici a été mené dans le cadre d'une convention de collaboration entre l'Université de Neuchâtel et l'Office Fédéral de la Statistique (OFS). Nous remercions pour leur aide précieuse Paul-André Salamin, Jean-Pierre Renfer et Daniel Assoulin du service de méthodes statistiques (METH) ainsi que Veronica Roth et Markus Däppen de la section Statistique Monétaire des Entreprises (MON).

Résumé

L'Office fédéral de la statistique produit la statistique de la valeur ajoutée à partir d'un panel, c'est-à-dire d'un même échantillon utilisé pendant plusieurs années consécutives. L'usage de panels est une stratégie qui permet d'estimer précisément les évolutions.

Nous avons estimé la variance des estimateurs des évolutions en prenant en compte toutes les composantes de ceux-ci: le plan de sondage, le traitement des non-réponses, le traitement des grosses entreprises, la corrélation de la non-réponse d'une vague à l'autre, la robustification et le calage.

Les résultats montrent que si l'on prend en compte l'effet du panel, les variances des évolutions sont considérablement plus petites que les variances qui auraient été calculées sous une hypothèse d'indépendance entre deux vagues successives. Les variances sont en effet réduites de 60%. Les évolutions sont estimées de manière précise au niveau de l'ensemble des entreprises. Cependant, au niveau des classes d'activité définies par la Nomenclature générale des activités économiques (NOGA), à cause de la faible taille des échantillons nets, les mesures des évolutions sont beaucoup moins précises, à tel point que pour la moitié des strates d'activité, il n'est pas possible d'affirmer si l'évolution est croissante ou décroissante.

Mots-clé

Rapport de méthode, variance, panel, entreprises.

Complément d'information: Lionel Qualité, tél. 032 717 24 54
lionel.qualite@bfs.admin.ch

Réalisation: Service de méthodes statistiques, OFS

Diffusion: Office fédéral de la statistique
CH-2010 Neuchâtel
Tél. 032 713 60 60 / Fax 032 713 60 61
Order@bfs.admin.ch

Internet: http://www.statistik.admin.ch

Numéro de commande: 338-0056

Prix: gratuit

Série: Statistique de la Suisse

Domaine: 0 Bases statistiques et produits généraux

Langue du texte original: Français

Graphisme/Layout: OFS

Copyright: OFS, Neuchâtel 2009
La reproduction est autorisée, sauf à des fins commerciales,
si la source est mentionnée.

ISBN: 978-3-303-00423-4

Table des matières

1 Introduction **5**

2 L'enquête sur la valeur ajoutée **5**
2.1 Préambule . . . 5
2.2 Le plan de sondage . . . 5
2.3 Pondération et estimation de base . . . 6
2.4 Traitement des très grosses entreprises . . . 6
2.5 Poststratification . . . 7
2.6 Calage . . . 7
2.7 Robustification . . . 8
2.8 Variance des estimations transversales . . . 8

3 Variance de l'évolution de la valeur ajoutée **9**
3.1 Le problème . . . 9
3.2 Mesure de l'évolution . . . 9
3.2.1 Evolution et effet de structure . . . 9
3.2.2 Evolutions en volume et en pourcentage . . . 10
3.3 Aspects de l'enquête sur la valeur ajoutée . . . 10
3.4 Estimation de la covariance et application . . . 12
3.4.1 Covariance entre les estimateurs transversaux . . . 12
3.4.2 Variance des estimateurs d'évolution . . . 13
3.5 Autres estimateurs de l'évolution . . . 13

4 Résultats pour l'enquête sur la valeur ajoutée **15**
4.1 Estimateur de base . . . 15
4.1.1 Corrélation, taux de recouvrement . . . 15
4.1.2 Variance . . . 16
4.2 Utilisation de la partie commune . . . 18
4.3 Comparaison entre les estimations sur échantillons complets et les estimations sur échantillon commun . . . 19

5 Conclusions **22**

Annexes **23**

A Résultats détaillés **23**
A.1 Comparaison entre les estimations OFS et UNI . . . 23
A.2 Taux de recouvrement et corrélations . . . 25
A.3 Evolutions entre 1999, 2000 et 2001 - échantillons complets . . . 28
A.4 Evolutions entre 1999, 2000 et 2001 - échantillon commun . . . 32

B Estimation de la précision des évolutions dans les plans simples **36**
B.1 Introduction . . . 36
B.2 Estimation dans des échantillons disjoints . . . 36
B.3 Estimation par la différence des estimateurs transversaux dans des échantillons qui se chevauchent . . . 37
B.3.1 Le problème et la notation . . . 37
B.3.2 L'estimateur et sa variance . . . 38
B.3.3 Cas particuliers . . . 39

B.3.4 Estimation de la variance . . . 40
B.3.5 Gain par rapport à des échantillons indépendants . . . 40
B.4 Estimation de la différence au moyen de la partie commune . . . 41
B.4.1 L'estimateur et sa variance . . . 41
B.4.2 Comparaison avec la différence des estimateurs transversaux . . . 41
B.5 Estimation optimale . . . 41

C Prise en compte de la non-réponse totale **43**
C.1 Introduction . . . 43
C.2 Le modèle de non-réponse . . . 43
C.3 Interaction entre le plan de sélection et le plan de réponse . . . 44

D Estimateurs optimaux **46**
D.1 Estimateur optimal parmi les combinaisons linéaires de $\widehat{\Delta}_d$ et $\widehat{\Delta}_c$ . . . 46
D.2 Estimateur optimal parmi les combinaisons linéaires des estimateurs sur les sous-échantillons . . . 46

E Glossaire **48**

Références **50**

1 Introduction

L'enquête sur la valeur ajoutée est réalisée tous les ans par l'Office Fédéral de la Statistique. Il s'agit d'un panel d'établissements dont l'échantillon est remis à jour périodiquement. Cette enquête doit donc permettre de mesurer assez correctement des évolutions à un niveau agrégé, et entre des vagues d'enquêtes qui utilisent les mêmes échantillons.

L'objet de ce rapport est de présenter la précision avec laquelle on pense estimer les évolutions des variables principales de l'enquête entre 1999, 2000 et 2001. Dans un premier temps on ne s'intéresse qu'aux estimateurs longitudinaux naturels, qui ont été utilisés jusqu'à présent : les différences entre les estimations transversales produites à chaque vague d'enquête. Le calcul de ces précisions est rendu difficile par les techniques particulières utilisées pour améliorer les estimations annuelles de la valeur ajoutée. Puis nous proposons d'autres estimateurs plus adaptés aux enquêtes répétées, mais qui présentent le désavantage de ne pas fournir des valeurs cohérentes avec les estimations transversales publiées.

Les deux premiers chapitres de ce rapport sont consacrés à une présentation de l'enquête sur la valeur ajoutée, et des méthodes utilisées pour estimer les précisions. Dans le troisième chapitre, on présente et on discute des résultats numériques obtenus . En annexe, on trouve des tableaux de résultats plus complets, et les fondements théoriques de ce travail : un développement sur le calcul de la variance dans des plans simples répétés, et une partie sur la prise en compte de la non-réponse totale. Les notations utilisées sont rappelées à la fin du document.

2 L'enquête sur la valeur ajoutée

2.1 Préambule

Dans ce chapitre, on propose un rappel de la méthodologie de l'enquête sur la valeur ajoutée. Ce chapitre est basé sur les rapports de méthodes de Renfer (2000) et de Peters et al. (2001). Il sert à définir clairement la notation et à rappeler tous les éléments à prendre en compte pour calculer la précision des estimateurs. Un estimateur de variance est présenté pour les estimations transversales.

2.2 Le plan de sondage

En 1999, un échantillon de 11'210 exploitations, de deux personnes ou plus, a été sélectionné et reconduit en 2000 et 2001. A ces trois vagues, respectivement 5'913, 5'957 et 5'897 exploitations ont répondu.

L'échantillon est stratifié par division économique, définie par les deux premiers chiffres de la nomenclature NOGA et par taille de l'entreprise. Pour ce faire, les activités économiques NOGA ont été reprises du fichier du recensement des entreprises de 1995 et intégrées dans le Registre des Entreprises et Etablissements qui sert de base de sondage. Les établissements qui avaient répondu à l'enquête de 1995 sur la valeur ajoutée ont automatiquement été sélectionnés.

Dans chaque strate d'activité NOGA2ST trois strates de taille sont constituées : les petites entreprises employant 2-19 personnes en Equivalent Temps Plein (ETP), les moyennes entreprises, de 20 à M ETP, et les grandes entreprises de plus de M ETP. La borne M est choisie différemment dans chaque strate d'activité afin de minimiser le coefficient de variation.

La strate contenant les grandes entreprises est recensée, tandis que les petites et moyennes entreprises sont sélectionnées aléatoirement avec des taux de sondage différents. On note $U_h, h = 1 \ldots, H$, les strates de tirage et leurs effectifs totaux sont $N_h = \mathrm{card}\, U_h$. Dans chaque strate, un échantillon de taille brute n_h^b est sélectionné. Après l'enquête, on n'observe réellement qu'un échantillon s_h de répondants, de taille n_h. Dans certaines NOGA2ST, les strates (de taille) comptant peu d'exploitations sont regroupées à posteriori.

2.3 Pondération et estimation de base

La pondération de base dépend de deux facteurs : la probabilité de sélection et la probabilité de réponse. On fait l'hypothèse d'une non-réponse uniforme dans chaque strate. Alors, quitte à raisonner conditionnellement à la taille de l'échantillon final, le poids de base d'une exploitation k répondante dans la strate h est :

$$w_k^0 = \frac{N_h}{n_h}.$$

Ce poids peut se décomposer en un produit de l'inverse de la probabilité d'inclusion N_h/n_h^b et de l'inverse de la probabilité de réponse estimée n_h^b/n_h.

L'estimateur de base est donc

$$\widehat{Y}_{HT} = \sum_{h=1}^{H} N_h \bar{y}_h = \sum_{h=1}^{H} \frac{N_h}{n_h} \sum_{k \in s_h} y_k,$$

où $\bar{y}_h$ est la moyenne des données observées dans la strate h.

2.4 Traitement des très grosses entreprises

Pour chaque NOGA2ST, les entreprises de taille supérieure à M sont recensées. Cependant toutes les entreprises n'ont pas accepté de répondre et il a été décidé à posteriori de traiter certaines très grosses entreprises séparément. On peut penser que le taux de réponse a été meilleur pour les entreprises les plus grosses qui ont une structure administrative plus apte à répondre aux questions de l'enquête. Leur appliquer un poids égal à celui des autres grandes entreprises introduirait un biais ainsi qu'une variabilité trop importante.

De ce fait, une poststrate "surprise" $s4$ a été créée, qui contient les 5% d'exploitations les plus grandes dans le fichier d'enquête. Ces dernières ont alors été considérées comme recensées et ont reçu le poids 1. Aucun autre traitement (calage, robustification) ne leur sera appliqué. Les strates de tirage de petites, moyennes et grandes entreprises ont été mises à jour. Si l'on accepte l'hypothèse un peu optimiste que les très grandes entreprises ont effectivement été recensées, l'estimateur qui en découle est sans biais, et la variance liée aux très grandes entreprises est nulle. On peut donc procéder, lors du calcul de la précision, comme si la strate $s4$ n'existait pas, et calculer uniquement la variance dans les autres strates mises à jour. Il serait toutefois souhaitable de vérifier que le taux de réponse est proche de 100% pour les très grandes entreprises, et d'essayer d'évaluer la sensibilité des résultats à une non-réponse même faible.

Les poids deviennent :

$$w_k^1 = \frac{N_h - n_h^4}{n_h - n_h^4} 1_{t_k < \ell(s)} + 1_{t_k \geq \ell(s)},$$

où n_h^4 est le nombre d'établissements de la strate h dont la taille t_k dépasse la limite $\ell(s)$ à partir de laquelle ils sont classés dans la strate $s4$.

L'estimateur est maintenant :

$$\widehat{Y} = \sum_{h=1}^{H} \frac{N_h - n_h^4}{n_h - n_h^4} \sum_{k \in s_h \setminus s4} y_k + \sum_{k \in s4} y_k = \widehat{Y}_{HT}^{\setminus 4} + y^4,$$

où $\widehat{Y}_{HT}^{\setminus 4}$ est l'estimateur de base du total de Y sur l'échantillon privé de $s4$, et y^4 est le total de Y sur la strate $s4$.

2.5 Poststratification

Lors de l'enquête, la catégorie d'activité économique était redemandée aux entreprises. Les estimations sont réalisées au niveau de ces NOGA2 déclarées et non au niveau des NOGA2ST de la base de sondage.

Pour la poststrate g, l'estimateur du total s'écrit :

$$\widehat{Y}_g = \sum_{h=1}^{H} \frac{N_h - n_h^4}{n_h - n_h^4} \sum_{k \in s_h \setminus s4} y_k 1_{k \in g} + \sum_{k \in s4} y_k 1_{k \in g} = \widehat{Y}_{HT,g}^{\setminus 4} + y_g^4.$$

2.6 Calage

Un calage sur le nombre d'équivalent temps plein (ETP) donné par la STATEM est ensuite réalisé au niveau des poststrates NOGA2 au moyen d'un estimateur par le quotient, en utilisant le nombre d'ETP relevé lors de l'enquête. L'estimateur du total dans la poststrate g devient

$$\widehat{Y}_g^Q = \widehat{R}_g^{\setminus 4}(X_g - x_g^4) + y_g^4,$$

où

$$\widehat{R}_g^{\setminus 4} = \frac{\displaystyle\sum_{h=1}^{H} \left(\frac{N_h - n_h^4}{n_h - n_h^4} \sum_{k \in s_h \setminus s4} y_k 1_{k \in g} \right)}{\displaystyle\sum_{h=1}^{H} \left(\frac{N_h - n_h^4}{n_h - n_h^4} \sum_{k \in s_h \setminus s4} x_k 1_{k \in g} \right)} = \frac{\widehat{Y}_{HT,g}^{\setminus 4}}{\widehat{X}_{HT,g}^{\setminus 4}},$$

x_k est le nombre d'ETP dans l'entreprise k, X_g est le nombre total d'ETP dans la NOGA2 g qui est connu par ailleurs, et x_g^4 est le nombre d'ETP déclarés par les entreprises de la poststrate $g \cap s4$.

Le poids d'une unité dans la poststrate $g \cap s_h \setminus s4$ s'écrit alors

$$w_k^2 = \frac{N_h - n_h^4}{n_h - n_h^4} \frac{X_g - x_g^4}{\widehat{X}_{HT,g}^{\setminus 4}},$$

soit dans le cas général :

$$w_k^2 = 1_{t_k \geq \ell(s)} + \left(\sum_{g=1}^{G} \sum_{h=1}^{H} \frac{N_h - n_h^4}{n_h - n_h^4} \frac{X_g - x_g^4}{\widehat{X}_{HT,g}^{\setminus 4}} 1_{k \in s_h} 1_{k \in g} \right) 1_{t_k < \ell(s)}.$$

2.7 Robustification

Une technique de robustification a été utilisée pour écréter la distribution de certaines variables dans l'échantillon des petites, moyennes et grandes entreprises (voir Hulliger, 1999; Peters et al., 2001). Les poids de sondage ont été modifiés par un facteur $u_k(s)$ compris entre 0 et 1. Les établissements dont les valeurs sont considérées comme normales reçoivent un facteur $u_k(s)$ égal à 1. Ceux dont les valeurs sont considérées comme aberrantes voient leur poid diminué par ce facteur.

La nouvelle pondération est :

$$w_k^3 = 1_{t_k \geq \ell(s)} + \left(\sum_{g=1}^{G} \sum_{h=1}^{H} \frac{N_h - n_h^4}{n_h - n_h^4} \frac{X_g - x_g^4}{\widehat{XU}_{HT,g}^{\setminus 4}} u_k 1_{k \in s_h} 1_{k \in g} \right) 1_{t_k < \ell(s)},$$

où $\widehat{XU}_{HT,g}^{\setminus 4}$ est l'estimateur standard du total de $x_k u_k$ dans la strate g, en dehors des très grandes entreprises. L'estimateur du total de Y pour la strate g devient :

$$\widehat{Y}_g^{Qr} = \frac{\widehat{YU}_{HT,g}^{\setminus 4}}{\widehat{XU}_{HT,g}^{\setminus 4}} (X_g - x_g^4) + y_g^4 \tag{1}$$

et l'estimateur du total dans la population est :

$$\widehat{Y}^{Qr} = \sum_{g=1}^{G} \widehat{Y}_g^{Qr}. \tag{2}$$

2.8 Variance des estimations transversales

Si l'impact du calage sur la variance de l'estimateur peut facilement être calculé, l'utilisation d'une procédure de robustification et la création de la poststrate surprise $s4$, dont la définition dépend de l'échantillon s, nécessitent quelques concessions et approximations.

La variance liée à la poststrate $s4$ est nulle si l'on accepte l'hypothèse que les très grandes entreprises ont toutes répondu.

La variance de l'estimateur robustifié peut être approchée en faisant l'hypothèse classique que les poids $u_k(s)$ ne dépendent que faiblement de l'échantillon s qui a été tiré. Il suffit alors de remplacer les variables x_k et y_k observées par $u_k x_k$ et $u_k y_k$ dans les estimateurs de variance.

Enfin, la variance d'un estimateur par le ratio peut être estimée soit par linéarisation, soit par une technique classique de résidus (voir à ce sujet Woodruff, 1971; Binder et Patak, 1994; Deville et Särndal, 1992; Deville, 1999). Le résidu estimé est

$$ue_k = uy_k - \frac{\widehat{YU}_{HT,g}^{\setminus 4}}{\widehat{XU}_{HT,g}^{\setminus 4}} ux_k, k \in g. \tag{3}$$

Ensuite ce résidu est inséré dans la formule de variance classique d'un plan stratifié,

$$\widehat{\mathrm{var}}(\widehat{Y}_g^{Q,r}) = \sum_{h=1}^{H} (N_h - n_h^4) \frac{N_h - n_h}{n_h - n_h^4} \frac{1}{n_h - n_h^4 - 1} \sum_{k \in g \cap s_h \setminus s4} (ue_k - \overline{ue}_{g \cap s_h \setminus s4})^2.$$

L'OFS a utilisé la technique de linéarisation pour calculer les variances. Nous privilégierons souvent l'approche 'résidus', asymptotiquement équivalente, mais qui donne des résultats numériques légèrement différents. Elle est plus adaptée aux calculs que nous allons réaliser pour les estimateurs longitudinaux.

3 Variance de l'évolution de la valeur ajoutée

3.1 Le problème

L'objectif est d'estimer correctement la variance des estimateurs d'évolution de la valeur ajoutée (voir Renfer, 2000; Peters et al., 2001). En calculant les variances sous une hypothèse d'indépendance des échantillons, on surestime largement les variances des évolutions. Une prise en compte correcte de tous les aspects du plan de sondage et du redressement devrait fournir des estimateurs crédibles. Les travaux portent sur les vagues d'enquête 1999, 2000 et 2001. Entre ces trois dates, l'échantillon brut n'a pas été modifié. Le fait que l'échantillon soit resté fixe devrait permettre d'estimer de manière fiable les évolutions, mais un taux de réponse proche de 50% peut faire perdre tout le bénéfice du panel, pour peu que le nombre de répondants communs aux vagues successives soit faible. Le cas de l'évolution entre deux vagues d'enquête où il y a eu une mise à jour de l'échantillon, et donc deux échantillons bruts et deux populations de référence différents, constitue un tout autre problème.

Dans le cas présent, plusieurs raisons contribuent conjointement à l'obtention de faibles variances :

1. *Plan optimal* : Le plan de sondage a été optimisé. Selon la stratification optimale, les grosses entreprises ont des probabilités d'inclusion plus élevées. La strate d'entreprises contribuant le plus à la valeur ajoutée est recensée. Les estimateurs transversaux ont pour cette raison une faible variance.
2. *Fraction de sondage élevée* : Dans la strate recensée des grandes entreprises, le taux de réponse avoisine 70%. La correction de population finie $(N-n)/N$ peut donc diviser la variance par 3.
3. *Effet panel* : L'échantillon est un panel, ce qui est la meilleure stratégie pour estimer des évolutions.
4. *Corrélation de la non-réponse* : La non-réponse à une vague est très liée à la vague suivante, et ne dégrade donc pas beaucoup le panel.
5. *Corrélation des variables entre les vagues* : Les variables valeurs ajoutées aux temps t et $t+1$ sont très corrélées, car il s'agit de la même variable mesurée à deux moments différents.
6. *Calage :* Les estimateurs sont calés dans les strates sur une variable liée à la variable d'intérêt, la variance des estimateurs peut alors s'écrire comme une variance résiduelle.

3.2 Mesure de l'évolution

3.2.1 Evolution et effet de structure

Les expressions de base, pour le calcul de la variance des plans en deux vagues, sont contenues dans l'annexe B. Les spécificités des estimateurs utilisés pour l'enquête sur la valeur ajoutée impliquent la mise en place d'une méthode d'estimation particulière.

Un traitement adapté devrait pouvoir séparer l'évolution directe de l'effet de structure. L'effet de structure est dû à l'évolution de la population (mort et naissance). L'évolution directe est l'évolution des valeurs pour les entreprises présentes aux deux vagues. L'enquête est un panel remis à jour tous les trois ans. Les créations d'entreprises sont donc assez mal prises en compte. Malheureusement, on ne sait pas si une absence de réponse est due à une mort ou à un simple refus. L'effet de structure ne peut donc pas être mesuré, mais il est en partie pris en compte du fait du calage.

Il serait souhaitable, dans la mesure du possible, de pouvoir différencier une non-réponse due à la disparition d'une entreprise d'un simple refus de réponse.

Faute de pouvoir faire mieux, on supposera donc que la population est fixe, qu'il n'y a ni morts, ni naissances, et qu'une absence de réponse est toujours due à un refus.

3.2.2 Evolutions en volume et en pourcentage

Si l'on note $\widehat{Y}_t$ (resp. $\widehat{Y}_{t-1}$) l'estimateur du total (2) de Y construit à partir de l'échantillon s_t (resp. s_{t-1}) obtenu à la date t (resp. $t-1$), l'évolution peut être mesurée de plusieurs manières
– La différence $\widehat{\Delta} = \widehat{Y}_t - \widehat{Y}_{t-1}$,
– L'évolution relative $\widehat{\Delta}_R = (\widehat{Y}_t - \widehat{Y}_{t-1})/\widehat{Y}_{t-1} = \widehat{Y}_t/\widehat{Y}_{t-1} - 1$,
– Le quotient $\widehat{Q} = \widehat{Y}_t/\widehat{Y}_{t-1}$.

La variance de $\widehat{\Delta}$ peut s'exprimer simplement en fonction des variances de $\widehat{Y}_t$ et $\widehat{Y}_{t-1}$ et de leur covariance. On peut approcher la variance de $\widehat{\Delta}_R$ et $\widehat{Q}$ par linéarisation, et l'estimer au moyen de $\widehat{Q}$, $\widehat{\mathrm{var}}(\widehat{Y}_t)$, $\widehat{\mathrm{var}}(\widehat{Y}_{t-1})$ et $\widehat{\mathrm{cov}}(\widehat{Y}_{t-1}, \widehat{Y}_t)$.

3.3 Aspects de l'enquête sur la valeur ajoutée

Nous décrivons dans cette partie comment les différents aspects de l'enquête sur la valeur ajoutée ont été pris en considération lors du calcul des variances.

Non-réponse

Sur les 11210 entreprises sélectionnées en 1999, 5162 ont répondu à la fois en 1999 et en 2000 ; et 5330 ont répondu aux vagues 2000 et 2001. La taille du panel est donc relativement modeste, et le traitement de la non-réponse va avoir une grande influence sur les résultats.

Afin de réaliser des estimations de variance, nous avons fait l'hypothèse que la non-réponse est ignorable (missing completely at random) au sein des strates de tirage. Autrement dit, le phénomène de non-réponse peut être assimilé à un plan bernoullien dans chaque strate. Dans l'annexe C, le Résultat 2, page 45 montre que, si le plan de sélection est simple et que la non-réponse est bernoullienne, le plan et la non-réponse peuvent être confondus comme s'il n'y avait qu'un seul plan bernoullien, ce qui simplifie considérablement le problème. Ce résultat reste vrai même si la non-réponse est liée d'une vague à l'autre. L'inférence est ensuite réalisée conditionnellement aux tailles des échantillons obtenues. On peut alors procéder comme si on avait affaire à des plans simples. On retombe dans le cas facile et bien connu décrit dans l'annexe B.

Très grosses entreprises

La taille d'entreprise utilisée est celle déclarée à chaque vague de l'enquête, et donc plusieurs facteurs sont à prendre en compte lors des estimations longitudinales : la limite de la strate $s4$ varie d'une vague à l'autre, même si elle reste proche de 424 ETP. Une entreprise peut également entrer ou sortir de $s4$ si elle est non-répondante à une vague, ou si son effectif change.

En 1999, 295 entreprises appartenaient à la strate $s4$, 46 d'entre elles n'ont pas répondu en 2000, et 12 ont répondu mais n'étaient plus considérées comme des très grandes entreprises. En 2000, La strate $s4$ compte 297 individus dont 237 étaient déjà dans $s4$ en 1999, 29 n'avaient pas répondu et 31 avaient répondu mais n'appartenaient pas à $s4$. Les échanges entre strates sont du même ordre entre les vagues 2000 et 2001. Les mouvements entre la strate $s4$ et les autres strates sont donc assez importants.

TABLE 1 Non-Réponse et appartenance à la strate $s4$ en 1999, 2000 et 2001

$99\backslash 00$	N.R.	s123	s4	$00\backslash 01$	N.R.	s123	s4
N.R.	4502	766	29	N.R.	4686	536	31
s123	705	4882	31	s123	600	5041	19
s4	46	12	237	s4	27	12	258

Plusieurs scenarii sont possibles pour calculer une variance malgré ces changements de strate. Une très courte comparaison des résultats fournis par deux de ces options, ainsi que par d'autres estimateurs de la corrélation sont disponibles dans Qualité et Tillé (2006).

– On peut choisir de négliger ces problèmes, et de considérer qu'une entreprise qui est dans $s4$ à l'une ou l'autre des deux vagues n'apporte pas de variance à la vague où elle est dans $s4$ ni de covariance entre une autre vague et celle-ci. C'est le cas idéal où il n'y a effectivement pas de non-réponse pour les très grandes entreprises à l'une au moins des deux vagues. C'est l'option qui a été retenue dans ce rapport.

– On peut décider de ne négliger que la variance liée aux entreprises qui sont dans la strate $s4$ aux deux vagues consécutives de l'enquête et de réintégrer les autres très grandes entreprises aux strates de petites moyennes et grandes entreprises. De cette façon, la variance liée à une bonne partie des grandes entreprises est totalement éliminée. Mais le fait de replacer certaines très grandes entreprises dans leur strate d'origine introduit un peu de variance, de sorte que la précision estimée est moins optimiste que celle du cas précédent.

– Une autre option est de simuler les résultats en fixant un taux de non-réponse pour les très grandes entreprises et de regarder l'influence que cela a sur les résultats. Il faudra alors essayer de calculer la variance malgré les traitements spécifiques à l'enquête, par exemple en modifiant les variables d'intérêt de manière adéquate et en utilisant une technique de linéarisation.

Poststratification

A chaque vague, le plan est poststratifié sur les NOGA2 déclarées. Cela implique la possibilité d'un changement de poststrate qu'il faudrait essayer de prendre en compte dans les estima-

tions longitudinales. Nous avons décidé de négliger l'impact de ces changements dans un premier temps, et de considérer, pour l'estimation de la covariance, que les strates sont fixes et données par la valeur déclarée à la première des deux vagues consécutives.

Cette simplification n'est pas abusive, dans la mesure où seules 30 entreprises (resp. 25) ont changé de strates entre 1999 et 2000 (resp. 2000 et 2001), ce qui représente moins de 0.5% (resp. 0.2%) des ETP de l'échantillon.

Calage

Le calage est réalisé chaque année pour les résultats de l'année en cours, et peut être pris en compte au moyen d'une technique de résidus, conformément à l'expression (3). Pour estimer la variances des différences, on utilisera simplement les résidus $e_k(s_t)$ calculés au temps t et $e_k(s_{t-1})$ calculés au temps $t-1$. Ces résidus sont calculés pour chaque vague à l'aide des variables auxiliaires $X(t)$ et $X(t-1)$ mesurées respectivement au temps t et au temps $t-1$. Une présentation plus détaillée de ces techniques et de leur application à l'enquête WS est disponible dans Qualité et Tillé (2008).

Robustification

Comme pour l'estimation de la précision des estimateurs transversaux, on prend en compte la robustification en pondérant les variables de l'enquête par les variables $u_k(s_t)$ et $u_k(s_{t-1})$. On calcule ensuite les estimateurs de variance pour les évolutions des variables pondérées.

3.4 Estimation de la covariance et application

3.4.1 Covariance entre les estimateurs transversaux

Soient, $y(t)$ et $y(t-1)$ les variables observées, dont on cherche à estimer le total. Conditionnellement aux tailles d'échantillons $n(s_t)$, $n(s_{t-1})$ et $n(s_t \cap s_{t-1})$, la covariance de l'estimateur de Horvitz-Thompson du total de y dans une strate h peut être estimée par (voir Annexe B.3) :

$$\widehat{\mathrm{cov}}_{HT}(\widehat{Y}_h(s_t), \widehat{Y}_h(s_{t-1})) = N_h^2 \left(\frac{n_h(s_t \cap s_{t-1})}{n_h(s_t) n_h(s_{t-1})} - \frac{1}{N_h} \right) s_h(y(t), y(t-1)).$$

Donc, si l'on fait l'hypothèse que les très grosses entreprises sont effectivement recensées et sont les mêmes aux vagues t et $t-1$, la covariance entre les estimateurs définis par (2) du total de $y(t)$ et $y(t-1)$ est estimable par :

$$\widehat{\mathrm{cov}}(\widehat{Y}(s_t), \widehat{Y}(s_{t-1})) = \sum_{h=1}^{H} N_{h\backslash s4}^2 \left(\frac{n_{h\backslash s4}(s_t \cap s_{t-1})}{n_{h\backslash s4}(s_t) n_{h\backslash s4}(s_{t-1})} - \frac{1}{N_{h\backslash s4}} \right) s_{h\backslash s4}(ue(t), ue(t-1)).$$

et dans le domaine d'activité g on a :

$$\widehat{\mathrm{cov}}(\widehat{Y}_g(s_t), \widehat{Y}_g(s_{t-1})) = \sum_{h=1}^{H} N_{h\backslash s4}^2 \left(\frac{n_{h\backslash s4}(s_t \cap s_{t-1})}{n_{h\backslash s4}(s_t) n_{h\backslash s4}(s_{t-1})} - \frac{1}{N_{h\backslash s4}} \right) s_{h\backslash s4}(ue1_g(t), ue1_g(t-1)).$$

où

– $ue1_g(t)$ est la variable $\{u_k(s_t) * e_k(s_t) * 1_{k \in g}\}_{k \in s_t \cap s_{t-1}}$,
– $s_{h\backslash s4}(ue1_g(t), ue1_g(t-1))$ est la covariance corrigée estimée entre $ue1_g(t)$ et $ue1_g(t-1)$ dans la strate $h\backslash s4$, sur la partie commune de l'échantillon $s_t \cap s_{t-1}$.

3.4.2 Variance des estimateurs d'évolution

On peut estimer la variance de la différence des estimateurs aux dates t et $t-1$ par :

$$\widehat{\text{var}}(\widehat{\Delta}) = \sum_{h=1}^{H} \left\{ \widehat{\text{var}}(\widehat{Y}_h(s_t)) + \widehat{\text{var}}(\widehat{Y}_h(s_{t-1})) - 2\widehat{\text{cov}}(\widehat{Y}_h(s_t), \widehat{Y}_h(s_{t-1})) \right\}.$$

Les variances des estimateurs $\widehat{\Delta}_R$ et $\widehat{Q}$ définis en (3.2.2) sont égales. On peut les estimer par linéarisation :

$$\widehat{Q} \approx \frac{Y_t}{Y_{t-1}} + \frac{1}{Y_{t-1}} \left(\widehat{Y(s_t)} - Y_t \right) - \frac{Y_t}{Y_{t-1}^2} \left(\widehat{Y(s_{t-1})} - Y_{t-1} \right),$$

et donc

$$\text{var}(\widehat{Q}) \approx \frac{1}{Y_{t-1}^2} \sum_{h=1}^{H} \left\{ \text{var}(\widehat{Y}_h(s_t)) + Q^2 \text{var}(\widehat{Y}_h(s_{t-1})) - 2Q\text{cov}(\widehat{Y}_h(s_t), \widehat{Y}_h(s_{t-1})) \right\}.$$

Ce qui justifie que :

$$\widehat{\text{var}}(\widehat{Q}) \approx \frac{1}{\widehat{Y}^2(s_{t-1})} \sum_{h=1}^{H} \left\{ \widehat{\text{var}}(\widehat{Y}_h(s_t)) + \widehat{Q}^2 \widehat{\text{var}}(\widehat{Y}_h(s_{t-1})) - 2\widehat{Q}\widehat{\text{cov}}(\widehat{Y}_h(s_t), \widehat{Y}_h(s_{t-1})) \right\}.$$

Les résultats de ces calculs appliqués à certaines variables de l'enquête sur la valeur ajoutée se trouvent dans le chapitre suivant.

3.5 Autres estimateurs de l'évolution

On présente ici des estimateurs spécifiques aux estimations longitudinales. Seul le premier d'entre eux, qui n'utilise que les entreprises ayant répondu aux deux vagues d'enquête sera effectivement calculé dans ce rapport, avec sa variance. Les autres feront l'objet d'une prochaine étude.

Evolution estimée au moyen de la partie commune

Dès que la corrélation entre y_t et y_{t-1} est suffisamment bonne, il est plus précis d'utiliser uniquement la partie commune aux deux échantillons plutôt que d'utiliser la différence des deux estimateurs transversaux (voir annexe B.4). Soit $\widehat{\Delta}_c = \widehat{Y}^{\backslash 4}(s_t \cap s_{t-1}) - \widehat{Y}^{\backslash 4}(s_{t-1} \cap s_t)$, l'estimateur $\widehat{\Delta}_c + (y_g^4(t) - y_g^4(t-1))$ a pour variance

$$\widehat{\text{var}}(\widehat{\Delta}_c) = \sum_{h=1}^{H} N_{h\backslash s4} \frac{N_{h\backslash s4} - n_{h\backslash s4}(s_t \cap s_{t-1})}{n_{h\backslash s4}(s_t \cap s_{t-1})} s^2_{h\backslash s4} \left[ue(t) - ue(t-1)\right], \tag{4}$$

où $s^2_{h\backslash s4}\left[ue(t) - ue(t-1)\right] = s^2_{h\backslash s4}[ue(t-1)] + s^2_{h\backslash s4}[ue(t)] - 2s_{h\backslash s4}[ue(t-1), ue(t)]$. Cet estimateur est positif lorsque les variables $ue(t)$ et $ue(t-1)$ sont observées sur le même échantillon aux temps t et $t-1$. Comme ce n'est en général pas le cas, l'estimateur peut prendre des valeurs négatives. Berger (2004) a étudié des estimateurs positifs pour les variances des évolutions. La variance de $\widehat{\Delta}_c$ est typiquement bien plus petite que celle de $\widehat{\Delta}$ car les résultats sont mieux corrélés entre les mêmes individus qu'entre deux échantillons qui ne se recouvrent que partiellement.

Combinaison linéaire optimale des estimateurs transversaux

Soit $\widehat{\Delta}_d = \widehat{Y}^{\backslash 4}(s_t \backslash s_{t-1}) - \widehat{Y}^{\backslash 4}(s_{t-1} \backslash s_t)$ la différence estimée uniquement à partir de la partie propre à l'échantillon s_t et de la partie propre à l'échantillon s_{t-1}. Les estimateurs appartenant à la famille

$$\alpha\widehat{\Delta}_d + (1-\alpha)\widehat{\Delta}_c + y_g^4(t) - y_g^4(t-1) \qquad \alpha \in [0,1]$$

sont sans biais, et on obtient la variance minimale lorsque

$$\alpha = \frac{\mathrm{var}(\widehat{\Delta}_c) - \mathrm{cov}(\widehat{\Delta}_c, \widehat{\Delta}_d)}{\mathrm{var}(\widehat{\Delta}_c) + \mathrm{var}(\widehat{\Delta}_d) - 2\mathrm{cov}(\widehat{\Delta}_c, \widehat{\Delta}_d)}.$$

Elle vaut alors (voir annexe D page 46) :

$$\mathrm{var}(\widehat{\Delta}_c) - \frac{\left[\mathrm{var}(\widehat{\Delta}_c) - \mathrm{cov}(\widehat{\Delta}_c, \widehat{\Delta}_d)\right]^2}{\mathrm{var}(\widehat{\Delta}_c) + \mathrm{var}(\widehat{\Delta}_d) - 2\mathrm{cov}(\widehat{\Delta}_c, \widehat{\Delta}_d)}.$$

On peut également trouver un estimateur sans biais de variance minimale parmi toutes les combinaisons linéaires de $\widehat{Y}^{\backslash 4}(s_t \cap s_{t-1})$, $\widehat{Y}^{\backslash 4}(s_{t-1} \cap s_t)$, $\widehat{Y}^{\backslash 4}(s_t \backslash s_{t-1})$ et $\widehat{Y}^{\backslash 4}(s_{t-1} \backslash s_t)$ (voir annexe D page 46).

Estimateur d'évolution calé

Au lieu de caler les estimateurs du total de Y par strate aux dates t et $t-1$ sur le total d'une variable X à ces mêmes dates, on peut caler les évolutions de Y sur les évolutions de X (voir Deville et Särndal, 1992). L'estimateur de l'évolution devient alors pour une strate g :

$$\widehat{\Delta}_g^{cale} = y_g^4(s_t) - y_g^4(s_{t-1}) + [X_g^{\backslash 4}(t) - X_g^{\backslash 4}(t-1)]\frac{\widehat{Y}_{HT,g}^{\backslash 4}(s_t) - \widehat{Y}_{HT,g}^{\backslash 4}(s_{t-1})}{\widehat{X}_{HT,g}^{\backslash 4}(s_t) - \widehat{X}_{HT,g}^{\backslash 4}(s_{t-1})}. \qquad (5)$$

La variance de ce nouvel estimateur sera d'autant plus faible que les évolutions de y sont bien corrélées avec les évolutions de x. En pratique cependant, l'estimateur par le ratio utilisé par l'OFS dans l'enquête WS semble avoir des performances comparables : la corrélation entre les évolutions de la valeur ajoutée et la taille des entreprises est au moins aussi bonne que la corrélation entre les évolutions de valeur ajoutée et les évolutions d'effectif.

Estimateur d'évolution robustifié

On peut également appliquer une technique de robustification aux entreprises présentes à deux vagues d'enquêtes successives, afin d'écrêter des évolutions jugées aberrantes. Soit $\widehat{\Delta}_c^{robust}(Y_g)$ (resp. $\widehat{\Delta}_c^{robust}(X_g)$) un estimateur robustifié de l'évolution de Y (resp. X) dans la strate g, sur la partie de l'échantillon commune aux deux vagues. Un tel estimateur peut être construit, à l'instar des estimateurs transversaux utilisés pour cette enquête, en diminuant le poids des entreprises qui ont connu une évolution considérée comme anormale. On peut alors estimer l'évolution de Y dans la strate g par :

$$\widehat{\Delta}_g^{robust} = y_g^4(t) - y_g^4(t-1) + [X_g^{\backslash 4}(t) - X_g^{\backslash 4}(t-1)]\widehat{R}_g^{robust}. \qquad (6)$$

où

$$\widehat{R}_g^{robust} = \frac{\alpha\widehat{\Delta}_d(Y_g) + (1-\alpha)\widehat{\Delta}_c^{robust}(Y_g)}{\beta\widehat{\Delta}_d(X_g) + (1-\beta)\widehat{\Delta}_c^{robust}(X_g)} \qquad \alpha, \beta \in [0,1].$$

On peut choisir les paramètres α et β pour minimiser la variance, ou, pour simplifier, les prendre égaux à zéro.

4 Résultats pour l'enquête sur la valeur ajoutée

Les résultats portent sur les variables T017 (valeur de production brute), T022 (consommation intermédiaire), T053 (salaires) et WS (valeur ajoutée $T017 - T022$).

4.1 Estimateur de base

4.1.1 Corrélation, taux de recouvrement

Comme on le montre dans l'annexe B page 40, lorsque l'on veut estimer l'évolution d'une variable Y entre deux dates, le gain de précision d'un panel par rapport à un échantillonnage indépendant dépend du taux de recouvrement q (défini en annexe) des échantillons et de la corrélation ρ entre les variables aux deux dates. Ce gain peut être approché par $1 - \rho q$.

Voici les résultats obtenus pour une partie des strates de tirage (le tableau complet est en annexe). Les corrélations sont estimées dans les strates d'enquête.

TABLE 2 Taux de recouvrement des échantillons et corrélations entre 2000 et 2001

noga2st	taille	Tx recouvrement	T017	T022	T053	WS
14	1	0.90	0.86	0.76	0.98	0.92
14	2	0.84	0.94	0.95	0.84	0.94
14	3	0.77	0.92	0.93	0.97	0.95
15	1	0.79	0.98	0.97	0.95	0.91
15	2	0.78	0.99	0.99	0.91	0.95
15	3	0.65	0.98	0.98	0.93	0.92
16	1	0.00	0.00	0.00	0.00	0.00
16	3	1.00	0.96	0.95	1.00	0.91
17	1	0.71	0.81	0.88	0.48	0.68
17	2	0.87	0.92	0.95	0.98	0.70
17	3	0.74	0.96	0.92	0.95	0.97
18	1	0.85	1.00	1.00	1.00	1.00
18	2	0.56	0.99	0.99	0.99	0.97
18	3	0.61	0.99	0.98	0.96	0.96
19	1	0.79	0.98	0.96	0.90	0.97
19	2	0.80	0.00	0.88	0.94	0.99
19	3	0.67	0.88	0.83	0.94	0.94
20	1	0.00	0.00	0.00	0.00	0.00
20	2	0.82	0.91	0.81	0.96	0.64
20	3	0.74	0.97	0.95	0.82	0.93
21	1	0.89	0.92	0.74	0.87	0.84
21	2	0.93	0.98	0.98	0.82	0.97
21	3	0.66	0.90	0.89	0.86	0.88
22	1	0.96	0.93	0.97	0.76	0.87
22	2	0.90	1.00	0.99	0.95	0.94
22	3	0.85	0.96	0.95	0.89	0.94
⋮	⋮	⋮	⋮	⋮	⋮	⋮
Valeur moyenne		0.74	0.81	0.82	0.80	0.80

On peut donc espérer voir la variance multipliée par $1 - 0.74 \times 0.80 \approx 0.41$ par rapport à l'hypothèse d'indépendance, et les écarts-types diminués d'un tiers. Les résultats sont similaires entre les vagues 1999 et 2000.

4.1.2 Variance

Les résultats sont présentés par poststrate NOGA pour T017 et WS.

TABLE 3 Evolution des variables T017 et WS, écart-type (Std.) - échantillons complets (en millions de francs suisses)

NOGA	t017v99	tO17v00	Evol.	Std.	signif.	WSv99	WSv00	Evol.	Std.	signif.
14	1832	1568	-14,4	9,1	0	802	750	-6,5	5,7	0
15	23190	23209	0,1	2,9	0	5790	5971	3,1	1,9	0
16	2418	3176	31,4	1,9	1	726	1048	44,4	3,4	1
17	3144	3497	11,2	5,2	1	1205	1368	13,5	5,5	1
18	1386	1598	15,3	6,5	1	444	422	-4,9	5,4	0
19	395	547	38,7	9,7	1	123	73	-40,4	6,7	1
20	6320	6547	3,6	6,6	0	2882	2901	0,6	6,2	0
21	4650	5203	11,9	2,2	1	1768	1807	2,2	2,3	0
22	10812	11354	5,0	4,8	0	5303	5754	8,5	5,6	0
24	34747	39002	12,2	3,1	1	11654	12016	3,1	4,0	0
25	6587	7055	7,1	2,7	1	2546	2726	7,1	2,2	1
26	4270	4805	12,5	4,6	1	1964	2106	7,2	3,9	0
27	5137	5696	10,9	3,1	1	1616	1673	3,5	5,4	0
28	16250	17425	7,2	2,9	1	8001	8412	5,1	2,3	1
29	29585	32355	9,4	1,9	1	11628	12280	5,6	1,7	1
30	809	1188	46,9	16,7	1	311	378	21,7	9,2	1
31	13551	14315	5,6	2,2	1	4890	4335	-11,4	2,1	1
32	5763	5982	3,8	7,2	0	2143	2210	3,1	4,7	0
33	20417	23211	13,7	5,3	1	8430	9304	10,4	2,1	1
34	1292	1348	4,4	2,7	0	452	484	7,1	3,0	1
35	3207	3406	6,2	3,6	0	1089	1271	16,6	2,4	1
36	5358	5678	6,0	4,2	0	2258	2184	-3,3	3,6	0
37	979	1030	5,2	7,2	0	367	362	-1,5	5,0	0
40	25538	22424	-12,2	2,0	1	10515	8451	-19,6	2,0	1
41	373	335	-10,0	1,9	1	263	226	-13,9	2,9	1
45	42760	45891	7,3	2,2	1	21428	23332	8,9	1,6	1
50	9037	9006	-0,4	3,5	0	5939	6074	2,3	3,3	0
51	36963	39977	8,2	2,7	1	24513	26329	7,4	3,2	1
52	31316	31220	-0,3	1,7	0	21502	21143	-1,7	2,0	0
55	13682	13596	-0,6	5,3	0	6866	6745	-1,8	4,6	0
60	6239	6103	-2,2	4,2	0	3542	3267	-7,8	4,6	0
61	284	352	24,3	7,8	1	148	194	31,3	8,2	1
62	6865	8057	17,4	1,4	1	2022	2043	1,0	2,2	0
63	10066	11491	14,2	4,8	1	3819	4310	12,8	2,4	1
64	3199	13717	328,7	47,1	1	1073	7072	558,8	36,7	1
70	6690	6673	-0,3	7,5	0	4115	4411	7,2	8,7	0
71	879	964	9,7	***	***	517	398	-23,0	10,0	1
72	11441	11194	-2,2	5,2	0	6185	6790	9,8	5,0	0
73	1900	2403	26,5	7,3	1	933	1073	14,9	8,7	0
74	41276	40499	-1,9	4,7	0	25061	25746	2,7	4,2	0
80	2050	2086	1,8	9,8	0	1077	1178	9,3	12,1	0
85	3227	3394	5,2	2,3	1	2341	2485	6,1	3,3	0
90	4045	3884	-4,0	3,5	0	1782	1836	3,0	2,9	0
91	1323	1603	21,1	8,5	1	598	594	-0,7	5,1	0
92	7971	7415	-7,0	15,6	0	2646	2618	-1,1	6,2	0
93	2420	2475	2,3	2,4	0	1598	1565	-2,0	2,4	0
Total	471640	503954	6,9	0,8	1	224877	237714	5,7	0,8	1

Le signe des évolutions est significatif dans la moitié des strates. C'est cohérent pour un panel de 5'000 individus répartis en plus de 40 strates. Si l'on omet la strate 14, les évolutions

importantes (à deux chiffres), semblent fiables.

TABLE 4 Estimation de l'évolution brute des variables T017 et WS, écart-type (Std.) - échantillons complets (en millions de francs suisses)

	Entre 1999 et 2000				entre 2000 et 2001			
	T017		WS		T017		WS	
NOGA	Evol.	Std.	Evol.	Std.	Evol.	Std.	Evol.	Std.
14	-264	192	-52	48	1	58	18	43
15	19	662	181	107	554	460	30	92
16	759	38	322	17	-351	8	58	2
17	353	178	163	64	-452	92	-138	40
18	212	88	-22	24	-33	79	16	18
19	152	29	-50	11	-160	54	45	21
20	227	405	19	178	181	353	44	211
21	553	98	39	40	-200	123	-34	50
22	542	484	451	267	101	470	-265	160
24	4255	1044	362	458	4450	1140	527	498
25	468	170	181	54	-24	229	92	51
26	535	219	142	76	-8	161	33	79
27	559	145	57	84	-541	118	129	43
28	1175	463	410	180	648	541	99	183
29	2770	521	652	190	-2339	672	-820	188
30	379	96	67	25	350	53	102	14
31	764	290	-555	105	1403	159	-579	79
32	219	402	67	99	641	147	371	73
33	2794	1078	874	171	3356	522	1223	215
34	57	34	32	13	118	35	6	12
35	199	110	181	23	765	69	-55	19
36	320	221	-74	83	-45	***	25	***
37	51	68	-5	19	274	102	54	25
40	-3114	538	-2064	252	3278	476	1069	152
41	-37	6	-36	7	274	14	221	14
45	3131	920	1905	336	-790	859	-353	391
50	-32	318	135	198	1029	445	819	351
51	3015	965	1816	768	-3177	599	-2256	418
52	-96	530	-360	431	1929	338	774	285
55	-86	729	-121	317	823	435	373	259
60	-137	265	-275	176	1167	560	759	211
61	69	22	46	11	-28	22	-7	13
62	1192	97	21	44	-3030	115	-600	40
63	1425	471	491	86	-470	378	-27	85
64	10517	340	5998	55	-1117	522	-1456	***
70	-17	503	296	400	1327	438	962	396
71	85	***	-119	163	324	***	233	***
72	-247	606	604	296	2132	481	1430	239
73	503	123	139	77	420	80	173	41
74	-777	1982	685	1044	9686	2796	5643	1728
80	36	200	101	135	-53	237	92	98
85	167	71	144	74	558	70	392	55
90	-161	147	53	52	283	150	193	52
91	280	116	-4	30	203	119	167	83
92	-557	1332	-28	167	1391	1240	173	275
93	55	57	-33	39	178	111	138	71
Total	32314	3675	12837	1757	25028	3892	9892	2124

4.2 Utilisation de la partie commune

Dans cette partie, les évolutions et la précision des estimateurs ont été calculées en n'utilisant que les entreprises qui ont répondu à deux vagues consécutives de l'enquête. La précision de ces estimateurs est souvent meilleure car la perte d'information due au fait que l'on écarte certaines entreprises est en général largement compensée par une bien meilleure corrélation entre les estimateurs des totaux aux différentes vagues.

TABLE 5 Evolution des variables T017 et WS ecart-type - échantillon commun (en millions de francs suisses)

	Entre 1999 et 2000				entre 2000 et 2001			
	T017		WS		T017		WS	
NOGA	Evol.	Std.	Evol.	Std.	Evol.	Std.	Evol.	Std.
14	-43	59	-40	36	27	43	16	24
15	761	288	115	71	100	245	1	77
16	831	15	362	10	-351	8	58	2
17	277	187	66	58	-430	140	-148	39
18	-33	20	-47	16	48	24	6	15
19	152	18	-63	9	-147	72	37	29
20	-353	67	-143	198	172	***	72	137
21	561	39	62	32	-76	75	-19	36
22	456	385	276	240	-244	198	-340	127
24	350	203	-924	121	2593	436	277	164
25	460	108	160	62	-22	108	69	36
26	520	183	200	54	9	235	66	75
27	1072	48	161	18	335	55	72	18
28	1365	436	406	158	406	529	115	188
29	2690	506	622	196	-1611	480	-724	153
30	99	28	11	10	454	34	154	12
31	722	171	-547	62	1867	93	-532	43
32	502	222	84	105	1077	179	418	79
33	1985	224	744	113	2828	307	1156	215
34	71	27	40	12	137	28	-6	9
35	-85	48	-29	17	-89	74	-33	20
36	-10	125	-178	68	211	***	145	***
37	135	42	28	14	314	14	86	4
40	-3561	469	-2308	250	2802	185	1484	111
41	-38	3	-35	5	280	11	217	11
45	2944	708	1935	330	-1142	830	-217	378
50	111	353	275	212	674	303	570	234
51	2327	664	1180	514	-2833	426	-1835	184
52	1253	384	551	313	1144	306	748	251
55	-986	550	-465	293	735	315	524	193
60	-18	328	-131	96	715	186	575	133
61	37	11	19	7	-20	8	-2	8
62	1265	85	-12	48	-463	114	26	28
63	1170	233	543	65	-5	202	42	75
64	7342	222	2129	24	-1302	87	-1403	***
70	-155	509	296	449	1410	441	1072	407
71	395	***	-61	***	427	***	245	***
72	142	397	268	269	2362	242	1609	203
73	509	144	140	86	431	66	193	32
74	3446	1029	2025	675	5888	2749	3856	1659
80	-250	128	-76	43	339	83	148	50
85	225	73	191	75	522	68	356	49
90	-264	90	-92	49	832	76	275	32
91	133	36	-15	22	86	42	58	41
92	716	88	1	50	-666	376	-288	133

voir page suivante

TABLE 5 – évolution de T017 et WS, échantillon commun

	T017		WS		T017		WS	
NOGA	Evol.	Std.	Evol.	Std.	Evol.	Std.	Evol.	Std.
93	39	49	-33	37	164	116	107	75
Total	29265	1736	7694	1281	19991	3114	9304	1889

Les écarts-types des estimateurs sont meilleurs que dans le cas des estimations sur échantillons complets, mais les résultats obtenus peuvent être assez différents suivant les strates.

4.3 Comparaison entre les estimations sur échantillons complets et les estimations sur échantillon commun

Les évolutions $\widehat{Y}_t/\widehat{Y}_{t-1}$ par strate, avec leurs intervalles de confiance à 95%, sont représentées dans les graphiques ci-dessous. Les chiffres détaillés se trouvent dans l'annexe A. Une évolution est jugée significative lorsque la barre verticale correspondant à l'intervalle de confiance de l'estimation ne coupe pas la ligne horizontale $\widehat{Y}_t/\widehat{Y}_{t-1} = 1$.

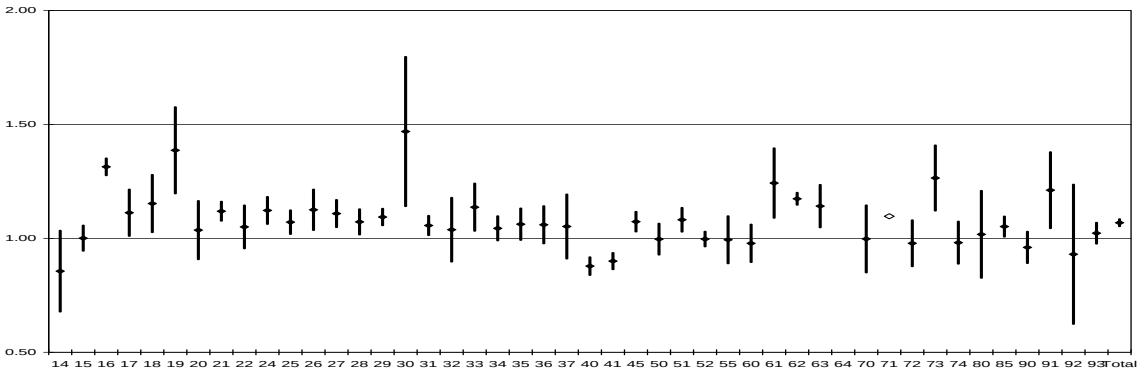


FIGURE 1 T017, 2000/1999, fichier complet

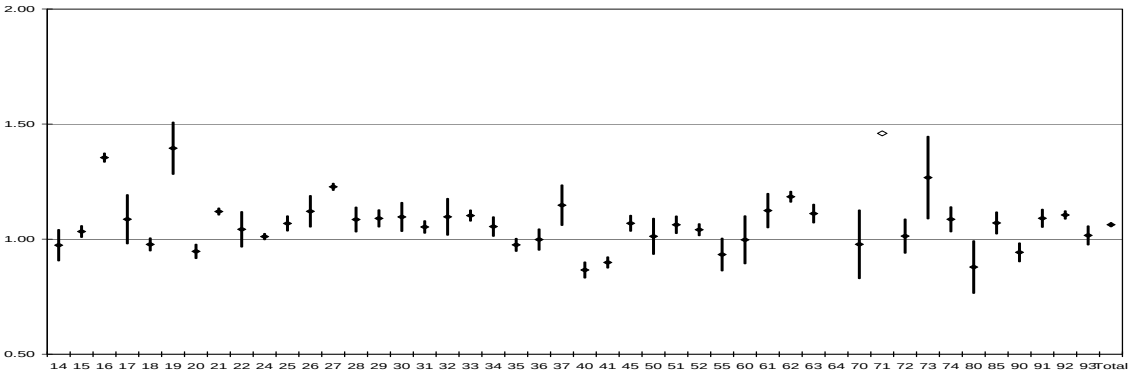


FIGURE 2 T017, 2000/1999, échantillon commun

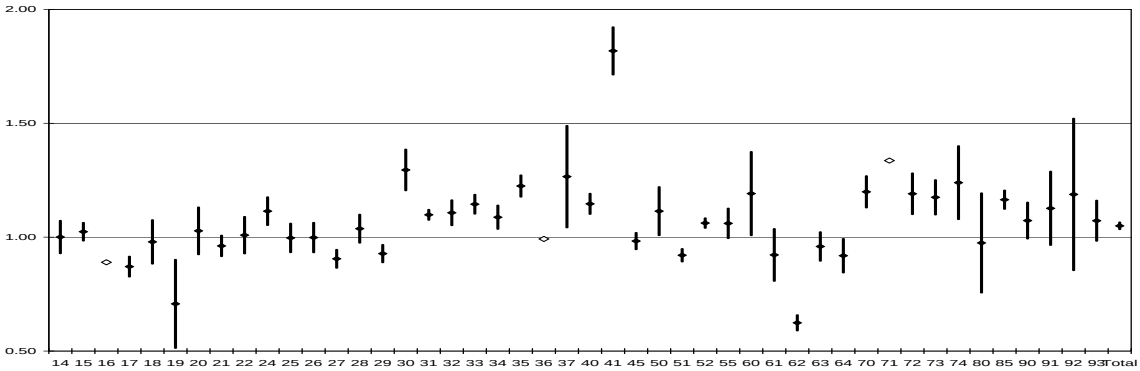


FIGURE 3 T017, 2001/2000, fichier complet

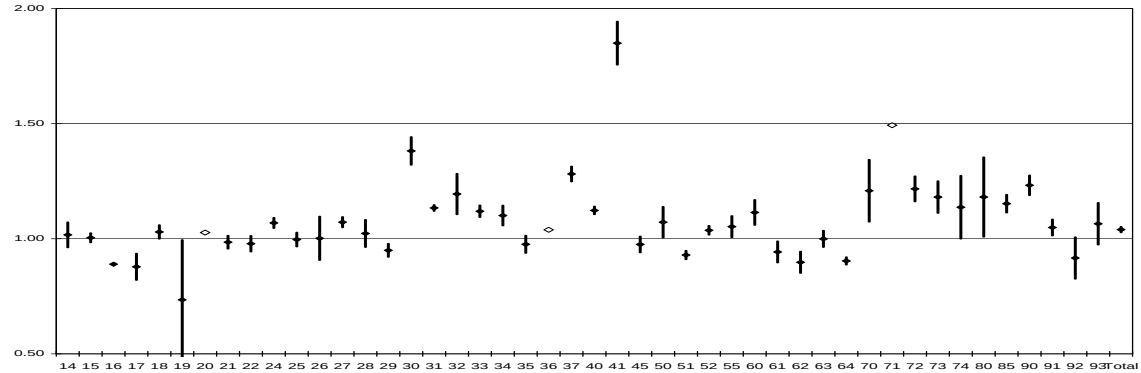


FIGURE 4 T017, 2001/2000, échantillon commun

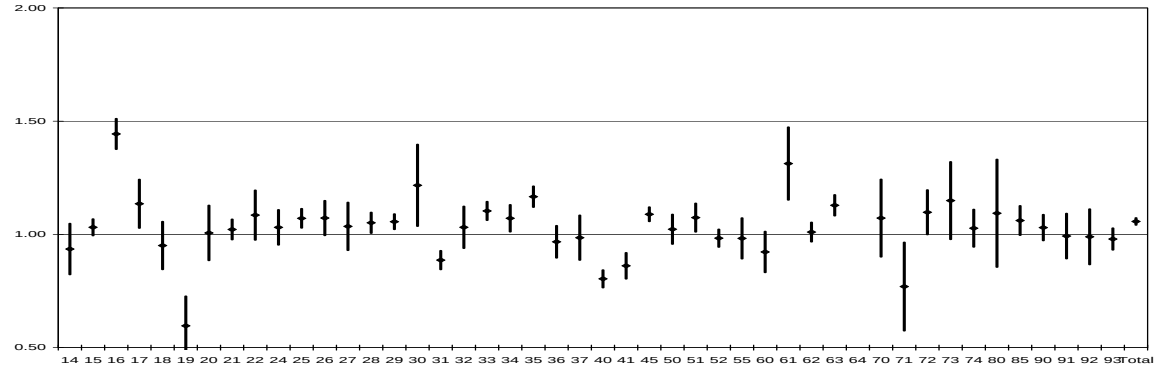


FIGURE 5 WS, 2000/1999, fichier complet

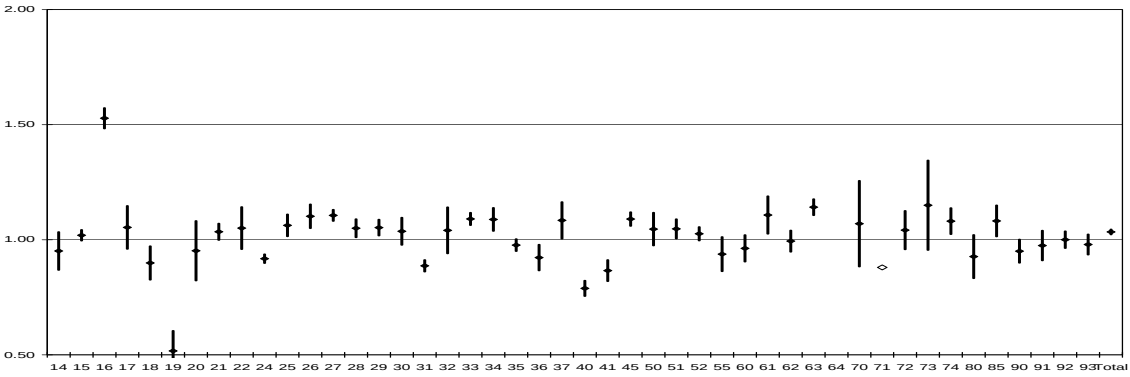


FIGURE 6 WS, 2000/1999, échantillon commun

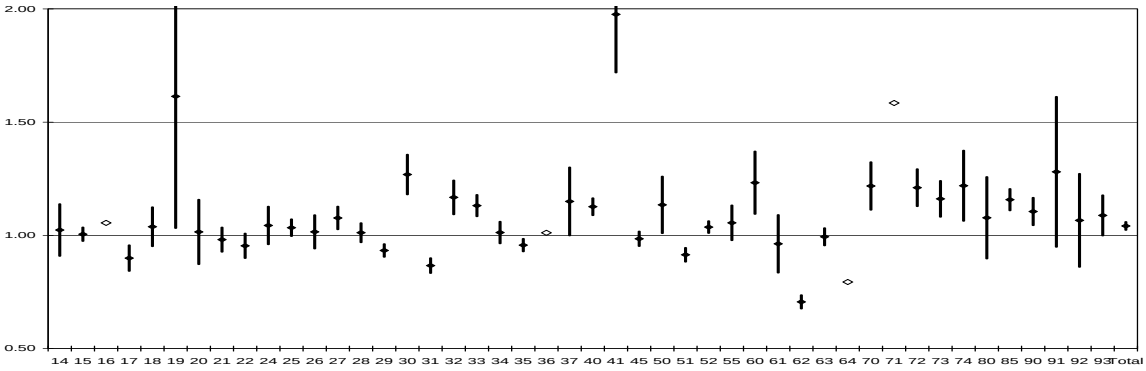


FIGURE 7 WS, 2001/2000, fichier complet

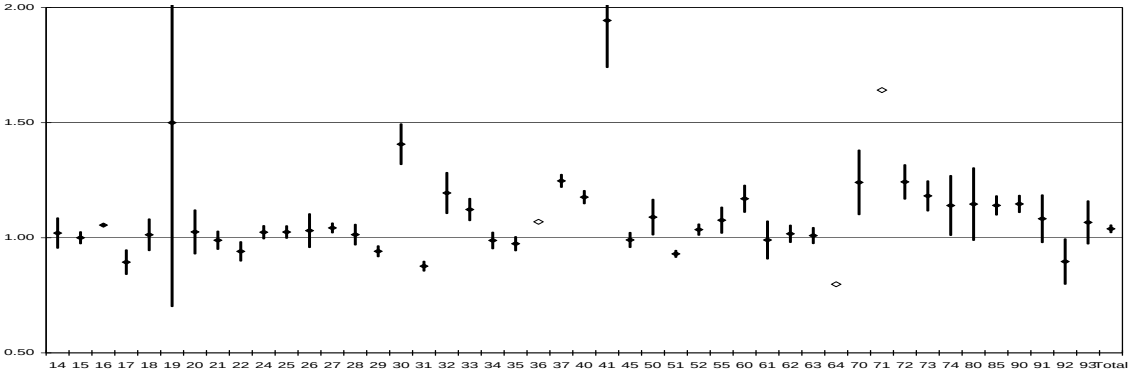


FIGURE 8 WS, 2001/2000, échantillon commun

Les intervalles de confiance des deux estimateurs se chevauchent dans quasiment toutes les strates. On observe que les estimations basées sur la partie commune de l'échantillon sont, pour la plupart, dans l'intervalle de confiance à 95% autour de l'estimateur utilisant les échantillons complets. Cela n'empêche pas d'avoir des différences parfois importantes entre ces deux estimateurs.

5 Conclusions

L'enquête sur la valeur ajoutée utilise le même échantillon pendant plusieurs années consécutives. De ce fait, elle devrait permettre d'estimer correctement des évolutions. C'est effectivement le cas lorsque l'on cherche à mesurer des évolutions sur l'ensemble des entreprises. Par contre, au niveau des classes d'activité définies par la NOGA, les mesures sont beaucoup moins précises. On ne détecte à ce niveau que les fortes variations. Il y a plusieurs raisons à cela.

1. L'échantillon compte 11'210 entreprises. A chaque vague, il y a environ 6'000 répondants, mais seules un peu plus de 5'100 entreprises répondent à deux années consécutives. Or, ce sont les établissements qui répondent chaque année qui fournissent l'information la plus fiable pour mesurer les évolutions. La non-réponse est donc une source de variabilité particulièrement gênante pour les estimations longitudinales.

2. Les évolutions sont beaucoup plus difficiles à mesurer que les niveaux. En effet, même avec un très bon échantillon on peut difficilement détecter de faibles variations.

3. L'échantillon a été conçu pour réaliser de bonnes estimations transversales : il est constitué de strates d'entreprises d'activités et de tailles homogènes, de ce fait les niveaux des variables fortement liées à la taille sont peu variables dans chaque strate. Par contre, rien ne garantit une telle homogénéité pour les évolutions de ces variables.

4. Les estimateurs sont conçus pour l'évaluation des niveaux et non des variations : le calage et la robustification, ainsi que la création d'une poststrate de très grosses entreprises permettent d'avoir de bons estimateurs pour une date donnée. Mais pour mesurer les évolutions il est préférable d'utiliser d'autres estimateurs, plus fiables, mais qui ne donnent pas les mêmes résultats que la simple comparaison des chiffres produits entre deux années.

Il faut en outre noter qu'on ne peut distinguer une disparition d'entreprise d'une non-réponse classique. Plus généralement, on ne peut suivre les évolutions de structure de la population des entreprises qu'au travers de la variable de calage : le nombre d'Equivalents Temps Plein par branche d'activité.

Au final, on ne peut en général affirmer qu'une des variables d'intérêt dans l'enquête a connu une évolution significative que pour la moitié des strates d'activité. Et souvent on ne peut pas donner un chiffre fiable pour cette évolution, tout juste dire s'il s'agit d'une hausse ou d'une baisse. L'utilisation d'estimateurs plus adaptés, n'utilisant que la partie de l'échantillon commune à deux vagues successives, permet de gagner sensiblement en précision, sans forcément que ce gain ne permette d'avoir d'estimations précises pour toutes les strates. D'autres estimateurs pourraient être utilisés et ont été proposés dans ce rapport. Leur efficacité fera l'objet d'un travail ultérieur.

A Résultats détaillés

A.1 Comparaison entre les estimations OFS et UNI

Les estimations transversales ont été recalculées en suivant la méthodologie utilisée par l'ofs. Pour simplifier et accélerer le travail, nous avons fait quelques approximations (par exemple sur la définition de la strate surprise de très grandes entreprises). De ce fait, les résultats sont très légèrement différents dans certaines strates. Nous avons recalculé les variances en utilisant un estimateur différent de celui de l'ofs, mais asymptotiquement équivalent. Ceci explique aussi quelques différences sur les écarts types.

TABLE 6 Variable T017 estimée en 1999 et 2000 (en millions de francs suisses)

	Chiffres OFS				Chiffres UNI			
NOGA	t017v99	t017v00	Std. t017v00	Evol. in%	t017v99	t017v00	Std. t017v00	Evol. in%
14	1832	1568	84	-14,4	1832	1568	94	-14,4
15	23170	23179	864	0,0	23190	23209	962	0,1
16	2418	3176	28	31,4	2418	3176	28	31,4
17	3144	3497	293	11,2	3144	3497	337	11,2
18	1386	1598	129	15,3	1386	1598	120	15,3
19	395	547	25	38,7	395	547	23	38,7
20	6320	6547	521	3,6	6320	6547	501	3,6
21	4652	5203	119	11,8	4650	5203	125	11,9
22	10831	11354	524	4,8	10812	11354	583	5,0
24	34747	39001	1374	12,2	34747	39002	1002	12,2
25	6591	7056	250	7,1	6587	7055	254	7,1
26	4270	4805	428	12,5	4270	4805	531	12,5
27	5137	5696	140	10,9	5137	5696	119	10,9
28	16250	17418	533	7,2	16250	17425	584	7,2
29	29585	32355	832	9,4	29585	32355	858	9,4
30	882	1188	101	34,6	809	1188	51	46,9
31	13545	14320	308	5,7	13551	14315	310	5,6
32	5763	5982	223	3,8	5763	5982	273	3,8
33	20457	23293	1404	13,9	20417	23211	1163	13,7
34	1292	1349	36	4,4	1292	1348	37	4,4
35	3207	3406	195	6,2	3207	3406	89	6,2
36	5358	5678	268	6,0	5358	5678	255	6,0
37	979	1030	92	5,2	979	1030	84	5,2
40	25538	22343	1071	-12,5	25538	22424	665	-12,2
41	373	335	15	-10,0	373	335	21	-10,0
45	42720	45867	1453	7,4	42760	45891	1685	7,3
50	9037	9006	413	-0,4	9037	9006	512	-0,4
51	36944	39908	1367	8,0	36963	39977	1554	8,2
52	31279	31192	731	-0,3	31316	31220	571	-0,3
55	13651	13577	488	-0,5	13682	13596	601	-0,6
60	6461	6103	334	-5,5	6239	6103	297	-2,2
61	284	352	23	24,3	284	352	26	24,3
62	6865	8057	104	17,4	6865	8057	142	17,4
63	10066	11507	616	14,3	10066	11491	647	14,2
64	3199	13780	437	330,7	3199	13717	271	328,7
70	6690	6673	1321	-0,3	6690	6673	1628	-0,3
71	879	964	195	9,7	879	964	181	9,7
72	11441	11194	586	-2,2	11441	11194	416	-2,2
73	1900	2403	76	26,5	1900	2403	163	26,5
74	41040	40522	3201	-1,3	41276	40499	3028	-1,9
80	2050	2086	404	1,8	2050	2086	370	1,8
85	3227	3394	75	5,2	3227	3394	81	5,2

voir page suivante

TABLE 6 – Variable T017 estimée en 1999 et 2000

NOGA	t017v99	t017v00	Std. t017v00	Evol. in%	t017v99	t017v00	Std. t017v00	Evol. in%
90	4045	3864	346	-4,5	4045	3884	149	-4,0
91	1323	1603	255	21,1	1323	1603	215	21,1
92	7971	7415	429	-7,0	7971	7415	374	-7,0
93	2420	2475	137	2,3	2420	2475	111	2,3
Total	471613	503864	5150	6,8	471640	503954	5041	6,9

TABLE 7 Variable T022 estimée en 1999 et 2000 (en millions de francs suisses)

	Chiffres OFS				Chiffres UNI			
NOGA	t022v99	t022v00	Std. t022v00	Evol. in%	t022v99	t022v00	Std. t022v00	Evol. in%
14	1031	819	55	-20,6	1031	819	62	-20,6
15	17389	17217	782	-1,0	17409	17238	870	-1,0
16	1692	2128	25	25,8	1692	2128	25	25,8
17	1940	2129	247	9,7	1940	2129	285	9,7
18	942	1176	93	24,8	942	1176	87	24,8
19	272	474	19	74,3	272	474	18	74,3
20	3438	3646	323	6,1	3438	3646	343	6,1
21	2883	3396	93	17,8	2881	3396	98	17,9
22	5514	5600	344	1,6	5510	5600	384	1,6
24	23093	26982	856	16,8	23093	26986	625	16,9
25	4046	4332	216	7,1	4041	4329	219	7,1
26	2306	2701	387	17,1	2306	2701	480	17,1
27	3535	4024	123	13,8	3535	4024	104	13,8
28	8249	9016	446	9,3	8249	9013	487	9,3
29	17984	20079	681	11,6	17984	20079	702	11,6
30	576	809	91	40,6	498	809	47	62,6
31	8656	9980	246	15,3	8660	9980	248	15,2
32	3620	3771	173	4,2	3620	3771	211	4,2
33	12022	14017	1250	16,6	11987	13957	1035	16,4
34	840	865	31	3,1	840	864	31	2,9
35	2118	2135	167	0,8	2118	2135	76	0,8
36	3107	3493	187	12,4	3107	3493	188	12,4
37	612	668	71	9,2	612	668	65	9,2
40	15023	13918	878	-7,4	15023	13973	545	-7,0
41	110	109	24	-0,7	110	109	33	-0,7
45	21297	22536	1098	5,8	21332	22559	1282	5,8
50	3103	2932	156	-5,5	3103	2932	193	-5,5
51	12447	13630	535	9,5	12455	13653	622	9,6
52	9799	10064	326	2,7	9816	10078	255	2,7
55	6762	6832	285	1,0	6815	6851	351	0,5
60	3087	2836	209	-8,2	2698	2836	194	5,1
61	135	158	18	16,6	135	158	20	16,6
62	4843	6014	116	24,2	4843	6014	159	24,2
63	6247	7198	612	15,2	6247	7181	642	15,0
64	2126	6706	391	215,5	2126	6645	243	212,6
70	2575	2262	544	-12,2	2575	2262	670	-12,2
71	362	566	145	56,5	362	566	152	56,5
72	5292	4404	458	-16,8	5292	4404	325	-16,8
73	968	1331	43	37,4	968	1331	93	37,5
74	16234	14816	1608	-8,7	16239	14794	1520	-8,9
80	974	914	159	-6,2	974	914	146	-6,2
85	891	913	58	2,4	891	913	63	2,4
90	2263	2036	342	-10,0	2263	2049	147	-9,5
91	746	1024	191	37,2	746	1024	162	37,2

voir page suivante

TABLE 7 – Variable T017 estimée en 1999 et 2000

NOGA	t022v99	t022v00	Std. t022v00	Evol. in%	t022v99	t022v00	Std. t022v00	Evol. in%
92	5330	4800	371	-9,9	5330	4800	324	-9,9
93	822	909	126	10,6	822	909	104	10,6
Total	247299	266367	3282	7,7	246929	266370	3159	7,9

A.2 Taux de recouvrement et corrélations

TABLE 8 Taux de recouvrement des échantillons et corrélations entre 2000 et 2001

NOGA2ST	TAILLE	Tx recouvrement	T017	T022	T053	WS
14	1	0.90	0.86	0.76	0.98	0.92
14	2	0.84	0.94	0.95	0.84	0.94
14	3	0.77	0.92	0.93	0.97	0.95
15	1	0.79	0.98	0.97	0.95	0.91
15	2	0.78	0.99	0.99	0.91	0.95
15	3	0.65	0.98	0.98	0.93	0.92
16	1	0.00	0.00	0.00	0.00	0.00
16	3	1.00	0.96	0.95	1.00	0.91
17	1	0.71	0.81	0.88	0.48	0.68
17	2	0.87	0.92	0.95	0.98	0.70
17	3	0.74	0.96	0.92	0.95	0.97
18	1	0.85	1.00	1.00	1.00	1.00
18	2	0.56	0.99	0.99	0.99	0.97
18	3	0.61	0.99	0.98	0.96	0.96
19	1	0.79	0.98	0.96	0.90	0.97
19	2	0.80	0.00	0.88	0.94	0.99
19	3	0.67	0.88	0.83	0.94	0.94
20	1	0.00	0.00	0.00	0.00	0.00
20	2	0.82	0.91	0.81	0.96	0.64
20	3	0.74	0.97	0.95	0.82	0.93
21	1	0.89	0.92	0.74	0.87	0.84
21	2	0.93	0.98	0.98	0.82	0.97
21	3	0.66	0.90	0.89	0.86	0.88
22	1	0.96	0.93	0.97	0.76	0.87
22	2	0.90	1.00	0.99	0.95	0.94
22	3	0.85	0.96	0.95	0.89	0.94
23	1	0.00	0.00	0.00	0.00	0.00
23	2	0.00	0.00	0.00	0.00	0.00
23	3	0.88	0.95	0.87	0.95	0.98
24	1	0.84	0.74	0.71	0.85	0.80
24	2	0.47	0.99	0.94	0.96	0.96
24	3	0.63	0.94	0.95	0.89	0.81
25	1	0.91	0.97	0.96	0.80	0.53
25	2	1.00	0.98	0.98	0.87	0.98
25	3	0.70	0.98	0.96	0.94	0.94
26	1	0.83	1.00	1.00	0.98	0.89
26	2	0.95	0.99	0.99	0.73	0.90
26	3	0.78	0.96	0.90	0.94	0.95
27	1	0.95	0.87	0.92	0.92	0.78
27	2	0.87	0.98	0.99	0.85	0.87
27	3	0.67	0.95	0.94	0.88	0.84
28	1	0.95	0.60	0.72	0.75	0.68
28	2	0.91	0.32	0.41	0.59	0.70
28	3	0.64	0.87	0.90	0.79	0.88
29	1	0.90	0.95	0.91	0.92	0.55

voir page suivante

TABLE 8 – taux de recouvrement et corrélations

NOGA2ST	TAILLE	Tx recouvrement	T017	T022	T053	WS
29	2	0.77	0.94	0.95	0.83	0.92
29	3	0.62	0.87	0.89	0.90	0.69
30	1	0.80	0.75	0.75	0.93	0.95
30	2	0.80	0.75	0.83	0.30	0.81
30	3	1.00	1.00	0.99	0.99	0.94
31	1	0.90	0.96	0.97	0.83	0.86
31	2	0.91	0.93	0.92	0.97	0.93
31	3	0.63	0.94	0.95	0.87	0.89
32	1	0.66	0.98	0.98	0.89	0.95
32	2	0.89	0.79	0.88	0.63	0.53
32	3	0.84	0.91	0.96	0.95	0.56
33	1	0.83	0.99	0.99	0.46	0.78
33	2	0.86	1.00	1.00	0.97	0.95
33	3	0.73	0.98	0.97	0.92	0.85
34	1	0.79	0.84	0.84	0.99	0.85
34	2	0.90	0.00	0.00	0.98	0.95
34	3	0.78	0.91	0.92	0.99	0.99
35	1	0.85	0.48	0.54	0.86	0.77
35	2	0.00	0.00	0.00	0.00	0.00
35	3	0.70	0.89	0.90	0.86	0.53
36	1	0.80	0.95	0.95	0.88	0.96
36	2	0.88	0.98	0.97	0.96	0.89
36	3	0.87	0.97	0.97	0.94	0.97
37	1	0.00	0.00	0.00	0.00	0.00
37	2	0.60	0.64	0.85	0.75	0.95
37	3	0.87	0.92	0.91	0.92	0.94
40	1	1.00	0.99	1.00	0.75	0.97
40	2	0.94	0.93	0.89	0.91	0.98
40	3	0.67	0.97	0.97	0.90	0.90
41	1	0.79	0.97	0.98	0.54	0.82
41	3	1.00	1.00	1.00	0.44	0.71
45	1	0.83	0.94	0.96	0.89	0.87
45	2	0.97	0.51	0.06	0.88	0.89
45	3	0.73	0.91	0.92	0.93	0.85
50	1	0.90	0.64	0.74	0.78	0.63
50	2	0.77	0.99	1.00	1.00	0.93
50	3	0.71	0.95	0.97	0.86	0.87
51	1	0.88	0.89	0.93	0.90	0.82
51	2	0.87	0.95	0.96	0.95	0.94
51	3	0.73	0.93	0.86	0.92	0.92
52	1	0.87	0.93	0.89	0.95	0.91
52	2	0.93	0.96	0.99	0.92	0.88
52	3	0.70	0.94	0.96	0.77	0.86
55	1	0.75	0.90	0.86	0.90	0.91
55	2	0.80	0.89	0.99	0.37	0.55
55	3	0.85	0.97	0.96	0.99	0.95
60	1	0.60	0.94	0.60	1.00	0.98
60	2	1.00	0.98	0.99	0.87	0.84
60	3	0.66	0.98	0.98	0.91	0.89
61	1	0.00	0.00	0.00	0.00	0.00
61	2	0.85	0.88	0.97	0.83	0.63
61	3	0.72	0.99	0.99	0.97	0.97
62	1	0.00	0.00	0.00	0.00	0.00
62	2	0.94	0.95	0.97	0.94	0.76
62	3	0.39	0.98	0.99	0.98	0.91
63	1	0.87	0.96	0.99	0.87	0.80
63	2	1.00	0.98	0.98	0.99	0.97
63	3	0.76	0.99	0.99	0.65	0.95
64	1	0.32	0.00	0.00	0.00	0.00

voir page suivante

TABLE 8 – taux de recouvrement et corrélations

NOGA2ST	TAILLE	Tx recouvrement	T017	T022	T053	WS
64	3	0.46	0.99	0.99	0.98	0.96
70	1	0.92	0.99	0.99	0.94	0.98
70	2	1.00	0.99	0.98	0.89	0.99
70	3	0.82	0.99	0.99	0.86	0.99
71	1	0.00	0.00	0.00	0.00	0.00
71	3	0.94	1.00	0.98	0.94	0.97
72	1	0.80	0.52	0.49	0.95	0.84
72	2	0.89	0.87	0.91	0.98	0.86
72	3	0.61	0.98	0.97	0.97	0.97
73	1	0.93	0.99	0.80	0.98	0.99
73	2	0.87	0.99	1.00	0.96	0.99
73	3	0.92	0.88	0.80	0.98	0.91
74	1	0.86	0.62	0.68	0.65	0.76
74	2	0.93	1.00	1.00	0.90	0.98
74	3	0.70	0.95	0.96	0.90	0.92
80	1	0.66	1.00	0.80	1.00	0.99
80	2	0.61	0.94	0.93	0.88	0.95
80	3	0.69	0.78	0.87	0.70	0.83
85	1	0.84	0.83	0.89	0.83	0.79
85	2	1.00	0.78	0.91	0.67	0.73
85	3	0.88	0.91	0.84	0.88	0.91
90	1	0.87	0.42	0.65	0.56	0.78
90	2	0.69	0.97	0.99	0.79	0.98
90	3	0.60	0.85	0.58	0.94	0.98
91	1	0.24	0.00	0.00	0.00	0.00
91	2	0.66	0.95	0.97	0.32	0.91
91	3	0.89	1.00	1.00	0.96	0.97
92	1	0.57	0.00	0.00	0.99	0.00
92	2	0.57	1.00	1.00	1.00	0.98
92	3	0.72	0.98	0.93	0.94	0.98
93	1	0.92	0.63	0.83	0.89	0.78
93	3	0.89	0.98	0.99	0.98	0.96
Valeur moyenne		0.74	0.81	0.82	0.80	0.80

A.3 Evolutions entre 1999, 2000 et 2001 - échantillons complets

TABLE 9 Evolutions de T017 - échantillons complets (en millions de francs suisses)

NOGA	t017v99	t017v00	t017v01	Evol. 99/00 (%)	Std.	signif.	Evol. 00/01 (%)	Std.	signif.
14	1832	1568	1569	-14,4	9,1	0	0,1	3,7	0
15	23190	23209	23763	0,1	2,9	0	2,4	2,0	0
16	2418	3176	2826	31,4	1,9	1	-11,0	0,3	1
17	3144	3497	3045	11,2	5,2	1	-12,9	2,3	1
18	1386	1598	1564	15,3	6,5	1	-2,1	4,9	0
19	395	547	387	38,7	9,7	1	-29,3	9,9	1
20	6320	6547	6728	3,6	6,6	0	2,8	5,3	0
21	4650	5203	5003	11,9	2,2	1	-3,8	2,3	0
22	10812	11354	11455	5,0	4,8	0	0,9	4,1	0
24	34747	39002	43453	12,2	3,1	1	11,4	3,2	1
25	6587	7055	7031	7,1	2,7	1	-0,3	3,2	0
26	4270	4805	4797	12,5	4,6	1	-0,2	3,3	0
27	5137	5696	5155	10,9	3,1	1	-9,5	2,0	1
28	16250	17425	18073	7,2	2,9	1	3,7	3,2	0
29	29585	32355	30016	9,4	1,9	1	-7,2	2,0	1
30	809	1188	1538	46,9	16,7	1	29,5	4,6	1
31	13551	14315	15718	5,6	2,2	1	9,8	1,2	1
32	5763	5982	6622	3,8	7,2	0	10,7	2,8	1
33	20417	23211	26566	13,7	5,3	1	14,5	2,2	1
34	1292	1348	1466	4,4	2,7	0	8,7	2,7	1
35	3207	3406	4171	6,2	3,6	0	22,4	2,4	1
36	5358	5678	5633	6,0	4,2	0	-0,8	***	
37	979	1030	1303	5,2	7,2	0	26,6	11,4	1
40	25538	22424	25702	-12,2	2,0	1	14,6	2,3	1
41	373	335	610	-10,0	1,9	1	81,8	5,3	1
45	42760	45891	45101	7,3	2,2	1	-1,7	1,9	0
50	9037	9006	10034	-0,4	3,5	0	11,4	5,4	1
51	36963	39977	36801	8,2	2,7	1	-7,9	1,4	1
52	31316	31220	33149	-0,3	1,7	0	6,2	1,1	1
55	13682	13596	14419	-0,6	5,3	0	6,1	3,3	0
60	6239	6103	7270	-2,2	4,2	0	19,1	9,4	1
61	284	352	325	24,3	7,8	1	-7,8	5,8	0
62	6865	8057	5028	17,4	1,4	1	-37,6	1,7	1
63	10066	11491	11022	14,2	4,8	1	-4,1	3,2	0
64	3199	13717	12599	328,7	47,1	1	-8,1	3,8	1
70	6690	6673	8000	-0,3	7,5	0	19,9	3,5	1
71	879	964	1288	9,7	***	***	33,6	***	
72	11441	11194	13326	-2,2	5,2	0	19,0	4,6	1
73	1900	2403	2823	26,5	7,3	1	17,5	3,9	1
74	41276	40499	50185	-1,9	4,7	0	23,9	8,2	1
80	2050	2086	2033	1,8	9,8	0	-2,5	11,1	0
85	3227	3394	3951	5,2	2,3	1	16,4	2,1	1
90	4045	3884	4167	-4,0	3,5	0	7,3	4,1	0
91	1323	1603	1806	21,1	8,5	1	12,7	8,2	0
92	7971	7415	8806	-7,0	15,6	0	18,8	17,0	0
93	2420	2475	2653	2,3	2,4	0	7,2	4,5	0
Total	471640	503954	528981	6,9	0,8	1	5,0	0,8	1

Les évolutions sont significatives pour 25 des 46 strates entre 1999 et 2000, et pour 26 entre 2000 et 2001. Pour les strates 36 et 71, la précision n'a pas pu être estimée.

TABLE 10 Evolutions de T022 - échantillons complets (en millions de francs suisses)

NOGA	t022v99	t022v00	t022v01	Evol. 99/00 (%)	Std.	signif.	Evol. 00/01 (%)	Std.	signif.
14	1031	819	802	-20,6	15,0	0	-2,1	5,6	0
15	17409	17238	17762	-1,0	3,5	0	3,0	2,6	0
16	1692	2128	1720	25,8	1,5	1	-19,2	0,5	1
17	1940	2129	1814	9,7	6,1	0	-14,8	2,6	1
18	942	1176	1127	24,8	8,2	1	-4,2	5,7	0
19	272	474	291	74,3	15,0	1	-38,6	15,6	1
20	3438	3646	3783	6,1	9,4	0	3,8	10,0	0
21	2881	3396	3231	17,9	2,9	1	-4,9	2,6	0
22	5510	5600	5967	1,6	5,5	0	6,5	6,0	0
24	23093	26986	30910	16,9	3,2	1	14,5	2,8	1
25	4041	4329	4213	7,1	3,5	1	-2,7	4,9	0
26	2306	2701	2661	17,1	6,7	1	-1,5	4,2	0
27	3535	4024	3354	13,8	4,4	1	-16,7	3,2	1
28	8249	9013	9562	9,3	4,2	1	6,1	5,0	0
29	17984	20079	18556	11,6	2,4	1	-7,6	2,5	1
30	498	809	1058	62,6	24,7	1	30,7	6,1	1
31	8660	9980	11962	15,2	2,8	1	19,9	1,4	1
32	3620	3771	4041	4,2	9,8	0	7,2	3,1	1
33	11987	13957	16040	16,4	8,2	1	14,9	2,8	1
34	840	864	976	2,9	3,3	0	12,9	3,9	1
35	2118	2135	2955	0,8	4,6	0	38,4	3,7	1
36	3107	3493	3424	12,4	5,4	1	-2,0	***	***
37	612	668	887	9,2	9,5	0	32,8	14,8	1
40	15023	13973	16182	-7,0	2,8	1	15,8	3,2	1
41	110	109	163	-0,7	9,0	0	49,0	24,6	1
45	21332	22559	22122	5,8	4,1	0	-1,9	2,9	0
50	3103	2932	3180	-5,5	7,8	0	8,5	4,5	0
51	12455	13653	12780	9,6	3,1	1	-6,4	1,8	1
52	9816	10078	11233	2,7	2,0	0	11,5	2,0	1
55	6815	6851	7300	0,5	7,6	0	6,6	3,8	0
60	2698	2836	3243	5,1	11,1	0	14,4	15,5	0
61	135	158	138	16,6	11,3	0	-12,8	7,5	0
62	4843	6014	3586	24,2	2,0	1	-40,4	2,4	1
63	6247	7181	6739	15,0	7,1	1	-6,2	4,5	0
64	2126	6645	6984	212,6	45,4	1	5,1	8,1	0
70	2575	2262	2628	-12,2	15,2	0	16,2	5,1	1
71	362	566	658	56,5	***	***	16,1	***	***
72	5292	4404	5106	-16,8	7,8	1	15,9	10,9	0
73	968	1331	1582	37,5	9,9	1	18,9	6,0	1
74	16239	14794	18817	-8,9	7,1	0	27,2	10,3	1
80	974	914	769	-6,2	11,4	0	-15,9	29,3	0
85	891	913	1086	2,4	2,9	0	19,0	4,2	1
90	2263	2049	2139	-9,5	5,7	0	4,4	7,8	0
91	746	1024	1046	37,2	13,6	1	2,1	8,2	0
92	5330	4800	6022	-9,9	20,2	0	25,5	22,2	0
93	822	909	950	10,6	4,0	1	4,4	6,9	0
Total	246929	266370	281546	7,9	1,2	1	5,7	0,9	1

Les évolutions sont significatives pour 23 des 46 strates entre 1999 et 2000, et pour 22 entre 2000 et 2001.

TABLE 11 Evolutions de T053 - échantillons complets (en millions de francs suisses)

NOGA	t053v99	t053v00	t053v01	Evol. 99/00 (%)	Std.	signif.	Evol. 00/01 (%)	Std.	signif.
14	414	418	411	0,9	2,5	0	-1,5	3,3	0
15	4129	4230	4298	2,4	1,3	0	1,6	1,3	0
16	297	288	285	-3,3	2,0	0	-1,0	0,2	1
17	978	1036	961	6,0	1,9	1	-7,2	2,1	1
18	356	360	354	1,2	4,5	0	-1,8	3,0	0
19	149	144	118	-3,3	2,5	0	-18,2	3,5	1
20	2242	2314	2478	3,2	7,2	0	7,1	***	***
21	1181	1215	1206	2,9	0,9	1	-0,7	0,8	0
22	4095	4273	4323	4,3	4,7	0	1,2	3,1	0
24	6391	6983	7251	9,3	0,9	1	3,8	1,2	1
25	1808	1940	2044	7,3	1,2	1	5,3	1,2	1
26	1407	1479	1541	5,1	2,0	1	4,2	2,4	0
27	1312	1305	1448	-0,6	3,2	0	11,0	2,3	1
28	5964	6127	6638	2,7	1,9	0	8,3	1,9	1
29	8942	9254	9302	3,5	1,3	1	0,5	1,3	0
30	237	318	360	33,9	4,5	1	13,4	2,1	1
31	3067	3252	3348	6,0	2,1	1	2,9	1,2	1
32	1618	1740	2072	7,5	2,7	1	19,1	1,9	1
33	5534	5895	6640	6,5	1,3	1	12,6	1,0	1
34	330	337	390	2,1	1,1	0	15,8	2,6	1
35	1040	980	1207	-5,8	0,8	1	23,1	1,0	1
36	1627	1693	1729	4,0	3,7	0	2,2	2,3	0
37	226	211	278	-6,8	4,1	0	31,9	3,9	1
40	2409	2185	2420	-9,3	0,7	1	10,8	0,7	1
41	75	68	132	-9,5	1,1	1	93,6	2,3	1
45	18452	19509	19274	5,7	1,2	1	-1,2	1,8	0
50	4484	4726	5075	5,4	3,1	0	7,4	2,6	1
51	16325	16959	15995	3,9	0,9	1	-5,7	1,0	1
52	15984	16398	17040	2,6	1,5	0	3,9	1,0	1
55	5586	5487	6035	-1,8	7,8	0	10,0	4,3	1
60	2639	2554	2807	-3,2	11,6	0	9,9	7,2	0
61	193	160	154	-17,1	2,4	1	-3,7	2,2	0
62	1135	1054	1000	-7,1	1,7	1	-5,2	1,1	1
63	2926	3291	3366	12,5	1,4	1	2,3	0,9	1
64	827	3115	3522	276,8	20,4	1	13,1	***	***
70	1424	1340	1657	-5,9	2,8	1	23,7	3,9	1
71	203	187	285	-8,2	***	***	52,4	***	***
72	4979	5800	6919	16,5	3,8	1	19,3	3,7	1
73	739	839	1042	13,6	4,5	1	24,2	4,8	1
74	19187	19553	24702	1,9	3,9	0	26,3	7,1	1
80	1536	1568	1580	2,1	8,2	0	0,7	18,8	0
85	2204	2352	2711	6,7	2,7	1	15,3	1,4	1
90	954	967	1197	1,4	1,7	0	23,8	2,6	1
91	609	639	677	5,0	3,7	0	6,0	2,7	1
92	2122	2158	2147	1,7	4,2	0	-0,5	7,8	0
93	1048	1020	1167	-2,6	2,8	0	14,3	5,1	1
Total	159385	167721	179585	5,2	0,7	1	7,1	0,9	1

Les évolutions sont significatives pour 23 des 46 strates entre 1999 et 2000, et pour 30 entre 2000 et 2001.

TABLE 12 Evolutions de WS - échantillons complets (en millions de francs suisses)

NOGA	wsv99	wsv00	wsv01	Evol. 99/00 (%)	Std.	signif.	Evol. 00/01 (%)	Std.	signif.
14	802	750	768	-6,5	5,7	0	2,4	5,9	0
15	5790	5971	6001	3,1	1,9	0	0,5	1,5	0
16	726	1048	1106	44,4	3,4	1	5,5	0,2	1
17	1205	1368	1231	13,5	5,5	1	-10,1	2,9	1
18	444	422	438	-4,9	5,4	0	3,8	4,4	0
19	123	73	118	-40,4	6,7	1	61,4	29,7	1
20	2882	2901	2945	0,6	6,2	0	1,5	7,3	0
21	1768	1807	1773	2,2	2,3	0	-1,9	2,8	0
22	5303	5754	5489	8,5	5,6	0	-4,6	2,8	0
24	11654	12016	12543	3,1	4,0	0	4,4	4,3	0
25	2546	2726	2819	7,1	2,2	1	3,4	1,9	0
26	1964	2106	2139	7,2	3,9	0	1,6	3,8	0
27	1616	1673	1801	3,5	5,4	0	7,7	2,6	1
28	8001	8412	8511	5,1	2,3	1	1,2	2,2	0
29	11628	12280	11460	5,6	1,7	1	-6,7	1,5	1
30	311	378	480	21,7	9,2	1	26,9	4,5	1
31	4890	4335	3756	-11,4	2,1	1	-13,3	1,7	1
32	2143	2210	2581	3,1	4,7	0	16,8	3,8	1
33	8430	9304	10527	10,4	2,1	1	13,1	2,4	1
34	452	484	490	7,1	3,0	1	1,3	2,5	0
35	1089	1271	1216	16,6	2,4	1	-4,3	1,4	1
36	2258	2184	2209	-3,3	3,6	0	1,1	***	***
37	367	362	416	-1,5	5,0	0	15,0	7,7	0
40	10515	8451	9520	-19,6	2,0	1	12,7	1,9	1
41	263	226	447	-13,9	2,9	1	97,6	13,1	1
45	21428	23332	22979	8,9	1,6	1	-1,5	1,7	0
50	5939	6074	6893	2,3	3,3	0	13,5	6,4	1
51	24513	26329	24073	7,4	3,2	1	-8,6	1,6	1
52	21502	21143	21916	-1,7	2,0	0	3,7	1,4	1
55	6866	6745	7119	-1,8	4,6	0	5,5	3,9	0
60	3542	3267	4027	-7,8	4,6	0	23,2	7,1	1
61	148	194	187	31,3	8,2	1	-3,7	6,5	0
62	2022	2043	1443	1,0	2,2	0	-29,4	1,5	1
63	3819	4310	4283	12,8	2,4	1	-0,6	2,0	0
64	1073	7072	5615	558,8	36,7	1	-20,6	***	***
70	4115	4411	5373	7,2	8,7	0	21,8	5,4	1
71	517	398	631	-23,0	10,0	1	58,5	***	***
72	6185	6790	8220	9,8	5,0	0	21,1	4,2	1
73	933	1073	1246	14,9	8,7	0	16,1	4,1	1
74	25061	25746	31390	2,7	4,2	0	21,9	7,9	1
80	1077	1178	1270	9,3	12,1	0	7,8	9,2	0
85	2341	2485	2876	6,1	3,3	0	15,8	2,4	1
90	1782	1836	2029	3,0	2,9	0	10,5	3,1	1
91	598	594	760	-0,7	5,1	0	28,1	16,9	0
92	2646	2618	2791	-1,1	6,2	0	6,6	10,5	0
93	1598	1565	1703	-2,0	2,4	0	8,8	4,6	0
Total	224877	237714	247606	5,7	0,8	1	4,2	0,9	1

Les évolutions sont significatives pour 19 des 46 strates entre 1999 et 2000, et pour 23 entre 2000 et 2001.

A.4 Evolutions entre 1999, 2000 et 2001 - échantillon commun

TABLE 13 Evolutions de T017 - échantillon commun (en millions de francs suisses)

NOGA	t017v99	t017v00a	t017v00b	t017v01	Evol. 99/00	Std.	signif.	Evol. 00/01	Std.	signif.
14	1645	1602	1573	1600	-2,6	3,4	0	1,7	2,8	0
15	23047	23808	23452	23552	3,3	1,3	1	0,4	1,0	0
16	2345	3176	3176	2826	35,4	1,0	1	-11,0	0,3	1
17	3202	3479	3530	3101	8,6	5,4	0	-12,2	2,9	1
18	1440	1407	1619	1667	-2,3	1,4	0	3,0	1,5	1
19	384	535	553	407	39,5	5,7	1	-26,5	13,3	1
20	6618	6265	6383	6555	-5,3	1,5	1	2,7	***	***
21	4676	5237	5121	5045	12,0	0,7	1	-1,5	1,5	0
22	10719	11175	11393	11148	4,3	3,9	0	-2,1	1,8	0
24	31749	32098	37715	40308	1,1	0,6	0	6,9	1,2	1
25	6739	7199	6888	6866	6,8	1,6	1	-0,3	1,6	0
26	4300	4819	4880	4890	12,1	3,4	1	0,2	4,8	0
27	4718	5790	4653	4988	22,7	0,7	1	7,2	1,2	1
28	16016	17381	17667	18073	8,5	2,7	1	2,3	3,0	0
29	29869	32560	32079	30469	9,0	1,9	1	-5,0	1,5	1
30	1028	1127	1191	1646	9,6	3,2	1	38,1	3,1	1
31	13659	14381	13944	15811	5,3	1,3	1	13,4	0,7	1
32	5178	5680	5542	6619	9,7	4,0	1	19,4	4,5	1
33	19385	21370	23637	26466	10,2	1,2	1	12,0	1,3	1
34	1298	1369	1352	1488	5,5	2,1	1	10,1	2,2	1
35	3424	3339	3693	3604	-2,5	1,4	0	-2,4	1,9	0
36	5388	5378	5411	5622	-0,2	2,3	0	3,9	***	***
37	913	1047	1116	1431	14,7	4,4	1	28,2	1,7	1
40	26554	22993	22764	25565	-13,4	1,7	1	12,3	0,9	1
41	375	337	330	610	-10,1	1,2	1	85,0	4,8	1
45	42773	45717	46227	45085	6,9	1,7	1	-2,5	1,8	0
50	9066	9176	9363	10037	1,2	4,0	0	7,2	3,4	1
51	37254	39581	40103	37271	6,2	1,9	1	-7,1	1,0	1
52	30650	31903	31073	32217	4,1	1,3	1	3,7	1,0	1
55	14706	13720	13964	14699	-6,7	3,6	0	5,3	2,4	1
60	6226	6208	6253	6968	-0,3	5,3	0	11,4	2,8	1
61	300	338	346	327	12,4	3,8	1	-5,7	2,4	1
62	6866	8131	4535	4073	18,4	1,2	1	-10,2	2,4	1
63	10530	11699	10999	10994	11,1	2,0	1	0,0	1,8	0
64	3522	10865	13529	12227	208,4	33,9	1	-9,6	0,8	1
70	6911	6756	6757	8166	-2,2	7,6	0	20,9	6,9	1
71	859	1253	865	1292	46,0	***	***	49,3	***	***
72	10766	10908	10896	13257	1,3	3,7	0	21,7	2,8	1
73	1902	2411	2379	2810	26,8	9,1	1	18,1	3,5	1
74	40012	43458	43001	48889	8,6	2,7	1	13,7	7,0	0
80	2057	1807	1868	2207	-12,1	5,8	1	18,1	8,8	1
85	3203	3428	3421	3943	7,0	2,4	1	15,2	2,0	1
90	4592	4328	3582	4414	-5,7	2,0	1	23,2	2,2	1
91	1470	1603	1747	1832	9,1	2,0	1	4,9	1,8	1
92	6817	7532	7960	7294	10,5	0,9	1	-8,4	4,6	0
93	2427	2466	2520	2684	1,6	2,1	0	6,5	4,6	0
Total	467575	496840	501050	521041	6,3	0,4	1	4,0	0,6	1

Les évolutions sont significatives pour 32 des 46 strates entre 1999 et 2000, et pour 30 entre 2000 et 2001. Les variables t017v99 et t017v00a sont calculées en utilisant l'échantillon commun entre la vague d'enquête de 1999 et celle de 2000. Les variables t017v00b et t017v01

sont calculées en utilisant l'échantillon commun entre la vague d'enquête de 2000 et celle de 2001.

TABLE 14 Evolutions de T022 - échantillon commun (en millions de francs suisses)

NOGA	t022v99	t022v00a	t022v00b	t022v01	Evol. 99/00	Std.	signif.	Evol. 00/01	Std.	signif.
14	827	824	814	826	-0,4	5,0	0	1,4	3,4	0
15	17166	17812	17449	17548	3,8	1,4	1	0,6	1,3	0
16	1659	2128	2128	1720	28,2	0,6	1	-19,2	0,5	1
17	1968	2178	2132	1851	10,7	6,1	0	-13,2	3,5	1
18	978	992	1168	1210	1,5	1,7	0	3,6	1,2	1
19	252	467	479	325	85,2	7,3	1	-32,2	21,3	0
20	3616	3406	3563	3664	-5,8	***	***	2,8	4,2	0
21	2905	3404	3311	3255	17,2	1,4	1	-1,7	1,6	0
22	5280	5460	5635	5731	3,4	3,3	0	1,7	2,2	0
24	20501	21774	26255	28572	6,2	0,5	1	8,8	1,2	1
25	4184	4483	4143	4052	7,2	1,5	1	-2,2	2,0	0
26	2338	2657	2761	2706	13,7	5,4	1	-2,0	7,3	0
27	3198	4109	2979	3243	28,5	0,9	1	8,9	1,5	1
28	7915	8875	9247	9538	12,1	3,9	1	3,1	4,6	0
29	18145	20187	19721	18830	11,3	2,3	1	-4,5	1,8	1
30	716	804	813	1113	12,2	4,8	1	37,0	5,3	1
31	8825	10094	9629	12028	14,4	2,0	1	24,9	1,0	1
32	3119	3536	3393	4052	13,4	3,9	1	19,4	5,0	1
33	11165	12460	14280	15895	11,6	1,6	1	11,3	1,6	1
34	845	876	860	1002	3,6	2,4	0	16,6	2,9	1
35	2161	2105	2388	2333	-2,6	2,2	0	-2,3	2,6	0
36	3087	3255	3328	3395	5,4	2,8	0	2,0	***	***
37	576	682	767	995	18,4	6,4	1	29,7	2,2	1
40	15623	14371	14366	15683	-8,0	2,2	1	9,2	1,0	1
41	116	112	100	163	-2,8	5,9	0	63,3	17,6	1
45	21279	22288	23004	22079	4,7	2,5	0	-4,0	2,5	0
50	3151	2982	3007	3153	-5,3	8,8	0	4,9	3,3	0
51	12365	13512	13856	12913	9,3	3,0	1	-6,8	1,2	1
52	9568	10272	9969	10363	7,4	1,7	1	4,0	1,7	1
55	7253	6732	7124	7335	-7,2	3,8	0	3,0	3,0	0
60	2710	2823	2861	3001	4,2	9,4	0	4,9	3,8	0
61	128	147	158	140	14,6	4,4	1	-11,4	3,4	1
62	4845	6122	3032	2543	26,4	0,9	1	-16,1	3,9	1
63	6696	7323	6649	6602	9,4	2,8	1	-0,7	2,6	0
64	2415	7628	6573	6674	215,8	45,6	1	1,5	3,2	0
70	2682	2232	2304	2642	-16,8	15,8	0	14,7	8,4	0
71	349	805	483	665	130,7	***	***	37,6	***	***
72	4385	4259	4269	5021	-2,9	4,8	0	17,6	3,5	1
73	969	1337	1320	1562	37,9	12,9	1	18,3	5,4	1
74	15054	16506	15553	17580	9,6	3,7	1	13,0	8,4	0
80	1019	850	863	1052	-16,6	9,9	0	21,9	10,0	1
85	877	909	898	1072	3,6	3,1	0	19,3	4,6	1
90	2740	2568	1709	2267	-6,3	4,4	0	32,6	2,7	1
91	884	1024	1040	1068	15,9	3,8	1	2,6	1,0	1
92	4159	4873	5173	4799	17,2	0,9	1	-7,2	4,6	0
93	827	899	926	984	8,7	3,6	1	6,2	7,2	0
Total	241519	263140	262480	273242	9,0	0,5	1	4,1	0,6	1

Les évolutions sont significatives pour 27 des 46 strates entre 1999 et 2000, et pour 24 entre 2000 et 2001.

TABLE 15 Evolutions de T053 - échantillon commun (en millions de francs suisses)

NOGA	t053v99	t053v00a	t053v00b	t053v01	Evol. 99/00	Std.	signif.	Evol. 00/01	Std.	signif.
14	416	424	417	420	2,1	1,6	0	0,8	2,6	0
15	4144	4217	4238	4353	1,7	1,0	0	2,7	0,7	1
16	284	288	288	285	1,1	0,8	0	-1,0	0,2	1
17	999	1019	1052	972	1,9	1,8	0	-7,7	3,2	1
18	369	368	385	364	-0,3	2,5	0	-5,6	2,2	1
19	147	147	146	116	-0,1	2,5	0	-20,7	3,6	1
20	2232	2070	2313	2508	-7,3	2,7	1	8,4	***	***
21	1185	1217	1218	1208	2,7	0,5	1	-0,9	0,5	0
22	4118	4280	4251	4259	4,0	4,9	0	0,2	2,3	0
24	6200	6487	6964	7012	4,6	0,5	1	0,7	1,1	0
25	1812	1915	1947	2037	5,7	1,0	1	4,6	0,7	1
26	1393	1488	1486	1551	6,8	1,6	1	4,4	3,1	0
27	1261	1302	1294	1412	3,3	0,5	1	9,1	0,5	1
28	6023	6132	6183	6637	1,8	1,2	0	7,3	1,9	1
29	8944	9249	9256	9222	3,4	1,3	1	-0,4	1,1	0
30	237	278	320	372	17,7	1,5	1	16,3	2,2	1
31	3021	3258	3248	3315	7,8	1,5	1	2,1	0,5	1
32	1566	1680	1630	2077	7,3	2,5	1	27,4	2,8	1
33	5493	5797	5951	6678	5,5	1,1	1	12,2	0,7	1
34	331	340	338	382	2,7	1,2	1	12,8	0,7	1
35	1011	991	989	1082	-2,0	0,6	1	9,3	0,6	1
36	1656	1632	1651	1740	-1,5	2,2	0	5,4	1,0	1
37	216	214	206	283	-1,2	2,5	0	37,4	0,7	1
40	2392	2198	2186	2404	-8,1	0,4	1	9,9	0,5	1
41	75	68	68	132	-9,4	1,1	1	94,5	2,1	1
45	18609	19840	19651	19379	6,6	1,1	1	-1,4	1,7	0
50	4556	4846	4728	5094	6,4	2,8	1	7,7	2,8	1
51	16500	16940	16968	16089	2,7	0,7	1	-5,2	0,6	1
52	16080	16513	16426	16891	2,7	0,9	1	2,8	0,7	1
55	5912	5761	5773	6296	-2,6	4,0	0	9,1	3,6	1
60	2638	2764	2689	2881	4,8	8,4	0	7,1	1,1	1
61	162	158	158	153	-2,7	2,5	0	-3,0	0,9	1
62	1135	1044	1033	991	-8,0	0,7	1	-4,1	0,8	1
63	2959	3277	3285	3334	10,8	1,4	1	1,5	0,8	0
64	830	2751	3025	3424	231,3	17,1	1	13,2	***	***
70	1431	1334	1352	1659	-6,8	3,3	1	22,7	4,9	1
71	205	231	176	285	12,6	***	***	62,5	***	***
72	4964	5663	5724	7037	14,1	2,6	1	22,9	2,9	1
73	733	828	833	1041	12,9	4,8	1	25,1	5,0	1
74	18718	19816	20684	24403	5,9	2,5	1	18,0	6,1	1
80	1651	1630	1672	2017	-1,2	3,3	0	20,6	2,3	1
85	2215	2364	2355	2717	6,7	2,9	1	15,4	1,2	1
90	941	956	1005	1211	1,6	2,3	0	20,5	1,9	1
91	629	639	663	684	1,6	0,9	0	3,1	1,8	0
92	2158	2199	2110	1915	1,9	2,4	0	-9,2	4,8	0
93	1043	1012	1036	1122	-3,0	2,6	0	8,4	3,5	1
Total	159596	167623	169372	179444	5,0	0,5	1	5,9	0,8	1

Les évolutions sont significatives pour 27 des 46 strates entre 1999 et 2000, et pour 33 entre 2000 et 2001.

TABLE 16 Evolutions de WS - échantillon commun (en millions de francs suisses)

NOGA	wsv99	wsv00a	wsv00b	wsv01	Evol. 99/00	Std.	signif.	Evol. 00/01	Std.	signif.
14	818	778	759	774	-4,9	4,2	0	2,0	3,3	0
15	5881	5997	6003	6004	2,0	1,2	0	0,0	1,3	0
16	686	1048	1048	1106	52,8	2,3	1	5,5	0,2	1
17	1234	1300	1398	1250	5,4	4,8	0	-10,6	2,7	1
18	462	416	452	458	-10,1	3,7	1	1,3	3,5	0
19	131	68	75	112	-48,3	4,5	1	50,0	40,7	0
20	3002	2859	2820	2892	-4,8	6,6	0	2,5	4,8	0
21	1771	1833	1810	1791	3,5	1,8	0	-1,1	2,0	0
22	5439	5715	5757	5418	5,1	4,7	0	-5,9	2,1	1
24	11248	10324	11460	11737	-8,2	1,0	1	2,4	1,4	0
25	2555	2715	2745	2814	6,3	2,4	1	2,5	1,3	0
26	1962	2162	2121	2187	10,2	2,7	1	3,1	3,7	0
27	1520	1681	1674	1746	10,6	1,3	1	4,3	1,1	1
28	8100	8506	8420	8535	5,0	2,0	1	1,4	2,2	0
29	11754	12377	12362	11638	5,3	1,8	1	-5,9	1,2	1
30	311	323	378	532	3,7	3,0	0	40,7	4,5	1
31	4833	4287	4315	3783	-11,3	1,3	1	-12,3	1,0	1
32	2059	2144	2149	2568	4,1	5,1	0	19,5	4,5	1
33	8220	8964	9415	10571	9,1	1,4	1	12,3	2,4	1
34	453	493	492	486	8,8	2,6	1	-1,2	1,8	0
35	1263	1234	1305	1271	-2,3	1,3	0	-2,6	1,5	0
36	2301	2123	2083	2228	-7,7	2,9	1	6,9	***	***
37	337	365	350	436	8,4	4,1	1	24,7	1,4	1
40	10931	8623	8398	9882	-21,1	1,7	1	17,7	1,4	1
41	259	224	230	447	-13,4	2,4	1	94,4	10,4	1
45	21494	23429	23223	23006	9,0	1,5	1	-0,9	1,6	0
50	5920	6194	6356	6927	4,6	3,7	0	9,0	3,9	1
51	24894	26074	26251	24417	4,7	2,2	1	-7,0	0,7	1
52	21081	21632	21105	21854	2,6	1,5	0	3,5	1,2	1
55	7453	6987	6840	7364	-6,2	3,8	0	7,7	2,9	1
60	3516	3385	3392	3966	-3,7	3,0	0	17,0	3,0	1
61	172	190	189	187	10,8	4,2	1	-1,0	4,2	0
62	2022	2010	1503	1530	-0,6	2,4	0	1,7	1,9	0
63	3834	4377	4350	4392	14,2	1,8	1	1,0	1,7	0
64	1107	3237	6956	5553	192,3	14,5	1	-20,2	***	***
70	4229	4524	4453	5524	7,0	9,5	0	24,1	7,1	1
71	510	448	382	627	-12,0	***	***	64,1	***	***
72	6381	6649	6627	8236	4,2	4,3	0	24,3	3,8	1
73	934	1074	1059	1252	15,0	9,9	0	18,2	3,3	1
74	24989	27014	27477	31333	8,1	2,9	1	14,0	6,6	1
80	1041	965	1012	1160	-7,3	4,8	0	14,6	8,0	0
85	2333	2524	2527	2883	8,2	3,5	1	14,1	2,1	1
90	1852	1761	1872	2148	-4,9	2,6	0	14,7	1,9	1
91	609	594	706	765	-2,5	3,3	0	8,3	5,2	0
92	2661	2662	2790	2502	0,0	1,9	0	-10,3	5,0	1
93	1600	1567	1594	1700	-2,0	2,3	0	6,7	4,7	0
Total	226162	233856	238683	247988	3,4	0,6	1	3,9	0,8	1

Les évolutions sont significatives pour 23 des 46 strates entre 1999 et 2000, et pour 24 entre 2000 et 2001.

B Estimation de la précision des évolutions dans les plans simples

B.1 Introduction

Dans les enquêtes longitudinales stratifiées, la précision des évolutions dépend directement du taux de recouvrement des échantillons. Dans ce chapitre, nous commençons par rappeler des résultats connus concernant les plans simples disjoints. Ensuite la variance des évolutions est calculée pour des plans simples dont les échantillons se superposent. La plupart de ces résultats sont bien connus et présentés par exemple dans Caron et Ravalet (2000) lorsque les taux de sondage sont négligeables. Cependant, nous avons calculé précisément les variances des estimateurs en tenant compte des corrections de population finies. Ces corrections peuvent jouer un rôle important dans les enquêtes auprès des entreprises. En effet, les entreprises de grandes tailles sont parfois sélectionnées avec des probabilités d'inclusion très élevées. Les corrections de population finie jouent alors un rôle important. Les calculs deviennent cependant beaucoup plus compliqués en population finie pour la raison suivante : si la taille de la population est finie, deux échantillons disjoints ne sont pas indépendants. Si la population est infinie, deux échantillons indépendants sont disjoints. Plusieurs estimateurs sont examinés : la différence des estimateurs transversaux, la différence estimée uniquement sur la partie commune, l'estimateur optimal.

B.2 Estimation dans des échantillons disjoints

Soit une population $U = \{1, ..., k, ..., N\}$ de taille N dans laquelle sont prélevés deux échantillons disjoints s_A et s_B de tailles respectives n_A et n_B. On s'intéresse à deux variables x et

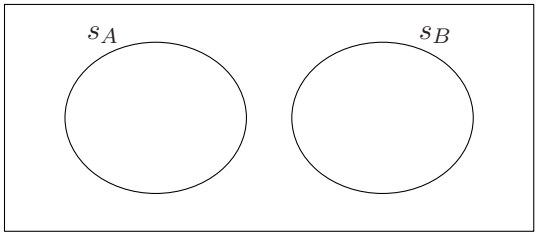


FIGURE 9 Echantillons disjoints

y dont les valeurs prises sur les unités de U sont notées respectivement x_k et $y_k, k \in U$. Les variables x et y peuvent représenter la même variable mesurée à deux moments différents.

L'objectif est d'estimer les moyennes

$$\overline{X} = \frac{1}{N} \sum_{k \in U} x_k \text{ et } \overline{Y} = \frac{1}{N} \sum_{k \in U} y_k.$$

On considère les estimateurs

$$\overline{x}_A = \frac{1}{n_A} \sum_{k \in s_A} x_k \text{ et } \overline{y}_B = \frac{1}{n_B} \sum_{k \in s_B} y_k,$$

où n_A et n_B sont les tailles respectives de s_A et s_B.

On dit alors que s_A et s_B suivent un plan simple bivarié (voir à ce sujet Goga, 2003) $p_{\text{simpl}}(s_A, s_B)$ si

$$p_{\text{simpl}}(s_A, s_B) = \frac{n_A! n_B! (N - n_A - n_B)!}{N!},$$

lorsque s_A contient n_A unités et s_B en contient n_B, et si $p_{\text{simpl}}(s_A, s_B) = 0$ sinon.

On peut trouver, par exemple dans Ardilly et Tillé (2003, page 24-28), une justification pour les résultats suivants :

$$\text{var}(\overline{x}_A) = \frac{N - n_A}{N n_A} S_x^2,$$

$$\text{var}(\overline{y}_B) = \frac{N - n_B}{N n_B} S_y^2,$$

$$\text{cov}(\overline{x}_A, \overline{y}_B) = -\frac{1}{N} S_{xy},$$

où

$$S_x^2 = \frac{1}{N-1} \sum_{k \in U} (x_k - \overline{X})^2,$$

$$S_y^2 = \frac{1}{N-1} \sum_{k \in U} (y_k - \overline{Y})^2,$$

$$S_{xy} = \frac{1}{N-1} \sum_{k \in U} (x_k - \overline{X})(y_k - \overline{Y}).$$

Il est surprenant de constater que $\text{cov}(\overline{x}_A, \overline{y}_B)$ ne dépend pas des tailles des échantillons n_A et n_B. La covariance est négative si x et y sont positivement corrélées, et tend vers 0 quand la taille de la population est grande.

B.3 Estimation par la différence des estimateurs transversaux dans des échantillons qui se chevauchent

B.3.1 Le problème et la notation

Supposons maintenant que l'intersection de deux échantillons s_1 et s_2 ne soit pas vide. On note

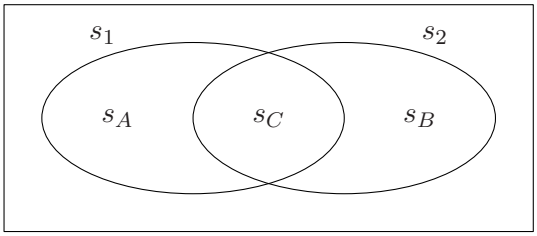


FIGURE 10 Echantillons qui se chevauchent

$$s_A = s_1 \backslash s_2, \quad s_B = s_2 \backslash s_1, \quad s_C = s_1 \cap s_2,$$

$$n_A = \text{card } s_A, \quad n_B = \text{card } s_B, \quad n_C = \text{card } s_C,$$

$$n_1 = n_A + n_C, \quad n_2 = n_B + n_C.$$

On suppose que s_1 et s_2 sont des échantillons prélevés selon un plan simple sans remise, les tailles n_1 et n_2 ne sont donc pas aléatoires. On suppose également que x ne peut être observé que pour s_1 et y pour s_2. Les tailles de s_A, s_B, et s_C, peuvent être aléatoires, mais on suppose que conditionnellement à n_A, n_B, et n_C, les échantillons s_A, s_B, et s_C, peuvent être considérés

comme simples, sans remise, de taille fixe. Plus formellement, si n_1 et n_2 sont donnés, on utilisera le plan $p_{simpl}(.,.,.)$ tel que :

$$p_{simpl}(s_A, s_B, s_C) = \frac{n_A! n_B! n_C! \left[N - n_A - n_B + n_C\right]!}{N!} \Pr\left[n_C = \text{card}\ (s_C)\right]$$

lorsque $n_A + n_C = n_1$ et $n_B + n_C = n_2$ et $p_{simpl}(s_A, s_B, s_C) = 0$ sinon (voir à ce sujet Goga, 2003).

Les cas particuliers suivants sont les plus importants :
– Si les deux échantillons forment un panel, $Pr(n_C = n_1 = n_2) = 1$.
– Si les échantillons sont disjoints, $Pr(n_C = 0) = 1$.
– Si un taux de rotation τ est fixé, $n_C = (1 - \tau)n_1$, et $n_1 = n_2$.
– Si les deux échantillons sont indépendants,

$$Pr(n_C = r) = \frac{\binom{n_2}{r}\binom{N-n_2}{n1-r}}{\binom{N}{n_1}}, r = 0, \ldots, \min(n_1, n_2).$$

On note

$$\begin{aligned} \overline{x}_A &= \frac{1}{n_A}\sum_{k \in s_A} x_k, \qquad & \overline{x}_C &= \frac{1}{n_C}\sum_{k \in s_C} x_k, \\ \overline{y}_B &= \frac{1}{n_B}\sum_{k \in s_B} y_k, \qquad & \overline{y}_C &= \frac{1}{n_C}\sum_{k \in s_C} y_k, \\ \overline{x}_1 &= \frac{n_A\overline{x}_A + n_C\overline{x}_C}{n_A + n_C}, \qquad & \overline{y}_2 &= \frac{n_B\overline{y}_B + n_C\overline{y}_C}{n_B + n_C}. \end{aligned}$$

B.3.2 L'estimateur et sa variance

Une première manière de procéder est d'estimer l'évolution $\Delta = \overline{X} - \overline{Y}$, par la différence des estimateurs transversaux

$$\widehat{\Delta}_1 = \overline{x}_1 - \overline{y}_2.$$

Cet estimateur est conditionnellement sans biais par rapport à n_A, n_B, n_C. En effet

$$\begin{aligned} &\mathrm{E}(\widehat{\Delta}_1 | n_A, n_B, n_C) \\ &= \frac{n_A \mathrm{E}(\overline{x}_A | n_A, n_B, n_C) + n_C \mathrm{E}(\overline{x}_C | n_A, n_B, n_C)}{n_A + n_C} - \frac{n_B \mathrm{E}(\overline{y}_B | n_A, n_B, n_C) + n_C \overline{\mathrm{E}}(y_C | n_A, n_B, n_C)}{n_B + n_C} \\ &= \frac{n_A\overline{X} + n_C\overline{X}}{n_A + n_C} - \frac{n_B\overline{Y} + n_C\overline{Y}}{n_B + n_C} \\ &= \overline{X} - \overline{Y}. \end{aligned}$$

L'estimateur $\widehat{\Delta}$ est donc également non-conditionnellement sans biais.

Résultat 1

$$\mathrm{var}(\widehat{\Delta}_1) = \frac{N - n_1}{N n_1} S_x^2 + \frac{N - n_2}{N n_2} S_y^2 - 2\left[\frac{\mathrm{E}(n_C)}{n_1 n_2} - \frac{1}{N}\right] S_{xy}$$

Démonstration
La réunion de deux échantillons disjoints tirés selon un plan simple est aussi un plan simple. Donc $s_1 = s_A \bigcup s_C$ et $s_2 = s_B \bigcup s_C$ sont sélectionnés selon un plan simple, ce qui donne

$$\mathrm{var}(\overline{x}_1) = \frac{N - n_1}{N n_1} S_x^2,$$

$$\text{var}(\overline{y}_2) = \frac{N-n_2}{Nn_2} S_y^2.$$

Comme

$$\text{var}(\widehat{\Delta}_1) = \text{var}(\overline{x}_1) + \text{var}(\overline{y}_2) - 2\text{cov}(\overline{x}_1, \overline{y}_2),$$

il reste à calculer

$$\text{cov}(\overline{x}_1, \overline{y}_2) = \text{E cov}(\overline{x}_1, \overline{y}_2 | n_A, n_B, n_C) + \text{cov}\left[E(\overline{x}_1 | n_A, n_B, n_C), E(\overline{y}_2 | n_A, n_B, n_C)\right].$$

Comme $\overline{x}_1$ et $\overline{y}_2$ sont sans biais conditionnellement à n_A, n_B et n_C,

$$\text{cov}\left[E(\overline{x}_1 | n_A, n_B, n_C), E(\overline{y}_2 | n_A, n_B, n_C)\right] = \text{cov}\left(\overline{X}, \overline{Y}\right) = 0.$$

On obtient donc

$$\text{cov}(\overline{x}_1, \overline{y}_2) = \text{E cov}(\overline{x}_1, \overline{y}_2 | n_A, n_B, n_C).$$

On commence par calculer la covariance conditionnelle :

$$\begin{aligned}
\text{cov}(\overline{x}_1, \overline{y}_2 | n_A, n_B, n_C) &= \text{cov}\left(\frac{n_A \overline{x}_A + n_C \overline{x}_C}{n_1}, \frac{n_B \overline{y}_B + n_C \overline{y}_C}{n_2} \middle| n_A, n_B, n_C \right) \\
&= \frac{n_A n_B}{n_1 n_2} \text{cov}\left(\overline{x}_A, \overline{y}_B | n_A, n_B, n_C\right) + \frac{n_A n_C}{n_1 n_2} \text{cov}\left(\overline{x}_A, \overline{y}_C | n_A, n_B, n_C\right) \\
&\quad + \frac{n_B n_C}{n_1 n_2} \text{cov}\left(\overline{x}_C, \overline{y}_B | n_A, n_B, n_C\right) + \frac{n_C^2}{n_1 n_2} \text{cov}\left(\overline{x}_C, \overline{y}_C | n_A, n_B, n_C\right) \\
&= \frac{-S_{xy}}{N n_1 n_2} \left(n_A n_B + n_A n_C + n_B n_C\right) + \frac{n_C^2}{n_1 n_2} \frac{N - n_C}{N n_C} S_{xy} \\
&= \frac{-S_{xy}}{N n_1 n_2} \left(n_A n_B + n_A n_C + n_B n_C + n_C^2\right) + \frac{n_C}{n_1 n_2} S_{xy} \\
&= \left(\frac{n_C}{n_1 n_2} - \frac{1}{N} \right) S_{xy},
\end{aligned}$$

et donc

$$\text{cov}(\overline{x}_1, \overline{y}_2) = \left(\frac{\text{E}(n_C)}{n_1 n_2} - \frac{1}{N} \right) S_{xy},$$

ce qui permet d'obtenir le résultat 1.

B.3.3 Cas particuliers

Le résultat 1 permet de traiter les cas particuliers de coordination suivants :

– Si les deux échantillons forment un panel, $n_C = n_1 = n_2$, alors

$$\text{var}(\widehat{\Delta}_1) = \frac{N - n_C}{N n_C} (S_x^2 + S_y^2 - 2S_{xy}).$$

– Si les échantillons sont disjoints, $n_C = 0$, alors

$$\text{var}(\widehat{\Delta}_1) = \frac{N - n_1}{N n_1} S_x^2 + \frac{N - n_2}{N n_2} S_y^2 + 2\frac{1}{N} S_{xy}.$$

– Si un taux de rotation τ est fixé et que $n_1 = n_2 = n$, alors

$$\text{var}(\widehat{\Delta}_1) = \frac{N - n}{N n} (S_x^2 + S_y^2) - 2\left[\frac{(1-\tau)}{n} - \frac{1}{N} \right] S_{xy}.$$

– Si les deux échantillons sont indépendants, $\text{E}(n_C) = n_1 n_2 / N$,

$$\text{var}(\widehat{\Delta}_1) = \frac{N - n_1}{N n_1} S_x^2 + \frac{N - n_2}{N n_2} S_y^2.$$

B.3.4 Estimation de la variance

Pour estimer la variance, il faut considérer deux cas :

– $\mathrm{E}(n_C)$ est connu, ce qui peut être le cas (par exemple quand l'on sait que les deux échantillons sont indépendants), alors

$$\widehat{\mathrm{var}}(\widehat{\Delta}_1) = \frac{N-n_1}{Nn_1}s_{x1}^2 + \frac{N-n_2}{Nn_2}s_{y2}^2 - 2\left(\frac{\mathrm{E}(n_C)}{n_1n_2} - \frac{1}{N}\right)s_{xyC}.$$

– $\mathrm{E}(n_C)$ n'est pas connu, et la seule information concernant la coordination est n_C,

$$\widehat{\mathrm{var}}(\widehat{\Delta}_1) = \frac{N-n_1}{Nn_1}s_{x1}^2 + \frac{N-n_2}{Nn_2}s_{y2}^2 - 2\left(\frac{n_C}{n_1n_2} - \frac{1}{N}\right)s_{xyC}, \tag{7}$$

où

$$s_{x1}^2 = \frac{1}{n_1-1}\sum_{s_1}(x_k - \bar{x}_1)^2,$$

$$s_{y2}^2 = \frac{1}{n_2-1}\sum_{s_2}(y_k - \bar{y}_2)^2,$$

et

$$s_{xyC} = \frac{1}{n_C-1}\sum_{s_C}(x_k - \bar{x}_C)(y_k - \bar{y}_C).$$

Ces estimateurs peuvent ne pas exister dans le cas où $n_C \leq 1$ et $\mathrm{E}(n_C) > 0$.

B.3.5 Gain par rapport à des échantillons indépendants

Si les deux échantillons étaient sélectionnés de manière indépendante, alors la variance serait

$$\mathrm{var}_{IND}\left(\widehat{\Delta}_1\right) = \frac{N-n_1}{Nn_1}S_x^2 + \frac{N-n_2}{Nn_2}S_y^2.$$

Le gain (ou la perte) due à une coordination positive ou négative est donc :

$$G = \frac{\mathrm{var}\left(\widehat{\Delta}_1\right)}{\mathrm{var}_{IND}\left(\widehat{\Delta}_1\right)} = 1 - 2\frac{(N\mathrm{E}(n_C) - n_1n_2)S_{xy}}{(Nn_2 - n_1n_2)S_x^2 + (Nn_1 - n_1n_2)S_y^2}.$$

Si S_x^2 est proche de S_y^2, on a alors

$$G \simeq 1 - 2\rho\frac{N\mathrm{E}(n_C) - n_1n_2}{N(n_1+n_2) - 2n_1n_2},$$

où ρ est le coefficient de corrélation entre x et y. Si N est grand, l'approximation devient

$$G = 1 - \rho q, \tag{8}$$

où q est le taux de recouvrement

$$q = \frac{2\mathrm{E}(n_C)}{n_1+n_2}.$$

L'expression (8) fournit un simple coefficient multiplicatif permettant de prendre en compte l'effet de la corrélation et du recouvrement lorsque la population est grande.

B.4 Estimation de la différence au moyen de la partie commune

B.4.1 L'estimateur et sa variance

On peut estimer la différence en utilisant uniquement la partie commune aux deux échantillons, ce qui donne l'estimateur

$$\widehat{\Delta}_c = \overline{y}_C - \overline{x}_C.$$

Cet estimateur est sans biais non-conditionnellement et conditionnellement à n_C. Sa variance conditionnelle vaut

$$\mathrm{var}(\widehat{\Delta}_c|n_C) = \frac{N-n_C}{Nn_C}(S_y^2 + S_x^2 - 2S_{xy}).$$

La variance non-conditionnelle vaut

$$\mathrm{var}(\widehat{\Delta}_c) = E\left(\frac{N-n_C}{Nn_C}\right)(S_y^2 + S_x^2 - 2S_{xy}) = \left[\mathrm{E}\left(\frac{1}{n_C}\right) - \frac{1}{N}\right](S_y^2 + S_x^2 - 2S_{xy}).$$

Cette variance non-conditionnelle peut être délicate à calculer quand n_C est aléatoire.

B.4.2 Comparaison avec la différence des estimateurs transversaux

Si on veut comparer les deux estimateurs de la différence

$$\begin{aligned}
&\mathrm{var}(\widehat{\Delta}_1) - \mathrm{var}(\widehat{\Delta}_c) \\
&= \frac{N-n_1}{Nn_1}S_x^2 + \frac{N-n_2}{Nn_2}S_y^2 - 2\left(\frac{\mathrm{E}(n_C)}{n_1n_2} - \frac{1}{N}\right)S_{xy} - \left[\mathrm{E}\left(\frac{1}{n_C}\right) - \frac{1}{N}\right](S_y^2 + S_x^2 - 2S_{xy}) \\
&= \left[\frac{1}{n_1} - \mathrm{E}\left(\frac{1}{n_C}\right)\right]S_y^2 + \left[\frac{1}{n_2} - \mathrm{E}\left(\frac{1}{n_C}\right)\right]S_x^2 - 2\left[\frac{\mathrm{E}(n_C)}{n_1n_2} - \mathrm{E}\left(\frac{1}{n_C}\right)\right]S_{xy}.
\end{aligned}$$

Si $n_1 = n_2 = n$, $S_x^2 = S_y^2 = S^2$, et $\mathrm{E}(1/n_C) \approx 1/\mathrm{E}(n_C)$, alors

$$\mathrm{var}(\widehat{\Delta}_1) - \mathrm{var}(\widehat{\Delta}_c) = \frac{1}{qn}\left[q-1\right]2S^2 - 2\frac{1}{qn}\left[q^2-1\right]\rho S^2 = \frac{2S^2}{qn}(1-q)\left[\rho(1+q)-1\right],$$

où $q = 2E(n_C)/(n_1+n_2)$ est le taux de recouvrement. L'estimateur $\widehat{\Delta}_c$ est plus précis que $\widehat{\Delta}_1$ si

$$\rho \geq \frac{1}{1+q}.$$

Par exemple, si $q = 0.7$, il est préférable de n'utiliser la partie commune dès que $\rho \geq 1/(1+0.7) = 0.588235294$ (voir à ce sujet Caron et Ravalet, 2000, p.346) Dans les cas, où la rotation est faible et la corrélation élevée, l'estimateur basé sur la différence des estimateurs transversaux n'est pas très pertinent.

B.5 Estimation optimale

Une troisième voie serait d'utiliser un estimateur optimal de la différence. Considérons le vecteur

$$\mathbf{z} = (\bar{x}_A, \bar{x}_C, \bar{y}_B, \bar{y}_C)',$$

de variance

$$\mathrm{var}(\mathbf{z}) = \begin{pmatrix}
\frac{N-n_A}{Nn_A}S_x^2 & -\frac{S_x^2}{N} & -\frac{S_{xy}}{N} & -\frac{S_{xy}}{N} \\
-\frac{S_x^2}{N} & \frac{N-n_C}{Nn_C}S_x^2 & -\frac{S_{xy}}{N} & \frac{N-n_C}{Nn_C}S_{xy} \\
-\frac{S_{xy}}{N} & -\frac{S_{xy}}{N} & \frac{N-n_B}{Nn_B}S_y^2 & -\frac{S_y^2}{N} \\
-\frac{S_{xy}}{N} & \frac{N-n_C}{Nn_C}S_{xy} & -\frac{S_y^2}{N} & \frac{N-n_C}{Nn_C}S_y^2
\end{pmatrix}.$$

On s'intéresse à une statistique T qui est linéaire en $\mathbf{z}$,

$$T = \mathbf{z}'\boldsymbol{\beta},$$

où $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)$.

Soit la matrice de contrainte

$$\mathbf{C} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

En imposant des contraintes *ad hoc* sur $\boldsymbol{\beta}$, on peut construire des estimateurs sans biais de plusieurs fonctions d'intérêt :

– Si $\mathbf{C}\boldsymbol{\beta} = (1, 0)'$, on a $\beta_1 + \beta_2 = 1$ et $\beta_3 + \beta_4 = 0$, autrement dit T est sans biais de $\overline{X}$.
– Si $\mathbf{C}\boldsymbol{\beta} = (0, 1)'$, on a $\beta_1 + \beta_2 = 0$ et $\beta_3 + \beta_4 = 1$, autrement dit T est sans biais de $\overline{Y}$.
– Si $\mathbf{C}\boldsymbol{\beta} = (1, -1)'$, on a $\beta_1 + \beta_2 = 1$ et $\beta_3 + \beta_4 = -1$, autrement dit T est sans biais de $\Delta = \overline{X} - \overline{Y}$.

On va chercher le vecteur $\boldsymbol{\beta}$ qui minimise la variance de T sous la contrainte $\mathbf{C}\boldsymbol{\beta} = \mathbf{c}$, où $\mathbf{c}$ est un vecteur quelconque de $\mathbb{R}^2$. Le programme s'écrit

$$\min_{\boldsymbol{\beta}} \operatorname{var}(\mathbf{z}'\boldsymbol{\beta}) \text{ sous la contrainte } \mathbf{C}\boldsymbol{\beta} = \mathbf{c}.$$

En utilisant la méthode de Lagrange, on obtient

$$\boldsymbol{\beta} = \boldsymbol{\gamma}\mathbf{c}$$

où

$$\boldsymbol{\gamma} = \left[\operatorname{var}(\mathbf{z})\right]^{-1} \mathbf{C}' \left\{ \mathbf{C} \left[\operatorname{var}(\mathbf{z})\right]^{-1} \mathbf{C}' \right\}^{-1}.$$

Après de longs calculs que nous avons réalisés avec un logiciel de calcul symbolique, on obtient

$$\boldsymbol{\gamma} = g \begin{pmatrix} n_A \left[n_B(1-\rho^2) + n_C\right] & n_A n_C b_{y|x} \\ n_C(n_B + n_C) & -n_A n_C b_{y|x} \\ n_B n_C b_{x|y} & n_B \left[n_A(1-\rho^2) + n_C\right] \\ -n_B n_C b_{x|y} & n_C(n_A + n_C) \end{pmatrix}$$

où

$$g = \left[(n_A n_B(1-\rho^2) + (n_A + n_B + n_C)n_C\right]^{-1},$$
$$\rho = S_{xy}/(S_x S_y), b_{y|x} = S_{xy}/S_x^2, b_{x|y} = S_{xy}/S_y^2.$$

Il est dès lors possible de construire les estimateur optimaux $T_{opt} = \mathbf{z}'\boldsymbol{\beta} = \mathbf{z}'\boldsymbol{\gamma}\mathbf{c}$ de plusieurs fonctions d'intérêt en utilisant le vecteur $\mathbf{c}$ adéquat.

– Si $\mathbf{c} = (1, 0)'$, on obtient l'estimateur optimal de $\overline{X}$

$$\widehat{\overline{X}}_{OPT} = g \left\{ n_A \left[n_B(1-\rho^2) + n_C\right] \bar{x}_A + n_C(n_B + n_C)\bar{x}_C + n_B n_C b_{x|y}(\bar{y}_B - \bar{y}_C) \right\}.$$

– Si $\mathbf{c} = (0, 1)'$, on obtient l'estimateur optimal de $\overline{Y}$

$$\widehat{\overline{Y}}_{OPT} = g \left\{ n_A n_C b_{y|x}(\bar{x}_A - \bar{x}_C) + n_B \left[n_A(1-\rho^2) + n_C\right] \bar{y}_B + n_C(n_A + n_C)\bar{y}_C \right\}.$$

– Si $\mathbf{c} = (1, -1)'$, on obtient l'estimateur optimal de Δ. On vérifie directement que

$$\widehat{\Delta}_{OPT} = \widehat{\overline{X}}_{OPT} - \widehat{\overline{Y}}_{OPT}.$$

La variance de ces estimateurs optimaux vaut

$$\mathrm{var}(T_{opt}) = \mathbf{c}'\boldsymbol{\gamma}'\mathrm{var}(\mathbf{z})\boldsymbol{\gamma}\mathbf{c} = \mathbf{c}'\left\{\mathbf{C}\left[\mathrm{var}(\mathbf{z})\right]^{-1}\mathbf{C}'\right\}^{-1}\mathbf{c}.$$

où

$$\begin{aligned}
&\left\{\mathbf{C}\left[\mathrm{var}(\mathbf{z})\right]^{-1}\mathbf{C}'\right\}^{-1} \\
&= \frac{g}{N}\left[n_C(N-n_A-n_B-n_C)\begin{pmatrix} S_x^2 & S_{xy} \\ S_{xy} & S_y^2 \end{pmatrix} + (1-\rho^2)\begin{pmatrix} n_B(N-n_A)S_x^2 & -n_An_BS_{xy} \\ -n_An_BS_{xy} & n_A(N-n_B)S_y^2 \end{pmatrix}\right] \\
&= \frac{g}{N}\left[\left\{n_C(N-n_A-n_B-n_C) - n_An_B(1-\rho^2)\right\}\begin{pmatrix} S_x^2 & S_{xy} \\ S_{xy} & S_y^2 \end{pmatrix} + (1-\rho^2)N\begin{pmatrix} n_BS_x^2 & 0 \\ 0 & n_AS_y^2 \end{pmatrix}\right].
\end{aligned}$$

C Prise en compte de la non-réponse totale

C.1 Introduction

Dans cette annexe, on modélise la non-réponse au moyen d'un plan de Bernoulli multivarié. On considère que la non-réponse est indépendante du plan de sélection. Selon le modèle, chaque entreprise décide de répondre ou non aléatoirement et les probabilités de réponse sont égales entre les entreprises. Il s'agit du modèle le plus élémentaire. Cependant, si une entreprise ne répond pas à la première vague, il est fortement probable qu'elle ne répondra pas non plus à la deuxième vague. Le modèle prend en compte cette dépendance en considérant séparément quatre cas :

– l'entreprise répond à la première vague et à la seconde,
– l'entreprise répond à la première vague mais pas à la seconde,
– l'entreprise ne répond pas à la première vague mais bien à la seconde,
– l'entreprise ne répond ni à la première vague, ni à la seconde.

C.2 Le modèle de non-réponse

La non-réponse est couramment modélisée par un plan bernoullien. La non-réponse est donc supposée être *missing completely at random*, ce qui signifie que la probabilité de répondre est la même pour toutes les unités statistiques. Cette modélisation implique que les probabilités de non-réponse sont égales, et qu'une entreprise décide de répondre indépendamment de la réponse des autres entreprises. Le plan de non-réponse est le suivant :

$$q(r_A, r_B, r_C, r_D) = \phi_A^{\mathrm{card}\ r_A}\phi_B^{\mathrm{card}\ r_B}\phi_C^{\mathrm{card}\ r_C}\phi_D^{\mathrm{card}\ r_D},$$

où $r_A, r_B, r_C, r_D \subset U$, et r_A, r_B, r_C, r_D sont mutuellement exclusifs. Ce plan peut représenter un échantillonnage en deux vagues où

– ϕ_A est la probabilité de répondre à la vague 1, mais pas à la vague 2,
– ϕ_B est la probabilité de répondre à la vague 2, mais pas à la vague 1,
– ϕ_C est la probabilité de répondre à la vague 1, et à la vague 2,
– ϕ_D est la probabilité de ne pas répondre ni à la vague 1, ni à la vague 2.

Les quantités ϕ_A, ϕ_B, ϕ_C, ϕ_D ne sont évidemment pas connues. Dans ce plan, les tailles des échantillons sont aléatoires et ont toutes une loi de Bernoulli. Evidemment, les non-réponses peuvent être liées ou non d'une vague à l'autre.

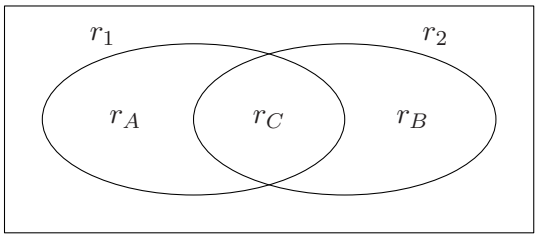


FIGURE 11 Plan de réponses

C.3 Interaction entre le plan de sélection et le plan de réponse

Supposons qu'un échantillon soit tiré selon un plan défini dans la Figure 10. Ce plan peut représenter un échantillonnage en deux vagues où
– s_A est l'échantillon des unités sélectionnées à la vague 1, mais pas à la vague 2,
– s_B est l'échantillon des unités sélectionnées à la vague 2, mais pas à la vague 1,
– s_C est l'échantillon des unités sélectionnées à la vague 1, et à la vague 2,
– s_D est l'échantillon des unités sélectionnées ni à la vague 1, ni à la vague 2.
Ce plan est supposé indépendant du plan de non-réponse. La Figure 12 présente toutes les intersections possibles. Dans le tableau 17, on indique quels sous-ensembles sont finalement observés.

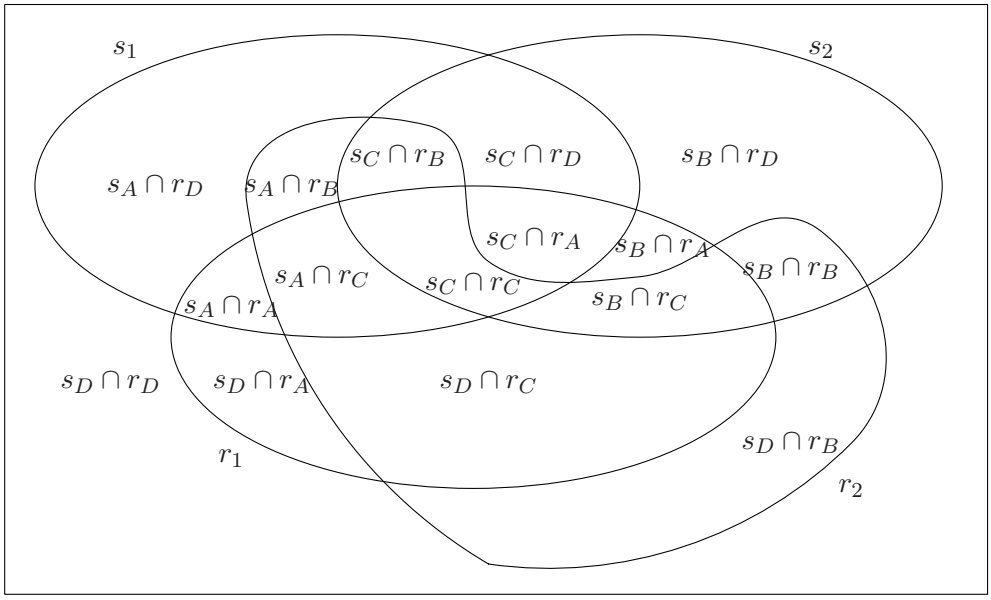


FIGURE 12 Interaction du plan de sélection et du plan de réponse

On peut définir les échantillons obtenus réellement :
– pour la vague 1 et pas pour la vague 2

$$t_A = (r_A \cap s_A) \cup (r_A \cap s_C) \cup (r_C \cap s_A),$$

– pour la vague 2 et pas pour la vague 1

$$t_B = (r_B \cap s_B) \cup (r_B \cap s_C) \cup (r_C \cap s_B),$$

TABLE 17 Interaction des deux plans

	uniquement vague 1	uniquement vague 2	vague 1 et vague 2	aucune vague
$s_A \cap r_A$	X			
$s_A \cap r_B$				X
$s_A \cap r_C$	X			
$s_A \cap r_D$				X
$s_B \cap r_A$				X
$s_B \cap r_B$		X		
$s_B \cap r_C$		X		
$s_B \cap r_D$				X
$s_C \cap r_A$	X			
$s_C \cap r_B$		X		
$s_C \cap r_C$			X	
$s_C \cap r_D$				X
$s_D \cap r_A$				X
$s_D \cap r_B$				X
$s_D \cap r_C$				X
$s_D \cap r_D$				X

– pour les deux vagues

$$t_C = r_C \cap s_C,$$

– pour aucune des deux vagues

$$t_D = r_D \cup s_D \cup (r_A \cap s_B) \cup (r_B \cap s_A).$$

De plus, on note, $\text{card}\,(t_A) = m_A$, $\text{card}\,(t_B) = m_B$, $\text{card}\,(t_C) = m_C$, et $\text{card}\,(t_D) = m_D$.

On a le résultat suivant :

Résultat 2 *Si le plan $p(s_A, s_B, s_C, s_D)$ est indépendant du plan $q(r_A, r_B, r_C, r_D)$ alors le plan*

$$p(t_A, t_B, t_C, t_D | m_A, m_B, m_C, m_D)$$

est un plan simple quadrivarié, autrement dit il s'écrit

$$p(t_A, t_B, t_C, t_D | m_A, m_B, m_C, m_D) = \frac{m_A! m_B! m_C! m_D!}{N!}.$$

On pourrait écrire une preuve formelle, mais les grandes lignes de la démonstration sont les suivantes. Un plan est dit simple si tous les échantillons de même taille ont la même probabilité d'être sélectionnés. Un plan simple peut donc avoir une taille aléatoire. Le plan de taille fixe et les plans bernoulliens sont simples. Cette définition s'étend également aux plans multivariés. Si deux plans simples sont indépendants, l'intersection des deux échantillons suit aussi un plan simple. Donc, les 16 échantillons disjoints du tableau 17 suivent tous des plans simples. Comme la réunion de deux échantillons disjoints sélectionnés selon un plan simple suit aussi un plan simple, les échantillons t_A, t_B, t_C, t_D suivent aussi un plan simple. En conditionnant par rapport aux tailles m_A, m_B, m_C et m_D, les échantillons suivent toujours un plan simple.
Autrement dit, en travaillant conditionnellement aux tailles d'échantillons, le plan est simple multivarié. Si l'inférence est menée conditionnellement aux tailles d'échantillons, l'estimation des

probabilités ϕ_A, ϕ_B, ϕ_C, ϕ_D n'est pas nécessaire et une inférence sans biais peut être menée, comme si on avait affaire à un plan simple. La théorie du chapitre précédent s'applique donc directement sur les répondants, et toute l'information sur le recouvrement des deux échantillons se trouve dans $\mathrm{card}\,(t_C)$, que ce recouvrement soit dû au plan ou au lien existant entre les non-réponses aux deux vagues.

D Estimateurs optimaux

D.1 Estimateur optimal parmi les combinaisons linéaires de $\widehat{\Delta}_d$ et $\widehat{\Delta}_c$

Comme on l'a annoncé, l'estimateur $\alpha\widehat{\Delta}_d + (1-\alpha)\widehat{\Delta}_c$, $\alpha \in [0,1]$ est de variance minimale lorsque

$$\alpha = \frac{\mathrm{var}(\widehat{\Delta}_c) - \mathrm{cov}(\widehat{\Delta}_c, \widehat{\Delta}_d)}{\mathrm{var}(\widehat{\Delta}_c) + \mathrm{var}(\widehat{\Delta}_d) - 2\mathrm{cov}(\widehat{\Delta}_c, \widehat{\Delta}_d)}, \tag{9}$$

avec pour variance

$$\mathrm{var}(\widehat{\Delta}_c) - \frac{(\mathrm{var}(\widehat{\Delta}_c) - \mathrm{cov}(\widehat{\Delta}_c, \widehat{\Delta}_d))^2}{\mathrm{var}(\widehat{\Delta}_c) + \mathrm{var}(\widehat{\Delta}_d) - 2\mathrm{cov}(\widehat{\Delta}_c, \widehat{\Delta}_d)}.$$

En effet, la variance de

$$\alpha\widehat{\Delta}_d + (1-\alpha)\widehat{\Delta}_c + y_g^4(t) - y_g^4(t-1) \qquad \alpha \in [0,1]$$

vaut

$$\alpha^2\mathrm{var}(\widehat{\Delta}_d) + (1-\alpha)^2\mathrm{var}(\widehat{\Delta}_c) + 2\alpha(1-\alpha)\mathrm{cov}(\widehat{\Delta}_c, \widehat{\Delta}_d),$$

et la dérivée en α de cette expression s'annule lorsque l'équation 9 est vérifiée.
On estime cet α et la variance en substituant :

$$\widehat{\mathrm{var}}(\widehat{\Delta}_c) = \sum_{h=1}^{H} N_{h\backslash s4} \frac{N_{h\backslash s4} - n_{h\backslash s4}(s_t \cap s_{t-1})}{n_{h\backslash s4}(s_t \cap s_{t-1})} s^2_{h\backslash s4}\left[ue(t) - ue(t-1)\right],$$

$$\widehat{\mathrm{var}}(\widehat{\Delta}_d) = \sum_{h=1}^{H} N^2_{h\backslash s4} \left\{ \frac{N_{h\backslash s4} - n_{h\backslash s4}(s_t \backslash s_{t-1})}{N_{h\backslash s4} n_{h\backslash s4}(s_t \backslash s_{t-1})} s^2_{h\backslash s4}[ue(t)] + \frac{N_{h\backslash s4} - n_{h\backslash s4}(s_{t-1} \backslash s_t)}{N_{h\backslash s4} n_{h\backslash s4}(s_{t-1} \backslash s_t)} s^2_{h\backslash s4}[ue(t-1)] \right.$$
$$\left. + \frac{2}{N_{h\backslash s4}} s_{h\backslash s4}[ue(t-1), ue(t)] \right\},$$

et

$$\widehat{\mathrm{cov}}(\widehat{\Delta}_c, \widehat{\Delta}_d) = -\sum_{h=1}^{H} N_{h\backslash s4} \left\{ s^2_{h\backslash s4}[ue(t)] + s^2_{h\backslash s4}[ue(t-1)] - 2s_{h\backslash s4}[ue(t-1), ue(t)] \right\}.$$

D.2 Estimateur optimal parmi les combinaisons linéaires des estimateurs sur les sous-échantillons

On cherche un estimateur sans biais de l'évolution parmi les combinaisons linéaires :

$$\alpha\widehat{Y}^{\backslash 4}(s_t \backslash s_{t-1}) + \beta\widehat{Y}^{\backslash 4}(s_t \cap s_{t-1}) + \gamma\widehat{Y}^{\backslash 4}(s_{t-1} \cap s_t) + \delta\widehat{Y}^{\backslash 4}(s_{t-1} \backslash s_t),$$

On impose la condition suffisante : $\alpha + \beta = 1$ et $\gamma + \delta = -1$.

Soit $\widehat{\delta}_t = \widehat{Y}^{\backslash 4}(s_t \backslash s_{t-1}) - \widehat{Y}^{\backslash 4}(s_t \cap s_{t-1})$ et $\widehat{\delta}_{t-1} = \widehat{Y}^{\backslash 4}(s_{t-1} \backslash s_t) - \widehat{Y}^{\backslash 4}(s_{t-1} \cap s_t)$

L'estimateur de variance minimale dans cette famille vaut :

$$\alpha \widehat{Y}^{\backslash 4}(s_t \backslash s_{t-1}) + (1-\alpha)\widehat{Y}^{\backslash 4}(s_t \cap s_{t-1}) - \beta \widehat{Y}^{\backslash 4}(s_{t-1} \backslash s_t) - (1-\beta)\widehat{Y}^{\backslash 4}(s_{t-1} \cap s_t)$$

pour

$$\begin{aligned}
\alpha &= \frac{\mathrm{cov}(\widehat{\Delta}_c, \widehat{\delta}_{t-1})\mathrm{cov}(\widehat{\delta}_t, \widehat{\delta}_{t-1}) - \mathrm{var}(\widehat{\delta}_{t-1})\mathrm{cov}(\widehat{\Delta}_c, \widehat{\delta}_t)}{\mathrm{var}(\widehat{\delta}_t)\mathrm{var}(\widehat{\delta}_{t-1}) - \mathrm{cov}^2(\widehat{\delta}_t, \widehat{\delta}_{t-1})} \\
\beta &= \frac{\mathrm{var}(\widehat{\delta}_t)\mathrm{cov}(\widehat{\Delta}_c, \widehat{\delta}_{t-1}) - \mathrm{cov}(\widehat{\Delta}_c, \widehat{\delta}_t)\mathrm{cov}(\widehat{\delta}_t, \widehat{\delta}_{t-1})}{\mathrm{var}(\widehat{\delta}_t)\mathrm{var}(\widehat{\delta}_{t-1}) - \mathrm{cov}^2(\widehat{\delta}_t, \widehat{\delta}_{t-1})}.
\end{aligned}$$

On peut estimer ces paramètres par substitution :

$$\begin{aligned}
\widehat{\mathrm{cov}}(\widehat{\Delta}_c, \widehat{\delta}_{t-1}) &= \sum_{h=1}^{H} \frac{N_{h\backslash s4}^2}{n_{h\backslash s4}(s_t \cap s_{t-1})} \left(s_{h\backslash s4}^2[ue(t-1)] - s_{h\backslash s4}[ue(t-1), ue(t)] \right), \\
\widehat{\mathrm{cov}}(\widehat{\Delta}_c, \widehat{\delta}_t) &= \sum_{h=1}^{H} \frac{N_{h\backslash s4}^2}{n_{h\backslash s4}(s_t \cap s_{t-1})} \left(s_{h\backslash s4}[ue(t-1), ue(t)] - s_{h\backslash s4}^2[ue(t)] \right), \\
\widehat{\mathrm{cov}}(\widehat{\delta}_t, \widehat{\delta}_{t-1}) &= \sum_{h=1}^{H} \frac{N_{h\backslash s4}^2}{n_{h\backslash s4}(s_t \cap s_{t-1})} s_{h\backslash s4}[ue(t-1), ue(t)], \\
\widehat{\mathrm{var}}(\widehat{\delta}_t) &= \sum_{h=1}^{H} \frac{N_{h\backslash s4}^2 n_{h\backslash s4}(s_t)}{n_{h\backslash s4}(s_t \backslash s_{t-1}) n_{h\backslash s4}(s_t \cap s_{t-1})} s_{h\backslash s4}^2[ue(t)], \\
\widehat{\mathrm{var}}(\widehat{\delta}_{t-1}) &= \sum_{h=1}^{H} \frac{N_{h\backslash s4}^2 n_{h\backslash s4}(s_{t-1})}{n_{h\backslash s4}(s_{t-1} \backslash s_t) n_{h\backslash s4}(s_t \cap s_{t-1})} s_{h\backslash s4}^2[ue(t-1)].
\end{aligned}$$

E Glossaire

vocabulaire de l'enquête WS

- NOGA : nomenclature d'activité
- NOGA2ST : strate d'activité dans la base de sondage, notées $h = 1, ..., H$ dans le rapport
- NOGA2 : poststrate d'activité déclarée dans l'enquête, notées $g = 1, ..., G$ dans le rapport
- ETP : équivalents temps plein
- T017 : valeur de production brute
- T022 : consommation intermédiaire
- T053 : salaires
- WS : valeur ajoutée $T017 - T022$

notations utilisées : estimations transversales

- N : effectif de la population U
- N_h : effectif de la strate U_h
- n_h^b : effectif de l'échantillon sélectionné dans la strate h
- s : échantillon observé
- s_h : échantillon observé dans la strate h, $s_h = s \cap U_h$
- n_h : effectif de s_h
- y_k : variable d'intérêt
- x_k : variable de calage, nombre d'ETP dans l'entreprise k
- t_k : taille de l'entreprise k (confondue avec x_k)
- $s4$: poststrate de très grosses entreprises
- $\ell(s)$: taille à partir de laquelle une entreprise est dans $s4$
- n_h^4 : effectif de $s_h \cap s4$
- $N_{h\backslash s4}$: effectif de $U_h \backslash s4$
- $n_{h\backslash s4}$: effectif de $U_h \backslash s4$
- $\widehat{Y}_{HT}$: estimateur de Horvitz-Thompson du total de Y
- $y^4(t)$, $x^4(t)$ (resp. $y_g^4(t)$, $x_g^4(t)$) : total de y, x dans la poststrate $s4$ (resp. dans la poststrate $g \cap s4$) à la date t
- $\widehat{Y}_{HT,g}^{\backslash 4}$: estimateur de Horvitz-Thompson du total de Y dans la poststrate $g \cap s4$
- $\widehat{Y}_g$: estimateur poststratifié du total de Y dans la poststrate g, $\widehat{Y}_g = \widehat{Y}_{HT,g}^{\backslash 4} + y_g^4$
- $\widehat{R}_g^{\backslash 4}$: ratio entre $\widehat{Y}_{HT,g}^{\backslash 4}$ et $\widehat{X}_{HT,g}^{\backslash 4}$
- $\widehat{Y}_g^Q$: estimateur poststratifié du total de Y dans la strate g, calé dans $g\backslash s4$ sur le total de X
- u_k : variable de robustification
- $\widehat{YU}_g^{\backslash 4}$, $\widehat{XU}_g^{\backslash 4}$: estimateurs robustifiés du total de Y, X
- $\widehat{Y}_g^{Qr}$: estimateur poststratifié, robustifié et calé du total de Y dans g
- ue_k : résidu de régression de $u_k * y_k$ sur $u_k * x_k$, $k \in g$ dans la poststrate g

notations utilisées : estimations longitudinales

- s_t, s_{t-1} : échantillons observés aux temps t et $t-1$
- $\widehat{\Delta}$: différence des estimateurs transversaux du total de Y aux temps t et $t-1$
- $\widehat{Q}$: rapport des estimateurs transversaux du total de Y aux temps t et $t-1$
- $\widehat{Y}^{\backslash 4}(s_t \cap s_{t-1})$: estimateur du total de Y au temps t dans les petites, moyennes et grandes entreprises, construit sur la partie commune de l'échantillon $s_t \cap s_{t-1}$

– $\widehat{Y}^{\backslash 4}(s_{t-1} \cap s_t)$: estimateur du total de Y au temps $t-1$ dans les petites, moyennes et grandes entreprises, construit sur la partie commune de l'échantillon
– $\widehat{\Delta}_c$: différence entre $\widehat{Y}^{\backslash 4}(s_t \cap s_{t-1})$ et $\widehat{Y}^{\backslash 4}(s_{t-1} \cap s_t)$
– $\widehat{Y}^{\backslash 4}(s_t \backslash s_{t-1})$ estimateur du total de Y (hors strate $s4$) au temps t, construit sur $s_t \backslash s_{t-1}$
– $\widehat{Y}^{\backslash 4}(s_{t-1} \backslash s_t)$ estimateur du total de Y (hors strate $s4$) au temps $t-1$, construit sur $s_{t-1} \backslash s_t$
– $\widehat{\Delta}_d$: différence entre $\widehat{Y}^{\backslash 4}(s_t \backslash s_{t-1})$ et $\widehat{Y}^{\backslash 4}(s_{t-1} \backslash s_t)$
– $s^2_{h \backslash s4}[z(t)]$: variance empirique corrigée de la variable z au temps t dans la strate $s_h(t) \backslash s4$
– $s_{h \backslash s4}[z(t), z(t-1)]$: covariance empirique corrigée entre z observée au temps t et z observée au temps $t-1$ dans la strate $(s_h(t) \cap s_h(t-1)) \backslash s4$

Références

Ardilly, P. et Tillé, Y. (2003). *Exercices corrigés de méthodes de sondage*. Ellipses, Paris.

Berger, Y. (2004). Variance estimation for measures of change in probability sampling. *Canadian Journal of Statistics*, 32, no.4 :451–467.

Binder, D. et Patak, Z. (1994). Use of estimating functions for estimation from complex surveys. *Journal of the American Statistical Association*, 89 :1035–1043.

Caron, N. et Ravalet, P. (2000). Estimation dans les enquêtes répétées : Application à l'enquête emploi en continu. Technical Report 0005, Méthodologie Statistique, INSEE, Paris.

Deville, J.-C. (1999). Variance estimation for complex statistics ans estimators : linearization and residual techniques. *Techniques d'enquête, Survey methodology*, 25 :193–204.

Deville, J.-C. et Särndal, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87 :376–382.

Goga, C. (2003). *Estimation de la variance dans les sondages à plusieurs échantillons et prise en compte de l'information auxiliaire par des modèles nonparamétriques*. PhD thesis, Université de Rennes II, Haute Bretagne, France.

Hulliger, B. (1999). Simple and robust estimators for sampling. In *Proceedings of the Section on Survey Research Methods*, pages 54–63. American Statistical Association.

Peters, R., Renfer, J.-P., et Hulliger, B. (2001). Statistique de la valeur ajoutée : procédure d'extrapolation des données. Technical report, Office fédéral de la Statistique.

Qualité, L. et Tillé, Y. (2006). *Méthodes d'enquêtes et sondages, pratiques européenne et nord-américaine*, Précision des estimations longitudinales dans l'enquête suisse sur la valeur ajoutée, pages 388–392. Dunod.

Qualité, L. et Tillé, Y. (2008). Variance estimation of changes in repeated surveys and its application to the swiss survey of value added. *Survey Methodology*, 34 :173–181.

Renfer, J.-P. (2000). Enquête sur la production et la valeur ajoutée : échantillonnage complémentaire. Technical report, Office fédéral de la Statistique.

Woodruff, R. (1971). A simple method for approximating de variance of a complicated estimate. *Journal of the American Statistical Association*, 66 :411–414.

Methodenberichte des Dienstes Statistische Methoden des BFS
Rapports de méthodes du Service de méthodes statistiques de l'OFS
Methodology reports published by the SFSO's Statistical Methods Unit

Qualité, L., Tillé, Y. (2009). Estimation de la précision d'évolutions dans l'enquête sur la valeur ajoutée. Numéro de commande: 338-0056

Renaud, A., Panchard, C. et Potterat, J. (2008). Statistique de l'emploi. Révision 2007 : méthodes d'estimation. Numéro de commande: 338-0055

Graf, E. (2008). Pondérations du PSM: PSM_I vague 8, PSM_II vague 3, PSM_I et PSM_II combinés. Numéro de commande: 338-0054

Andrade, B., Graf, M. (2008). Enquête suisse sur la structure des salaires 2006. Aspects méthodologiques du modèle des salaires "Salarium".
Numéro de commande: 338-0053

Renaud, A. (2008). Statistique de l'emploi. Révision 2007 : cadre de sondage et échantillonnage. Numéro de commande: 338-0052

Graf, E. (2008). Pondérations du SILC pilote: SILC_I vague 2, SILC_II vague 1, SILC_I et SILC_II combinés. Numéro de commande: 338-0051

Kilchmann, D. (2008). Statistik der sozialmedizinischen Institutionen 1999-2004 und Krankenhausstatistik 1999-2002. Einsetzungen für fehlende Daten. Bestellnummer: 338-0050

Renaud, A. (2008). Technologies de l'information et de la communication. Estimations sur la base de la statistique de la valeur ajoutée. Numéro de commande: 338-0049

Assoulin, D. (2007). Wertschöpfungsstatistik. Einsetzungsversuche für fehlende Antworten grosser Unternehmen. Bestellnummer: 338-0048

Kilchmann, D. (2007). Beherbergungsstatistik Campingplätze.Stichprobenrahmen und Schätzverfahren 2005/06. Bestellnummer: 338-0047

Gabler, S., Häder, S. (2007). Haushalts- und Personenerhebungen. Machbarkeit von Random Digit Dialing in der Schweiz. Bestellnummer: 338-0046

Ferrez, J., Graf, M. (2007). Enquête suisse sur la structure des salaires. Programmes R pour l'intervalle de confiance de la médiane. Numéro de commande: 338-0045

Renaud, A. (2007). Harmonisation de la scolarité obligatoire en Suisse (HarmoS). Design général de l'enquête et échantillon des écoles. Numéro de commande: 338-0044

Potterat, J. (2007). Betriebszählung 2005. Statistische Methoden zur Schätzung der provisorischen Ergebnisse. Bestellnummer: 338-0043

Hulliger, B. (2006). Umweltschutzausgaben der Unternehmen 2003, Stichprobenplan, Datenaufbereitung und Schätzverfahren. Bestellnummer: 338-0042

Renfer, J.-P. (2006). Enquête sur les chiffres d'affaires du commerce de détail. Plan d'échantillonnage et méthodes d'estimation. Numéro de commande: 338-0041

Salamin, P.-A. (2006). Statistique de l'aide sociale dans le domaine de l'asile. Plan de sondage et extrapolations pour l'enquête pilote 2005. Numéro de commande: 338-0040

Renaud, A. (2006). Statistique suisse des bénéficiaires de l'aide sociale. Pondération des communes 2004. Numéro de commande: 338-0039

Graf, M. (2006). Swiss Earnings Structure Survey 2002-2004. Compositional data in a stratified two-stage sample: Analysis and precision assessment of wage components. Order number: 338-0038

Potterat, J. (2006). Pensionskassenstatistik 2004. Statistische Methoden zur Schätzung der provisorischen Ergebnisse. Bestellnummer: 338-0037

Potterat, J. (2006). Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe im Jahr 2004. Stichprobenplan, Gewichtung und Schätzverfahren. Bestellnummer: 338-0036

Kilchmann, D. (2006). Vierteljährliche Wohnbaustatistik. Stichprobenplan, statistische Datenaufarbeitung und Schätzverfahren 2005. Bestellnummer: 338-0035

Kilchmann, D. (2006). Erhebung über Forschung und Entwicklung in der schweizerischen Privatwirtschaft 2004. Bereinigung der Stichprobe, Ersatz fehlender Werte und Schätzverfahren. Bestellnummer: 338-0034

Kilchmann, D., Eichenberger, P., Potterat, J. (2005). Volkszählung 2000. Statistische Einsetzungsverfahren Band 2. Bestellnummer: 338-0033

Kilchmann, D., Eichenberger, P., Potterat, J. (2005). Volkszählung 2000. Statistische Einsetzungsverfahren Band 1. Bestellnummer: 338-0032

Graf, M., Matei, A. (2005). Enquête suisse sur la structure des salaires 2002. La précision du salaire brut standardisé médian. Numéro de commande: 338-0031

Graf, E., Renfer, J.-P. (2005). Enquête suisse sur la santé 2002. Plan d'échantillonnage, pondération et estimation de la précision. Numéro de commande: 338-0030

Potterat, J. (2005). Mietpreis-Strukturerhebung 2003. Gewichtung und Schätzverfahren. Bestellnummer: 338-0029

Potterat, J. (2005). Landwirtschaftliche Betriebszählung 2003. Schätzverfahren für die Zusatzerhebung. Bestellnummer: 338-0028

Renaud, A. (2004). Coverage estimation for the Swiss population census 2000. Estimation methodology and results. Order number: 338-0027

Kilchmann, D. (2004). Revision des Schweizerischen Lohnindex. Schätzmethoden der Lohnindices und deren Varianzschätzer. Bestellnummer: 338-0026

Graf, M. (2004). Enquête suisse sur la structure des salaires 2002. Plan d'échantillonnage et extrapolation pour le secteur privé. Numéro de commande: 338-0025

Renaud, A. (2004). Analyse de données d'enquêtes. Quelques méthodes et illustration avec des données de l'OFS. Numéro de commande 338-0024

Renaud, A., Potterat, J. (2004). Estimation de la couverture du recensement de la population de l'an 2000. Echantillon pour l'estimation de la sous-couverture (P-sample) et qualité du cadre de sondage des bâtiments. Numéro de commande: 338-0023

Graf, M. (2004). Fusion de données. Etude de faisabilité. Numéro de commande: 338-0022

Potterat, J. (2003). Mietpreis-Strukturerhebung 2003. Entwicklung des Stichprobenplans und Ziehung der Stichprobe. Bestellnummer: 338-0021

Potterat, J. (2003). Landwirtschaftliche Betriebszählung 2003. Stichprobenplan der Zusatzerhebung. Bestellnummer: 338-0020

Renaud, A. (2003). Estimation de la couverture du recensement de la population de l'an 2000. Echantillon pour l'estimation de la sur-couverture (E-sample). Numéro de commande: 338-0019

Hulliger, B. (2003). Erhebung über Forschung und Entwicklung in der schweizerischen Privatwirtschaft 2000. Bereinigung der Stichprobe, Ersatz fehlender Werte und Schätzverfahren. Bestellnummer: 338-0018

Renfer, J.-P. (2003). Enquête 2000 sur la recherche et le développement dans l'économie privée en Suisse. Plan d'échantillonnage. Numéro de commande: 338-0017

Potterat, J. (2003). Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe. Schätzverfahren. Bestellnummer:
338-0016

Graf, M., Matei, A. (2003). Stratégie de choix des modèles de désaisonnalisation. Application aux séries de l'emploi total. Numéro de commande : 338-0015

Potterat, J., Salamin, P.A. (2002). Betriebszählung 2001. Methoden für die Datenbereinigung. Bestellnummer:
338-0014

Renaud, A. (2002). Programme international pour le suivi des acquis des élèves (PISA). Plans d'échantillonnage pour PISA 2000 en Suisse. Numéro de commande: 338-0013

Renfer, J.-P. (2002). Enquête 2001 sur les coûts et l'utilité de la formation des apprentis du point de vue des établissements. Plan d'échantillonnage. Numéro de commande: 338-0012

Potterat, J., Salamin, P.A. (2002). Betriebszählung 2001. Stichprobenplan und Schätzverfahren für die provisorischen Ergebnisse. Bestellnummer: 338-0011

Graf, M. (2002). Enquête suisse sur la structure des salaires 2000. Plan d'échantillonnage, pondération et méthode d'estimation pour le secteur privé. Numéro de commande: 338-0010

Renaud, A., Eichenberger P. (2002). Estimation de la couverture du recensement de la population de l'an 2000. Procédure d'enquête et plan d'échantillonnage de l'enquête de couverture. Numéro de commande: 338-0009

Kilchmann, D., Hulliger, B. (2002). Stichprobenplan für die Obstbaumzählung 2001. Bestellnummer: 338-0008

Graf, M. (2002). Passage du concept établissement au concept entreprise. Numéro de commande: 338-0007

Salamin, P.A. (2001). La technique de la double enquête pour la statistique du transport routier de marchandise. Numéro de commande: 338-0006

Peters, R., Renfer, J.-P. et Hulliger, B. (2001). Statistique de la valeur ajoutée 1997-1998. Procédure d'extrapolation des données. Numéro de commande: 338-0005

Potterat, J., Hulliger, B. (2001). Schätzung der Sägereiproduktion mit der Sägerei-Erhebung PAUL. Bestellnummer: 338-0004

Graf, M. (2001). Désaisonnalisation. Aspects méthodologiques et application à la statistique de l'emploi. Numéro de commande: 338-0003

Hüsler, J., Müller, S. (2001). Schlussbericht Betriebszählung 1995 (BZ 95), Mehrfach imputierte Umsatzzahlen. Bestellnummer: 338-0002

Renaud, A. (2001). Statistique suisse des bénéficiaires de l'aide sociale. Plan d'échantillonnage des communes. Numéro de commande: 338-0001

Hulliger, B., Eichenberger, P. (2000). Stichprobenregister für Haushalterhebungen: Umstellung auf Telefonnummern ohne Namen und Adressen, Abläufe für Erstellung und Stichprobenziehung. Bestellnummer: 338-0000

de Rossi, F.-X. (1998). Méthodes statistiques pour le compte routier suisse.

Hulliger, B., Kassab, M. (1998). Evaluation of Estimation Methods for the Survey on Environment Protection Expenditures of Swiss Communes.

Salamin, P.A. (1998). Etablissement d'une clef de passage pondérée entre l'ancienne (NGAE 85) et la nouvelle nomenclature (NOGA 95) générale des activités économiques.

Peters, R. (1998). Extrapolation des données de l'enquête de structure sur les loyers.

Bender, A., Hulliger, B. (1997). Enquête suisse sur la population active: rapport de pondération pour 1996.

Salamin, P.A. (1997). Evaluation de la Statistique de l'emploi.

Peters, R. (1997). Etablissement du plan d'échantillonnage pour l'enquête 1996 sur la recherche et le développement dans l'économie privée en Suisse.

Peters, R. (1997). Enquête 1996 sur la structure des salaires en Suisse: établissement du plan d'échantillonnage.

Peters, R. (1996). Pondération des données de l'enquête sur la famille en Suisse.

Comment, T., Hulliger, B., Ries, A. (1996). Gewichtungsverfahren für die Schweizerische Arbeitskräfteerhebung (1991-1995).

Hulliger, B. (1996). Haushalterhebung Familie 1994: Stichprobenplan, Stichprobenziehung und Reservestichproben.

Peters, R., Hulliger, B. (1996). Schätzverfahren für die Lohnstruktur-Erhebung 1994 / Procédure d'estimation pour l'enquête de 1994 sur la structure des salaires.

Peters, R. (1996). Schéma de pondération des indices PAUL.

Hulliger, B., Peters, R. (1996). Enquête sur le comportement de la population suisse en matière de transport en 1994: plan d'échantillonnage et pondération.

Hulliger, B. (1996). Gütertransportstatistik 1993: Schätzverfahren mit Kompensation der Antwortausfälle.

Salamin, P.A. (1995). Estimation des flux pour le module II des comptes globaux du marché de travail.

Peters, R. (1995). Enquête de structure sur les loyers: établissement d'un plan d'échantillonnage stratifié.

Hulliger, B. (1995). Konjunkturelle Mietpreiserhebung: Stichprobenplan und Schätzverfahren.

Schwendener, P. (1995). Verbrauchserhebung 1990 - Vertrauensintervalle.

Peters, R., Hulliger, B. (1994). La technique de pondération des données: application à l'enquête suisse sur la santé.

Hulliger, B., Peters, R. (1994). Enquête sur la structure des salaires en Suisse: stratégie d'échantillonnage pour le secteur privé.

Publikationsprogramm BFS

Das Bundesamt für Statistik (BFS) hat – als zentrale Statistikstelle des Bundes – die Aufgabe, statistische Informationen breiten Benutzerkreisen zur Verfügung zu stellen.

Die Verbreitung der statistischen Information geschieht gegliedert nach Fachbereichen (vgl. Umschlagseite 2) und mit verschiedenen Mitteln

Programme des publications de l'OFS

En sa qualité de service central de statistique de la Confédération, l'Office fédéral de la statistique (OFS) a pour tâche de rendre les informations statistiques accessibles à un large public.

L'information statistique est diffusée par domaine (cf. verso de la première page de couverture); elle emprunte diverses voies:

Diffusionsmittel	*Kontakt* *N° à composer*	*Moyen de diffusion*
Individuelle Auskünfte	032 713 60 11 info@bfs.admin.ch	Service de renseignements individuels
Das BFS im Internet	www.statistik.admin.ch	L'OFS sur Internet
Medienmitteilungen zur raschen Information der Öffentlichkeit über die neusten Ergebnisse	www.news-stat.admin.ch	Communiqués de presse: information rapide concernant les résultats les plus récents
Publikationen zur vertieften Information (zum Teil auch als Diskette/CD-Rom)	032 713 60 60 order@bfs.admin.ch	Publications: information approfondie (certaines sont disponibles sur disquette/CD-Rom)
Online-Datenbank	032 713 60 86 www.statweb.admin.ch	Banque de données (accessible en ligne)

Nähere Angaben zu den verschiedenen Diffusionsmitteln liefert das laufend nachgeführte Publikationsverzeichnis im Internet unter der Adresse www.statistik.admin.ch→Aktuell→Publikationen.

La liste des publications, mise à jour régulièrement, donne davantage de détails sur les divers moyens de diffusion. Elle se trouve sur Internet à l'adresse www.statistique.admin.ch →Actualités→Publications.

Methodenberichte des Dienstes Statistische Methoden
Rapports de méthodes du Service de méthodes statistiques
Methodology Reports by the Statistical Methods Unit

Die Methodenberichte beschreiben die mathematischen und statistischen Methoden, die den Resultaten und Analysen der öffentlichen Statistik zu Grunde liegen. Sie enthalten ausserdem die Evaluation und Entwicklung von neuen Methoden im Hinblick auf eine zukünftige Anwendung. Diese Publikationen sollen einerseits die verwendeten Methoden dokumentieren, um Transparenz und Wissenschaftlichkeit sicher zu stellen, und sie sollen andererseits die Zusammenarbeit mit den Hochschulen und der Wissenschaft fördern.

Zur Illustration der beschriebenen mathematischen Konzepte, werden im Bericht numerische Resultate aufgeführt. Diese sind allerdings nicht als offizielle Resultate der betreffenden Erhebungen zu verstehen. Ebenfalls können die tatsächlich angewendeten Methoden leicht von den hier beschriebenen abweichen.

Die Methodenberichte sind auf der Internetseite des BFS in elektronischer Form verfügbar.

Les rapports de méthodes décrivent les méthodes mathématiques et statistiques à la base des résultats et des analyses de la statistique publique. Ils présentent également l'évaluation et le développement de nouvelles méthodes en vue d'une application future. Ces publications visent d'une part à documenter les méthodes utilisées ou envisagées dans un souci de transparence et de rigueur scientifique, et d'autre part à favoriser la collaboration avec le monde scientifique et universitaire.

Les résultats numériques présentés dans les rapports de méthodes illustrent les concepts mathématiques décrits, mais ne sont pas des résultats officiels des enquêtes concernées. De même, les méthodes réellement appliquées peuvent différer légèrement de celles décrites dans ces rapports.

Les rapports de méthodes sont disponibles sous forme électronique sur le site internet de l'OFS.

L'Office fédéral de la statistique produit la statistique de la valeur ajoutée à partir d'un panel, c'est-à-dire d'un même échantillon utilisé pendant plusieurs années consécutives. L'usage de panels est une stratégie qui permet d'estimer précisément les évolutions.

Nous avons estimé la variance des estimateurs des évolutions en prenant en compte toutes les composantes de ceux-ci: le plan de sondage, le traitement des non-réponses, le traitement des grosses entreprises, la corrélation de la non-réponse d'une vague à l'autre, la robustification et le calage.

Les résultats montrent que si l'on prend en compte l'effet du panel, les variances des évolutions sont considérablement plus petite que les variances qui auraient été calculées sous une hypothèse d'indépendance entre deux vagues successives. Les variances sont en effet réduites de 60%. Les évolutions sont estimées de manière précise au niveau de l'ensemble des entreprises. Cependant, au niveau des classes d'activité définies par la Nomenclature générale des activités économiques (NOGA), à cause de la faible taille des échantillons nets, les mesures des évolutions sont beaucoup moins précises, à tel point que pour la moitié des strates d'activité, il n'est pas possible d'affirmer si l'évolution est croissante ou décroissante.

N° de commande: 338-0056
Commandes: 032 713 60 60 Fax: 032 713 60 61 E-Mail: order@bfs.admin.ch
Prix: gratuit
ISBN 978-3-303-00423-4