Rapport de méthodes

Technologies de l'information et de la communication

Estimations sur la base de la statistique de la valeur ajoutée



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

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

Neuchâtel, 2008

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 Verkehr und Nachrichtenwesen
12 Geld, Banken, Versicherungen
13 Soziale Sicherheit
14 Gesundheit
15 Bildung und Wissenschaft
16 Kultur, 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 Transports et communications
12 Monnaie, banques, assurances
13 Protection sociale
14 Santé
15 Education et science
16 Culture, 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

Technologies de l'information et de la communication

Estimations sur la base de la statistique de la valeur ajoutée

Auteur **Anne Renaud**
Office fédéral de la Statistique

Editeur **Office fédéral de la statistique**

Neuchâtel, 2008

Préambule

La section «Structures et conjoncture économiques» (SUKO) de l'Office fédéral de la statistique a demandé au «Service de méthodes statistiques» (METH) de faire une étude de faisabilité à propos de l'utilisation des données de la statistique de la valeur ajoutée pour faire des estimations dans le domaine des technologies de l'information et de la communication (TIC). Les résultats de l'étude préliminaire étant satisfaisants, des estimations ont été effectuées sur les données de la statistique de la valeur ajoutée de 1997 à 2003.

Merci à Philippe Stauffer, Gregory Rais et Pierre Sollberger de la section SUKO pour le suivi du projet. Merci également à Beat Hulliger, Paul-André Salamin et Jean-Pierre Renfer de METH pour leurs explications sur le fonctionnement de la WS, les méthodes utilisées et les fichiers de données. Merci encore à Pascal Utz de la section «Registre des entreprises et des établissements» pour les données de reclassification des cas spéciaux, ainsi qu'à Markus Däppen et Veronica Roth de la section «Statistiques monétaires des entreprises» (MON) pour les informations sur la WS. Merci finalement à Gregory Rais (SUKO), Pierre Sollberger (SUKO) et Paul-André Salamin (METH) pour la relecture du rapport.

Résumé

Les données de la statistique de la valeur ajoutée (WS) sont utilisées pour la production d'estimations dans le domaine des technologies de l'information et de la communication (TIC). La méthode utilisée pour la WS est adaptée afin de produire des estimations du total de la valeur de production brute, du total de la consommation intermédiaire et du total de la charge de personnel dans les TIC. Il s'agit d'un domaine transversal aux résultats habituels de la WS. La nouvelle méthode est appliquée sur les données de la période 1997 à 2003 afin d'estimer les totaux des grandeurs économiques d'intérêt et leur variabilité.

Mots-clé

rapport de méthodes; estimation; plan d'échantillonnage; domaine; technologie de l'information et de la communication; TIC.

Complément d'information: Anne Renaud, tél. 032 713 62 65
Anne.Renaud@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: 338-0049

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 2008
La reproduction est autorisée, sauf à des fins commerciales,
si la source est mentionnée.

ISBN: 978-3-303-00401-2

Table des matières

1 Introduction **5**

2 La statistique de la valeur ajoutée **5**

2.1 Population et échantillons 5

2.2 Relevé et analyses 6

2.3 Pondération 6

3 Définition du domaine TIC **8**

3.1 Nomenclatures NOGA 8

3.2 Application et cas particuliers 9

4 Méthode d'estimation d'un total **12**

4.1 Estimation d'un total 12

4.2 Estimation de la variance d'un total 13

4.3 Application 15

4.4 Un peu plus sur la méthode d'estimation 16

5 Jeux de données à disposition **16**

5.1 Fichiers de base 17

5.2 Fichiers de relevé 17

6 Résultats **18**

6.1 Reconstruction des résultats globaux 18

6.2 Estimation des totaux dans le domaine des TIC 19

6.3 Un peu plus sur le domaine des TIC 21

7 Conclusion **22**

A Code NOGA au niveau 2 **23**

B Résultats TIC par classe NOGA niveau 2 **24**

C Codes SAS **26**

C.1 Macros 26

C.2 Exemple pour 2003 31

Bibliographie **35**

1 Introduction

La section "Structures et conjoncture économiques" (SUKO) de l'Office fédéral de la statistique s'intéresse aux branches économiques d'activités productrices de technologies de l'information et de la communication (TIC). Les besoins d'analyses portent sur l'impact du secteur TIC sur la croissance de l'économie et de la productivité du travail.

Ce document présente une analyse du secteur TIC sur la base des données de la statistique de la valeur ajoutée (WS). Cette statistique annuelle auprès d'un échantillon d'entreprises permet d'estimer trois montants à la base des comptes nationaux : la valeur de production brute, la consommation intermédiaire et la charge de personnel. L'idée est de reprendre les données de la WS et de proposer une estimation des trois statistiques d'intérêt pour la sous-population des entreprises TIC.

Les données de la WS2002 ont été utilisées pour tester la faisabilité de l'analyse. Les résultats s'étant avérés intéressants pour la section SUKO et raisonnables, avec quelques contraintes d'utilisation, du point de vue des méthodes, les estimations ont été étendues à la WS2003, puis aux WS1997-WS2001. Des résultats sont donc disponibles pour la période 1997 à 2003. Il faut cependant noter que la qualité des résultats est moins bonne lorsque l'on remonte dans les années. Les contraintes d'utilisation et les réserves dans l'interprétation sont, elles, de plus en plus nombreuses.

Le présent rapport est relativement technique. Il documente les travaux effectués et présente la méthode d'estimation développée pour les TIC à partir de la méthode générale appliquée dans la WS. La section 2 présente les échantillons et la pondération de la WS. La section 3 décrit les travaux liés à la définition du secteur TIC. Les sections 4 et 5 présentent les méthodes d'estimation et les jeux de données disponibles. Les sections 6 et 7 présentent les résultats et la conclusion. Les programmes SAS et des résultats détaillés sont proposés dans l'annexe.

2 La statistique de la valeur ajoutée

La statistique de la valeur ajoutée (WS) est basée sur une enquête annuelle organisée par la section "Statistiques monétaires des entreprises" (MON) de l'OFS.

2.1 Population et échantillons

La population d'intérêt de la WS correspond à la quasi-totalité des entreprises des secteurs secondaire et tertiaire, à l'exception des banques et des assurances. Seules les entreprises ayant au moins 2 emplois en équivalents plein temps sont considérées.

Le cadre de sondage est extrait du Registre des entreprises et des établissements (REE) de l'OFS. Les règles d'extraction utilisent les codes d'activités économiques et les formes juridiques selon la définition fournie par la section "Comptes nationaux" (VGR). Pour des raisons historiques, le groupe de l'hôtellerie est traité séparément.

Les échantillons d'entreprises sont tirés pour 2 à 3 ans. Durant la période 1997 à 2003, on a un premier échantillon pour 1997 et 1998 (état de la base : fin 1997), un deuxième pour 1999 à 2001 (état de la base : fin 1999) et un troisième pour 2002 et 2003 (état de la base : 16 janvier 2003).

Chaque échantillon est tiré selon un plan stratifié. La stratification est définie par les deux

premiers chiffres - sur 5 - de l'activité économique selon la Nomenclature générale des activités économiques (NOGA, variable `noga2str`) et la taille de l'entreprise (nombre de plein temps) en 3 catégories (variable `clt`) ; voir Peters, Renfer et Hulliger (2001) pour des informations sur la statistique de la valeur ajoutée, et l'annexe A pour la liste des classes NOGA au niveau 2.

2.2 Relevé et analyses

Le questionnaire soumis aux entreprises se rapporte aux résultats comptables de l'exercice écoulé.

Les trois variables d'intérêt principales sont construites sur la base du relevé :
– Valeur de production brute (1000 CHF) : `T017`
– Consommation intermédiaire (1000 CHF) : `T022`
– Charge de personnel (1000 CHF) : `T053`

Les statistiques d'intérêt sont les totaux des trois variables d'intérêt $y =$ `T017`, `T022` et `T053` dans la population des entreprises U : $Y = \sum_{i \in U} y_i$.

Les estimations de totaux sont donnés au niveau global suisse et au niveau 2 de l'activité économique selon la NOGA.

On effectue des analyses transversales et non pas longitudinales. Les totaux et leurs variabilités sont donc estimés pour chaque année. On n'estime pas d'évolutions entre années successives.

Il est important de noter que l'unité d'analyse est l'entreprise. Ainsi l'activité économique est l'activité *principale* de l'entreprise. L'activité principale est l'activité majoritaire sur le plan des emplois exprimés en équivalents plein temps (EPT). Une entreprise formée d'un établissement (ou unité locale) de 10 EPT dans la classe NOGA '32.10A' (Fabrication de composants électroniques) et d'un établissement de 15 EPT dans la classe NOGA '74.13A' (Études de marché et sondages) sera classifiée dans la classe NOGA '74.13A' (Études de marché et sondages). Le code NOGA de l'entreprise est repris du registre au moment du tirage de l'échantillon. Il est modifié uniquement dans des cas avérés d'erreurs détectées lors du traitement des données.

2.3 Pondération

Les poids d'échantillonnage sont ajustés pour tenir compte de la non-réponse des entreprises, d'une robustification et d'un calage sur des variables auxiliaires (équivalents plein temps).

Les différentes étapes de la construction des poids sont décrites brièvement ci-dessous. Elles sont reprises de Peters, Renfer et Hulliger (2001), Hulliger (1999) et de la macro SAS de production `%estvar` écrite par Markus Eichenberger (Office fédéral de l'informatique et de la télécommunication).

Modifications préliminaires

Quelques modifications sont effectuées avant le calcul des poids (information reprise des documents de Paul-André Salamin, METH) :
– Les très grandes entreprises de chaque classe NOGA au niveau 2 relevée durant l'enquête (`noga2`) sont mises dans une classe de taille `taille=4` (4ème strate dans chaque NOGA

au niveau 2[1]. On considère alors que ces entreprises ont été tirées avec probabilité 1. Les autres entreprises restent dans leur classe de taille de stratification `taille=clt`.
– On définit des groupes (indice c) : `grou= noga2 * 10 + gros`, avec `gros`=0 si `taille` entre 1 et 3, et `gros`=1 si `taille`=4.
– Strates utilisées pour les estimations : `stra= noga2str * 10 + taille`, avec `noga2str` l'activité économique utilisée pour la définition des strates de tirage.
– Les entreprises avec un nombre d'emplois équivalent plein temps (EPT) relevé inférieur à 2 ont un poids 0 (hors population).

Poids d'échantillonnage

Le poids d'échantillonnage ω_i pour l'entreprise i dans la strate h est défini par :

$$\omega_i = \frac{1}{\pi_i} = \frac{N_h}{n_h}, \tag{1}$$

avec π_i la probabilité d'inclusion de i dans l'échantillon, N_h le nombre d'entreprises dans la strate h du cadre de sondage et n_h la taille de l'échantillon brut dans h.

Non réponse

Les cellules d'ajustement pour la non-réponse correspondent aux strates d'échantillonnage. Le facteur d'ajustement τ_i pour i dans h est défini par :

$$\tau_i = \frac{n_h}{m_h}, \tag{2}$$

avec m_h le nombre d'entreprises dans échantillon net dans h.

Le poids "initial" avec ajustement pour la non-réponse est noté $w_i = \omega_i\, \tau_i = N_h/m_h$.

Robustification

Un ajustement pour la robustification est appliqué au niveau des groupes `grou` (indice c). Le facteur d'ajustement u_i pour i dans le groupe c est donné par :

$$u_i = \min(u_1, u_2, u_3), \tag{3}$$

avec u_1 déterminé pour y =`T017`, u_2 pour y =`T022` et u_3 pour y =`T053` selon la méthode suivante :

$$u_i = \begin{cases} 1 & \text{si } |r_i| \leq c_0\widehat{\sigma} \\ c_0\widehat{\sigma}/|r_i| & \text{si } |r_i| > c_0\widehat{\sigma} \end{cases}, \tag{4}$$

avec

$$r_i = r_i(\beta_0) = (y_i - \beta_0 x_i)/\sqrt{x_i} \tag{5}$$
$$\beta_0 = \texttt{med}(y_i)/\texttt{med}(x_i) \tag{6}$$
$$\widehat{\sigma} = \texttt{med}(|r_i(\beta_0)|)/0.67 \tag{7}$$
$$c_0 = 10 \text{ (constante de robustification)} \tag{8}$$

et x_i le nombre d'équivalents plein temps selon l'enquête (variable `s14`).

Les calculs de β_0 et $\widehat{\sigma}$ se basent sur des médianes non pondérées.

[1] La classe des très grandes entreprises est définie par l'ensemble des env. 5% plus grandes entreprises. La limite est définie au niveau des EPT.

Calage

Un calage est effectué sur des totaux d'équivalents plein temps au niveau des groupes `grou` (indice c).

Les valeurs de référence X_c sont construites sur la base du recensement des entreprises. Une clé de passage est appliquée pour passer du concept entreprise au concept établissement. Les valeurs ainsi obtenues sont alors adaptées au fil des années pour suivre l'évolution des résultats de la Statistique de l'emploi (STATEM). Les résultats se réfèrent à l'univers WS.

Le facteur d'ajustement pour le calage g_i pour i dans le groupe c est donné par

$$g_i = \frac{X_c}{\sum_{i \in c} w_i u_i x_i}, \tag{9}$$

avec X_c le total de référence et x_i le nombre d'équivalents plein temps (variable `s14`) pour i.

Poids final

Le poids final $\tilde{w}_i$ pour i devient alors :

$$\tilde{w}_i = \omega_i \, \tau_i \, u_i \, g_i = w_i \, u_i \, g_i. \tag{10}$$

Application

La nouvelle macro `%outuigi` permet de déterminer explicitement les u_i de robustification, les g_i de calage et le poids final $\tilde{w}_i$ pour chaque unité de l'échantillon ; voir Annexe C. Cette décomposition des poids permet une flexibilité dans les estimations (voir notamment plus bas pour le domaine des TIC).

3 Définition du domaine TIC

Le domaine TIC est défini comme un ensemble d'activités économiques de la NOGA au niveau le plus détaillé (niveau 5, nommé "genre) ; voir Table 1 pour la liste selon la NOGA 2002. Seule l'activité principale de l'entreprise est prise en compte.

3.1 Nomenclatures NOGA

Trois nomenclatures NOGA sont en application durant la période de 1997 à 2003 : NOGA 1995, NOGA 1998 et NOGA 2002 ; voir Table 2. La NOGA 1995, partiellement adaptée pour la NOGA 1998, est utilisée pour l'échantillon 1997-1998. La NOGA 1998 est appliquée pour l'échantillon 1999-2001 et la NOGA 2002 est appliquée pour l'échantillon 2002-2003. Le traitement de certains codes n'a cependant pas toujours été effectué complètement ; voir exemple de 51.65A ci-dessous.

Les modifications entre les nomenclatures sont de deux types : transfert du numéro ou morcellement (split). Il n'y a pas de regroupements. Le morcellement de la classe '64.20A' en trois parties n'a pas d'implication majeure sur la définition des TIC car toutes ces classes restent

TAB. 1 Liste des 21 activités TIC selon la NOGA 2002.

NOGA 2002	Intitulé
30.02A	Fabrication appareils pour traitement données et équipements informatiques
31.30A	Fabrication câbles, conducteurs et fils isolés
32.10A	Fabrication composants électroniques
32.20A	Fabrication appareils et équipements émission et transmission
32.30A	Fabrication appareils réception, enregistrement ou reproduction son et image
33.20A	Fabrication instruments et appareils mesure, contrôle, navigation, etc.
33.30A	Fabrication équipements contrôle processus industriels
51.84A	Commerce de gros appareils pour trait. données et équipements info. périphériques
51.84B	Commerce de gros logiciels
51.86A	Commerce de gros composants électroniques
64.20A	Télécommunications (sans transmission radio et télévision)
64.20B	Transmission programmes radio et télévision
64.20C	Fourniture d'accès à Internet
71.33A	Location machines bureau, appareils trait. données et équipements informatiques
72.10A	Conseil en systèmes informatiques
72.21A	Édition de logiciels
72.22A	Conseil et développement de logiciels
72.30A	Traitement de données
72.40A	Banques de données
72.50A	Entretien et réparation machines bureau, appareils trait. données et équipements info.
72.60A	Autres activités liées à l'informatique

dans les TIC. De même, les deux transferts '51.64B' → '51.84A' et '51.64C' → '51.84B', et le morcellement de la classe '72.20A' demandent uniquement une adaptation de la liste.

Le morcellement de la classe '51.65A' (Commerce de gros d'autres machines et équipements pour l'industrie, le commerce et la navigation) est plus complexe. Cette classe contient une partie TIC et une partie hors TIC. Avant la NOGA 2002, le partage n'est possible que par l'utilisation de sources auxiliaires. Après la mise en place de la NOGA 2002, le morcellement permet théoriquement de différencier les deux parties car '51.86A' (Commerce de gros composants électroniques) est TIC et '51.87A' (Commerce de gros d'autres machines, équip. et accessoires) est hors TIC. En pratique, on verra qu'un traitement spécial est nécessaire pour toutes les années considérées.

3.2 Application et cas particuliers

Le code NOGA au niveau 5 est donné par la variable `s0007` dans les données WS. Ce code provient dans la grande majorité des cas du REE au moment du tirage de l'échantillon (variable `nogacd`). Il est modifié durant les travaux de la production uniquement dans des cas avérés d'erreurs.

Le partage de la classe '51.65A' en entreprises TIC et entreprises hors TIC est effectué grâce à des compléments d'information fournis par la section "Registre des entreprises et des établissements" (BUR) de l'OFS :

– 1997 - 1998 : reclassification des entreprises `s0007`='51.65A' basée sur le recensement des entreprises de 1998 (BZ98, données reçues de Pascal Utz, BUR) ;
– 1999 - 2001 : reclassification des entreprises `s0007`='51.65A' basée sur le recensement des entreprises de 2001 (BZ01, données reçues de Pascal Utz, BUR) ;

TAB. 2 Définition des TIC selon les classifications NOGA 1995, 1998 et 2002.

NOGA 1995	Modif	NOGA 1998	Modif	NOGA 2002
30.02A		30.02A		30.02A
31.30A		31.30A		31.30A
32.10A		32.10A		32.10A
32.20A		32.20A		32.20A
32.30A		32.30A		32.30A
33.20A		33.20A		33.20A
33.30A		33.30A		33.30A
51.64B		51.64B	→ transfert →	51.84A
51.64C		51.64C	→ transfert →	51.84B
51.65A*		51.65A*	→ split →	51.86A (51.87A)
64.20A	→ split →	64.20A 64.20B 64.20C		64.20A 64.20B 64.20C
71.33A		71.33A		71.33A
72.10A		72.10A		72.10A
72.20A		72.20A	→ split →	72.21A 72.22A
72.30A		72.30A		72.30A
72.40A		72.40A		72.40A
72.50A		72.50A		72.50A
72.60A		72.60A		72.60A

* : traitement spécial car classe partiellement TIC et partiellement hors TIC.

– 2002 : les entreprises de la classe ’51.65A’ ont toutes été codées ’51.87A’ et le partage en ’51.86A’ et ’51.87A’ n’est pas encore réalisé. On choisit donc de réutiliser les informations de la BZ01 pour faire le partage ;
– 2003 : trois entreprises sont codées ’51.86A’. Les modifications liées au partage de la classe ’51.87A’ ne sont cependant pas terminées et le fichier basé sur la BZ01 est à nouveau utilisé pour compléter les données (décision sur la base des informations de Markus Däppen, MON).

Notons que le transfert direct du code ’51.65A’ dans le code ’51.87A’ est lié à la maintenance du REE. Ce procédé se retrouve par exemple pour la classe ’72.20A’ (passage dans la classe ’72.22A’). Dans ce cas cependant, le détail des codes n’est pas en ordre mais les résultats globaux TIC ne sont pas perturbés. Ces exemples montrent qu’il est très important de travailler en étroite collaboration avec la section BUR lorsque des travaux se font à un niveau NOGA détaillé.

Le nombre d’entreprises TIC dans l’échantillon net de la WS est de 200-210 pour le premier échantillon, 250-260 pour le deuxième et 480-500 pour le troisième ; voir Table 3. Certaines classes NOGA sont très petites ou même vides. Le but des analyses est cependant d’obtenir des résultats au niveau global TIC. Le niveau 5 n’est utilisé que pour la définition du domaine. Au vu des tailles, les résultats de 2002 et 2003 ont de grandes chances d’avoir une meilleure précision que les années précédentes. Les résultats de 1997 et 1998, avec un maximum de 210 entreprises, risquent notamment d’être peu précis.

TAB. 3 Entreprises TIC dans les échantillons net WS.

NOGA 1998	1997	1998	1999	2000	2001	NOGA 2002	2002	2003
30.02A	10	10	9	10	13	30.02A	33	31
31.30A	7	8	10	9	8	31.30A	13	13
32.10A	34	37	50	50	44	32.10A	80	84
32.20A	12	9	14	12	9	32.20A	19	18
32.30A	5	4	3	5	5	32.30A	9	9
33.20A	41	39	43	49	48	33.20A	41	39
33.30A	5	4	8	8	8	33.30A	7	8
51.64B	26	26	18	20	17	51.84A	18	18
51.64C	2	4	2	1	1	51.84B	1	1
51.65A (TIC)	13	15	25	21	22			
51.65A (hors)	(67)	(68)	(64)	(45)	(43)			
						51.86A	0	3
						51.87A (TIC)	14	9
						51.87A (hors)	(41)	(37)
64.20A	7	5	10	12	9	64.20A	34	36
64.20B		1	0	1	2	64.20B	2	3
64.20C		0	1	1	0	64.20C	6	11
71.33A	0	0	0	0	0	71.33A	1	1
72.10A	7	5	7	9	7	72.10A	25	22
72.20A	33	28	45	41	46			
						72.21A	0	0
						72.22A	178	157
72.30A	4	4	4	2	5	72.30A	8	6
72.40A	0	1	3	4	4	72.40A	5	4
72.50A	1	2	1	2	2	72.50A	2	5
72.60A	3	1	2	2	1	72.60A	1	1
Total TIC	210	203	255	259	251		497	479
propTIC [%]	3.7%	3.6%	4.2%	4.2%	4.1%		7.4%	7.3%
Total hors TIC	5416	5497	5861	5891	5836		6188	6072
Total WS	5626	5700	6116	6150	6087		6685	6551

4 Méthode d'estimation d'un total

Les estimations officielles de la WS sont des totaux pour l'ensemble de la population d'intérêt et les classes NOGA au niveau 2. La méthode d'estimation est décrite brièvement ci-dessous. Pour plus de détails, voir Peters, Renfer et Hulliger (2001), Hulliger (1999) et la macro SAS `%estvar` de Markus Eichenberger (Office fédéral de l'informatique et de la télécommunication).

L'adaptation au domaine des TIC développée plus bas part du principe qu'une estimation directe est possible ; sans passer par des méthodes du type des petits domaines avec modèles ("small area estimation"). La méthode, décrite ci-dessous, reprend les idées des estimations officielles de la WS avec quelques adaptations pour passer au domaine des TIC. Les idées de base sont reprises de Hulliger (1999) et Cochran (1977).

4.1 Estimation d'un total

On définit l'estimateur du total global et l'estimateur du total dans un domaine de la population.

Estimation du total global

Le total Y pour la population globale est estimé sur la base de l'échantillon net s par :

$$\widehat{Y} = \sum_{i \in s} w_i \, u_i \, g_i \, y_i, \tag{11}$$

avec w_i le poids d'échantillonnage ajusté pour la non-réponse, u_i l'ajustement de robustification et g_i la correction due au calage.

Estimation du total dans le domaine d

Le total Y_d pour le domaine d est estimé sur la base de l'échantillon net s :

$$\widehat{Y}_d = \sum_{i \in s} w_i \, u_i \, g_i \, y_i \, I_{id}, \tag{12}$$

avec l'identificateur de domaine I_{id} pour l'entreprise i :

$$I_{id} = \begin{cases} 1 & \text{si } i \in d \\ 0 & \text{sinon} \end{cases} . \tag{13}$$

Dans le cadre de la production des résultats de la WS, les domaines correspondent aux classes NOGA au niveau 2. Ces domaines correspondent à des agrégats de groupes `grou` c (NOGA2 x 2 classes de taille). Dans le cas des TIC, le domaine est transversal aux groupes.

Nous remarquons que le jeux de poids WS est utilisé pour les estimations TIC. Cela signifie que les ajustements u_i et g_i sont estimés sur la base des groupes entiers c et non pas l'intersection $c \cap d$. On fait donc l'hypothèse que les entreprises du domaine d se comportent en moyenne de façon similaire à l'ensemble de leur groupe (robustification, non-réponse). L'avantage de cette méthode est de fournir des résultats de domaines dont le total pour la population est identique au total officiel ("cohérence"). Des alternatives sont discutées dans la section 4.4.

4.2 Estimation de la variance d'un total

On définit l'estimateur de la variance du total estimé global et dans un domaine de la population.

Estimation de la variance du total global

Références : Peters, Renfer et Hulliger (2001) et Hulliger (1999).

La variance $\widehat{var}(\widehat{Y})$ du total $\widehat{Y}$ est estimée dans la WS par la somme des estimations de la variance dans les différents groupes `grou` c (classe NOGA de relevé au niveau 2 croisée avec la variable `gros`). Elle prend en compte le calage effectué sur les valeurs de références X_c :

$$\widehat{var}(\widehat{Y}) = \sum_c \widehat{var}(\widehat{Y}_c), \tag{14}$$

avec la variance du groupe c :

$$\widehat{var}(\widehat{Y}_c) = \widehat{var}(X_c\, b_c) = X_c^2\, \widehat{var}(b_c) \tag{15}$$

et le coefficient :

$$b_c = \frac{\sum_{i\in c} w_i u_i y_i}{\sum_{i\in c} w_i u_i x_i}. \tag{16}$$

Le coefficient b_c est défini par le rapport entre les totaux estimés sur la base des w_i et u_i, voir pondération dans section 2.3.

A l'aide d'une linéarisation et d'une approximation (u_i fixes), on obtient :

$$\widehat{var}(b_c) = \frac{\widehat{var}(\sum_{i\in c} \tilde{e}_i)}{(\sum_{i\in c} w_i u_i x_i)^2}, \tag{17}$$

avec le résidu :

$$\tilde{e}_i = w_i u_i e_i = w_i u_i (y_i - b_c\, x_i). \tag{18}$$

En tenant compte du plan stratifié, on obtient :

$$\widehat{var}(\sum_{i\in c} \tilde{e}_i) = \sum_h \widehat{var}(\sum\nolimits_{i\in(c\cap h)} \tilde{e}_i), \tag{19}$$

avec :

$$\widehat{var}(\sum\nolimits_{i\in(c\,\cap\, h)} \tilde{e}_i) = (1-f_h)\, D_h^2\, n_h, \tag{20}$$

où $(1-f_h)$ est une correction de population finie, et D_h^2 est la variance de $\tilde{e}$ dans h.

L'estimation du facteur de correction de population finie $1-f_h$ est donné par :

$$(1-f_h) = 1 - \frac{1}{\bar{w}_h} = 1 - \frac{n_h}{\sum_{i\in h} w_i}. \tag{21}$$

On remarque que le facteur de correction prend en compte les poids initiaux (échantillonnage et correction non réponse) et qu'il est calculé sur la strate entière (pas seulement sur la partie correspondant au groupe).

L'estimateur D_h^2 est donné par :

$$D_h^2 = \frac{1}{n_h - 1} \sum_{i\in h} (\tilde{\varepsilon}_i - \bar{\tilde{\varepsilon}}_h)^2, \tag{22}$$

avec :

$$\tilde{\varepsilon}_i = \tilde{e}_i \, I_{ic} \tag{23}$$

$$\bar{\tilde{\varepsilon}}_h = \frac{\sum_{i \in h} \tilde{\varepsilon}_i}{n_h} \tag{24}$$

et

$$I_{ic} = \begin{cases} 1 & \text{si } i \in c \\ 0 & \text{sinon.} \end{cases} \tag{25}$$

Estimation de la variance du total dans le domaine d

L'estimateur de la variance du total dans un domaine dépend du type de domaine considéré.

Domaine = ensemble de groupes `grou`

La variance de $\widehat{Y}_d$ dans un domaine d défini par un ensemble Ω de groupes c est estimée comme suit :

$$\widehat{var}(\widehat{Y}_d) = \sum_{c \in \Omega} \widehat{var}(\widehat{Y}_c). \tag{26}$$

C'est la méthode utilisée dans la WS pour toutes les estimations pour les domaines correspondants aux NOGA niveau 2.

Domaine "à travers" les groupes `grou`

Lorsque le domaine, tel celui des TIC, coupe les groupes `grou` c, on peut adapter l'estimateur global au cas des domaines.

La variance *var*(Y_d) du total Y_d peut être estimée par :

$$\widehat{var}(\widehat{Y}_d) = \sum_{c} \widehat{var}(\widehat{Y}_{c,d}), \tag{27}$$

avec :

$$\widehat{var}(\widehat{Y}_{c,d}) = \widehat{var}(X_{c,d} \, b_{c,d}) = X_{c,d}^2 \widehat{var}(b_{c,d}). \tag{28}$$

La valeur de $X_{c,d}$ n'étant pas connue, on fait dans notre cas l'hypothèse que la proportion d'équivalents plein temps dans le domaine est identique à celle observée dans l'échantillon et que la variance de $X_{c,d}$ est nulle :

$$X_{c,d} = X_c \frac{\sum_{i \in c \cap d} w_i u_i x_i}{\sum_{i \in c} w_i u_i x_i} = X_c \frac{\widehat{X}_{c,d}}{\widehat{X}_c}. \tag{29}$$

La valeur de $b_{c,d}$ est calculée comme suit :

$$b_{c,d} = \frac{\sum_{i \in c \cap d} w_i u_i y_i}{\sum_{i \in c \cap d} w_i u_i x_i} = \frac{\sum_{i \in c} w_i u_i y_i I_{id}}{\sum_{i \in c} w_i u_i x_i I_{id}}, \tag{30}$$

avec I_{id}=1 si $i \in d$ et 0 sinon.

Notons cependant que l'on aurait pu choisir $b_{c,d} = b_c$ pour avoir plus de stabilité dans l'estimation (plus de données). La valeur choisie est calculée dans $c \cap d$ pour avoir une cohérence générale avec les autres parties de l'estimation.

La variance $\widehat{var}(b_{c,d})$ est donnée par :

$$\widehat{var}(b_{c,d}) = \frac{\widehat{var}(\sum_{i\in c\cap d}\tilde{e}_i)}{(\sum_{i\in c\cap d} w_i u_i x_i)^2}, \tag{31}$$

avec :

$$\tilde{e}_i = w_i u_i e_i = w_i u_i (y_i - b_{c,d}\, x_i). \tag{32}$$

En tenant compte du plan stratifié, on obtient :

$$\widehat{var}(\sum_{i\in c\cap d}\tilde{e}_i) = \sum_h \widehat{var}(\sum_{i\in(c\cap d\cap h)}\tilde{e}_i), \tag{33}$$

avec :

$$\widehat{var}(\sum_{i\in(c\cap d\cap h)}\tilde{e}_i) = (1-f_h) D_h^2 n_h. \tag{34}$$

L'estimation du facteur de correction de population finie $1-f_h$ est donné par :

$$(1-f_h) = 1 - \frac{1}{\bar{w}_h} = 1 - \frac{n_h}{\sum_{i\in h} w_i}. \tag{35}$$

Il est fixé indépendamment du domaine d.

L'estimateur D_h^2 est donné par :

$$D_h^2 = \frac{1}{n_h - 1}\sum_{i\in h}(\tilde{\varepsilon}_i - \bar{\tilde{\varepsilon}}_h)^2, \tag{36}$$

avec

$$\tilde{\varepsilon}_i = \tilde{e}_i\, I_{icd} \tag{37}$$

$$\bar{\tilde{\varepsilon}}_h = \frac{\sum_{i\in h}\tilde{\varepsilon}_i}{n_h} \tag{38}$$

et

$$I_{icd} = \begin{cases} 1 & \text{si } i \in c\cap d \\ 0 & \text{sinon.} \end{cases} \tag{39}$$

Il dépend du domaine d.

4.3 Application

Estimations WS

La macro `%estvar` de Markus Eichenberger appliquée en production WS donne des résultats finaux par groupe `grou` et classes NOGA au niveau 2 `noga2`. Les résultats peuvent être agrégés au niveau 2 de la NOGA avec `%agg`.

La nouvelle macro `%EstVarDec` fournit l'estimation du total et de la variance pour les groupes `grou` et l'ensemble de la population. L'agrégation des résultats par NOGA au niveau 2 est donnée par la nouvelle macro `%aggResNoga`. Le fichier créé par `%outuigi` (détail des poids) est utilisé en input, tout comme les informations de référence pour le calage.

Les résultats des nouvelles macros sont identiques à ceux obtenus avec les macros `%estvar` et `%agg` de Markus Eichenberger. L'avantage de leur utilisation est cependant de sortir les résultats sous la forme d'un data set directement utilisable dans SAS et non seulement d'un output "print".

Estimations pour les TIC

La nouvelle macro `%EstVarDecDom` fournit l'estimation du total et de la variance pour le domaine TIC (identificateur `iddom`=0 ou 1) dans les groupes `grou` et l'ensemble de la population. L'agrégation des résultats par NOGA au niveau 2 est donnée par la nouvelle macro `%aggResNogaDom`. Le fichier créé par `%outuigi` (détail des poids) est utilisé en input de `%EstVarDecDom`.

Les estimations dans le domaine des TIC interviennent après regroupement manuel des strates non exhaustives avec moins de 5 entreprises TIC. Cette procédure permet de stabiliser l'estimation de la variance. Les regroupements sont effectués en priorité dans la même classe NOGA au niveau 2 (regroupement de classes de taille). Si cela n'est pas suffisant pour atteindre la taille de 5, le regroupement s'effectue avec la même classe de taille dans la classe NOGA au niveau 1. On effectue de 6 à 10 regroupements de strates pour les estimations de 1997 à 2003.

Les macros créées pour le projet des TIC et un exemple de code pour l'application sur les données de 2003 sont disponibles dans l'Annexe C.

4.4 Un peu plus sur la méthode d'estimation

Dans la méthode d'estimation proposée, les ajustements u_i et g_i sont déterminés sur la base des groupes entiers c et non pas les intersections $c \cap d$. La méthode a l'avantage de donner des résultats de domaines dont la somme correspond au résultat global (cohérence) et d'utiliser un unique jeu de poids pour toutes les estimations. Cette méthode repose cependant sur l'hypothèse que les entreprises du domaine d se comportent en moyenne de façon similaire à l'ensemble de leur groupe.

On peut envisager deux modifications :

1. Robustification : les u_i pourraient être estimés dans les sous-ensembles $c \cap d$ au lieu des groupes c. Quelques sous-ensembles seraient cependant très petits, *i.e.* estimateurs peu stables.

2. Calage : les valeurs de référence pourraient être déterminées au niveau $X_{c\cap d}$ pour la détermination des g_i (pas disponible dans notre cas). La valeur g_i calculée dans le groupe entier est cependant équivalente à la valeur calculée sur l'intersection entre le groupe et le domaine si on fait l'hypothèse suivante : "la proportion d'emplois dans le domaine selon la population de référence est identique à celle estimée sur la base de l'échantillon" : $X_{c,d}/X_c = \widehat{X}_{c,d}/\widehat{X}_c$.

Les deux modifications auraient pour conséquence une perte de cohérence dans les agrégats et l'abandon du jeu de poids unique pour les estimations globales et TIC sur les données de la WS. Elles n'ont pas été retenues.

5 Jeux de données à disposition

Les jeux de données à disposition sont de deux types : (1) les fichiers de base utilisés dans la production des résultats officiels WS et (2) les fichiers de travail de la section MON qui contiennent plus d'informations sur le relevé.

5.1 Fichiers de base

Deux fichiers de base sont utilisés pour les estimations WS pour les années xx $\in \{97, 98, 99, 00, 01, 02, 03\}$:

- ech1WSxx : échantillon net de l'année xx considérée. Le poids d'extrapolation est donné par la variable gew. Le fichier contient les variables principale T017, T022 et T053 tout comme la variable de l'emploi s14, la NOGA du cadre noga2str et de relevé noga2, et la variable grou.
- pro2WSxx : valeurs de référence pour le calage de l'année xx considérée. Une valeur par combinaison de noga2 et gros (binaire), i.e. variable grou.

La variable définissant la classe NOGA au niveau 2 dans les estimations est noga2, le code selon le relevé de la statistique de l'année en cours.

La strate pour les calculs de variance est donnée par la variable stra = noga2st*10+taille, avec noga2str, la classe NOGA du cadre de sondage et taille, la classe de taille de l'entreprise définie sur la base de la classe de taille clt du cadre de sondage ; voir les modifications préliminaires dans la Section 2.3. La variable stra peut être construite sur la base des données de ech1WSxx.

L'échantillon net utilisé pour les estimations est un extrait de l'échantillon ech1WSxx de l'année considérée. L'idée est d'exclure les données hors population, *i.e.* les entreprises avec moins de 2 EPT voient leur poids fixé à 0 durant la pondération[2].

5.2 Fichiers de relevé

Les fichiers de base utilisés pour les estimations officielles ne contiennent pas l'information sur l'activité économique au niveau 5. Il faut donc utiliser des fichiers contenant l'ensemble des valeurs relevées durant l'enquête.

Nous avons à disposition les fichiers de relevé sburxx, pour xx $\in \{98, 99, 00, 01, 02, 03\}$. Ces fichiers contiennent la variable s0007 correspondant à l'activité économique de relevé au niveau 5. Un appariement entre ech1WSxx et sburxx permet d'intégrer cette information aux données de l'échantillon net[3].

Le fichier sbur97 n'est pas disponible pour 1997. On prend alors les fichiers sfile97 et sfile_n qui contiennent également les données du relevé mais dans un état pas connu de manière exacte. Ces deux fichiers ont des informations identiques dans les variables qui nous intéressent.

Cas particuliers

Il y a une incertitude sur le statut des fichiers disponibles pour 1997 et 1998. L'ancienneté de ces données ne permet en effet pas d'assurer que nous avons les fichiers de référence

[2] Nous notons que l'extrait effectué pour la production des résultats officiels utilise le critère 'if s14>1', avec s14 le nombre d'équivalents plein temps relevés dans l'enquête, et non pas 'if gew ne 0' ou, équivalent, 'if s14>=2'. Il reste donc quelques entreprises avec poids 0 dans les estimations. Les totaux estimés ne sont pas influencés mais la variance peut être légèrement modifiée.

[3] Cet appariement utilise l'identificateur de l'entreprise burnr. Les entreprises sont identifiées par le burnr de la tête dans le registre des entreprises et établissements et par s0001 dans la section MON. Note : avant l'appariement, il faut intégrer quelques modifications d'identificateurs (variable burnr) effectuées par la section MON en 1999, 2000, 2001 et 2003.

utilisés dans les estimations officielles. Quelques cas spéciaux sont listés ci-dessous dans un but documentaire. Ils ne touchent cependant pas le domaine des TIC.

Remarques sur le fichier ech1WS97 :
- 16 observations ont des valeurs manquantes dans les variables T017, T022 et T053. Il ne s'agit pas d'entreprises dans le domaine des TIC (NOGA niveau 2 possédant des TIC touchées mais pas les NOGA niveau 5 des TIC selon information dans sfile97).
- 51 observations ont une valeur grou=0 ou 1 au lieu d'un code à 3 positions. On reconstitue la variable sur la base de la NOGA au niveau 2 et de la variable taille : grou=noga2*10+gros (rem : 14 cas dans classe noga2=64, gros=0 et 1).
- Comparaison des variables communes avec sfile97 : 4 modifications de T022, 1 modification de T053, 256 modifications de noga2 (52→51, 56→55), et 16 différences dues aux valeurs manquantes dans ech1WS97 mais existantes dans sfile97. Aucune entreprise TIC n'est touchée.
- L'échantillon ech1WS97 contient 5629 observations. Trois observations sont retirées avec le critère s14>1 avant les estimations. Il reste donc 5626 observations (dont 24 avec poids 0, voir remarque dans Section 5.1).

Remarques sur les données de 1998 :
- pas de valeurs manquantes dans les variables importantes de ech1WS98.
- Comparaison de ech1WS98 avec sbur98 :
  - un identifiant s0001 existe dans ech1WS98 mais pas dans sbur98 et un autre existe dans sbur98 mais pas dans ech1WS98. On dirait une modification d'identificateur mais les informations sur la NOGA niveau 2 et les variables T0xx sont très différentes. Sans information claire, on laisse les données tel quel car les TIC ne sont pas touchés.
  - 201 modifications de NOGA niveau 2 (56→55) : ne touche pas les TIC.
  - Énormes différences dans les variables T0xx pour une entreprise (noga5='5146A'). On laisse tel quel car cela ne touche pas les TIC.

6 Résultats

Les résultats globaux pour l'ensemble de la population et les différentes classes NOGA sont calculés sur la base des fichiers disponibles avant de faire des estimations dans le domaine des TIC. Cette étape préliminaire permet de contrôler que les fichiers à disposition sont bien les fichiers utilisés dans les estimations officielles. Elle permet également de définir les estimations globales de référence pour l'estimation de la proportion des TIC.

Pour rappel : les résultats numériques présentés dans les rapports de méthodes illustrent les concepts mathématiques décrits, mais ne sont pas des résultats officiels des enquêtes concernées.

6.1 Reconstruction des résultats globaux

Les totaux estimés sont identiques aux résultats officiels pour toutes les années sauf 1997, 1998 et 2002. Les tailles d'échantillon et leur répartition dans les NOGA au niveau 2 sont identiques, mais des différences sont observées dans les totaux et, dans une moindre mesure, les CV.

La différence en 2002 est très faible. Elle se trouve dans la classe NOGA=56 (hôtels, classe avec traitement spécial) et ne touche donc pas le domaine des TIC.

Les différences en 1997 et 1998, avec un maximum de 0.7% sur les totaux en 1997, proviennent très certainement des problèmes cités dans les cas spéciaux de la Section 5.2. Les valeurs de calage sont identiques.

6.2 Estimation des totaux dans le domaine des TIC

Les estimations des totaux TIC pour les trois variables correspondent à 5.5%-9.2% des totaux estimés globaux ; voir Table 4. La précision des estimations est clairement moins bonne pour les TIC que pour les totaux globaux avec des coefficients de variation entre 4.2% et 18.7%. Il est probable que les estimations des coefficients de variation soient elles-même également relativement variables.

TAB. 4 Estimations dans le domaine des TIC pour les trois variables (valeur en 10^9 CHF), avec le nombre d'entreprises total `n` et TIC `nTIC` dans l'échantillon, les totaux estimés `Tot` (valeurs officielles) et TIC `TotTIC`. Les coefficients de variations `CvTot` et `CvTIC` et la variable `PropTIC=TotTIC/Tot` sont exprimés en %.

T017	n	nTIC	Tot	CvTot	TotTIC	CvTIC	PropTIC
1997	5626	210	446.3	1.1	24.3	15.8	5.5
1998	5700	203	474.1	1.2	28.1	6.8	5.9
1999	6116	255	479.2	1.1	30.5	12.7	6.4
2000	6150	259	511.7	1.0	40.6	6.3	7.9
2001	6087	251	535.1	1.0	41.9	5.8	7.8
2002	6685	497	564.8	1.0	51.9	5.1	9.2
2003	6551	479	544.1	1.6	49.0	7.6	9.0

T022	n	nTIC	Tot	CvTot	TotTIC	CvTIC	PropTIC
1997	5626	210	230.9	1.4	11.7	10.9	5.0
1998	5700	203	249.8	1.9	13.9	11.5	5.6
1999	6116	255	250.5	1.4	15.5	13.7	6.2
2000	6150	259	269.7	1.2	19.5	7.3	7.2
2001	6087	251	284.2	1.3	20.6	7.5	7.3
2002	6685	497	307.7	1.4	27.9	6.3	9.1
2003	6551	479	291.7	1.5	25.6	6.5	8.8

T053	n	nTIC	Tot	CvTot	TotTIC	CvTIC	PropTIC
1997	5626	210	153.4	1.0	9.3	9.0	6.1
1998	5700	203	156.9	1.1	10.0	18.7	6.4
1999	6116	255	162.0	1.1	11.3	10.0	7.0
2000	6150	259	170.8	0.9	14.2	5.4	8.3
2001	6087	251	182.5	0.9	16.1	4.2	8.8
2002	6685	497	194.5	0.6	17.0	4.4	8.8
2003	6551	479	191.3	0.6	15.2	5.9	8.0

Les totaux TIC des trois variables tendent à croître en 1997 et 2002 puis à se stabiliser entre 2002 et 2003 ; voir Figure 1. Les intervalles de confiance de deux années successives se recoupent dans tous les cas.

Il est important de garder en mémoire que seul trois échantillons sont utilisés pour les estimations des sept années de 1997 à 2003. Les nouvelles entreprises ne sont pas intégrées chaque année dans la statistique. Ainsi, une entreprise créée en 2000 a une chance d'entrer dans l'échantillon seulement à partir de 2002. Le calage effectué durant l'estimation prend en compte l'évolution du REE entre les années. Le calage est cependant intégré de manière globale par NOGA niveau 2 et non pas différenciée selon les secteurs TIC et non TIC. Une partie des modifications de structures dans les TIC n'est donc pas mesurable sur la base d'une enquête telle la WS. Cet effet est inexistant dans les classes NOGA entièrement TIC (NOGA 32 et 72). Il peut être important dans les classes NOGA avec mélange d'entreprises TIC et non

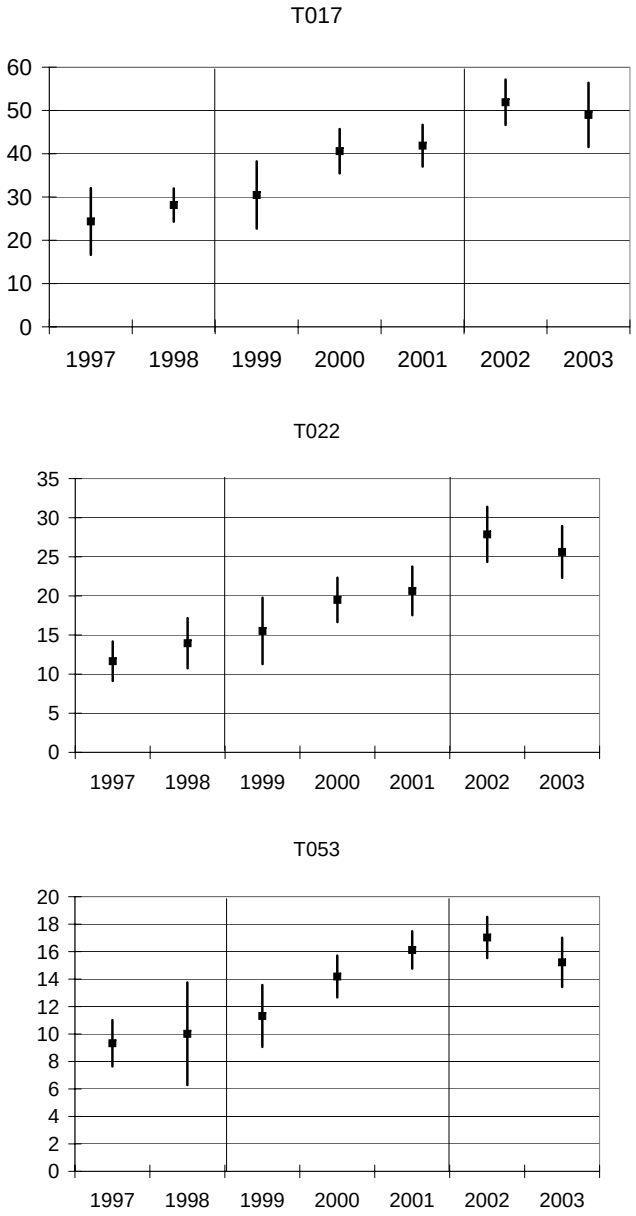


FIG. 1 Estimations TIC pour les variables `T017`, `T022` et `T053` (en 10^9 CHF), avec estimation des intervalles de confiance à 95%.

TIC.

Il est bon de noter également que les disparitions d'entreprises de l'échantillon de la WS sont traitées comme de la non-réponse. La pondération est basée sur l'ensemble des répondants, sans ajout de l'information de la mort éventuelle d'entreprises. Le calage sur les données de l'année de référence est ainsi prévu pour rectifier les effets combinés des naissances et des morts d'entreprises.

Les résultats des estimations TIC par classe NOGA au niveau 2 sont fournis dans l'annexe B. On note les coefficients de variation élevés dans presque toutes les classes. Ces résultats sont intéressants dans un but de documentation. Les résultats TIC à ce niveau ne sont cependant pas assez précis (biais et variance) pour être utilisés comme sous-domaines si la classe NOGA au niveau 2 n'est pas entièrement TIC.

6.3 Un peu plus sur le domaine des TIC

Sur la base des échantillons, la proportion d'entreprises TIC dans la population globale est estimée de l'ordre de 3.2 à 3.6% entre 1997 et 2000 ; voir Table 5. Elle passe à 2.8% en 2000 puis à plus de 4% en 2002 et 2003. Les estimations de la proportion d'équivalents plein temps dans les TIC semblent plus réalistes, avec une croissance puis une stabilisation comme dans le cas des totaux. En 2000, la chute du nombre d'entreprises est compensée par une augmentation des EPT estimés, au niveau global et dans les TIC. Ces remarques mènent à la recommandation suivante : une étude du domaine des TIC et de son évolution devrait être réalisée sur la base du registre des entreprises et des établissements. Cette étude, hors du cadre décrit dans ce document, compléterait judicieusement les résultats obtenus sur la base des échantillons de la WS.

TAB. 5 Estimation de la part TIC des entreprises et des EPT, avec les nombres d'entreprises `n` et `nTIC` dans l'échantillon, les tailles estimées `npond` et `nTICpond`, la proportion `pTIC=nTICpond/npond` de TIC [%], les EPT `EPT` et `EPTTIC` (valeurs en 1000 EPT) et la proportion `pTICEPT=EPTTIC/EPT` [%].

WS	n	nTIC	npond	nTICpond	pTIC	EPT	EPTTIC	pTICEPT
1997	5626	210	156455	5061	3.2	2141	93	4.3
1998	5700	203	155225	5661	3.6	2150	104	4.9
1999	6116	255	141268	4842	3.4	2165	112	5.2
2000	6150	259	145953	5212	3.6	2221	135	6.1
2001	6087	251	143062	3972	2.8	2285	150	6.6
2002	6685	497	153062	7098	4.6	2374	149	6.3
2003	6551	479	141754	5995	4.2	2314	136	5.9

Une étude plus détaillée sur les données 2002 montre que la proportion de TIC dans le cadre d'échantillonnage est semblable à la proportion de TIC dans les échantillons brut et net pour la grande majorité des 22 strates de tirage. Les différences importantes sont observées uniquement dans les petites strates. Globalement, on observe une proportion de TIC égale à 33% dans le cadre, 43% dans le brut et 39% dans le net. Le taux de réponse des entreprises TIC vaut 57%. Il est inférieur au taux global de 62%. On note également que 6.4% de l'ensemble des entreprises et 7.7% des entreprises TIC font partie de l'échantillon net.

La relation entre les variables d'intérêt `T017`, `T022` et `T053` et le nombre d'équivalents plein temps est utilisé dans le calage. Sur la base des données de 2002, on note que la relation peut varier de façon sensible entre TIC et non TIC dans un même groupe d'estimation c. Les différences les plus sensibles s'observent cependant dans les cas habituels d'estimations difficiles tels une petite taille de l'échantillon.

Les hypothèses considérées dans les estimations pour le domaine des TIC semblent raisonnables au niveau général. Il faut cependant noter que ces hypothèses ne sont plus contrôlées

dans le cas de résultats à un niveau plus détaillé que le domaine des TIC dans son ensemble. Cette remarque est par exemple valable pour les résultats au niveau 2 des classes NOGA, notamment si la classe NOGA ne fait que partiellement partie des TIC.

La définition des TIC dans les données de la WS est complexe. Les informations se réfèrent à l'entreprise et le code NOGA au niveau 5 est de moins bonne qualité que celui au niveau 2 par exemple. On prend également conscience de ce flou de définition en comparant le code NOGA des entreprises présentes dans l'échantillon net plusieurs années de suite. Ainsi, sur les 2082 entreprises présentes dans tous les échantillons de 1999 à 2003, 1972 sont toujours hors TIC, 87 sont toujours dans les TIC et 23 entrent ou sortent des TIC. Sans information complémentaire, nous ne savons pas si ces modifications sont réelles ou dues aux difficultés de la définition.

7 Conclusion

Les données de la statistique de la valeur ajoutée (WS) ont permis de faire des estimations spécifiques au secteur TIC pour la période de 1997 à 2003[4]. La méthode d'estimation se base sur la méthode appliquée pour les résultats globaux officiels de la WS. Elle prend en compte le fait que l'ensemble des TIC forme un domaine de la population WS.

Quelques difficultés d'estimations sont apparues lors de l'analyse. On note par exemple la complexité de la définition du domaine TIC dans une enquête auprès des entreprises. Ceci d'autant que le domaine se définit sur la base de l'activité économique au niveau le plus détaillé de la NOGA. De plus, le fait de conserver le même échantillon durant 2 à 3 ans peut potentiellement amener des biais dans les estimations dans un domaine fluctuant tel que les TIC.

Il serait intéressant de compléter l'analyse présentée dans ce rapport par une analyse des TIC dans le Registre des entreprises et des établissements (REE) entre 1997 et 2003. Cette analyse permettrait notamment d'évaluer l'évolution du nombre d'entreprises et d'emplois ; sans les sauts dus aux changements d'échantillons. Les données économiques ne sont pas disponibles mais l'intérêt serait de mieux comprendre la démographie des entreprises TIC et, peut-être également, de comparer l'approche entreprise de la WS avec une approche établissements telle que disponible dans le REE.

La méthode d'estimation développée pour le domaine TIC est applicable pour d'autres domaines. Les hypothèses de travail devraient cependant être évaluées pour chaque nouveau domaine afin d'éviter des biais dans les estimations. Une taille minimale d'échantillon et une définition précise du domaine sont par exemple requises pour faire des estimations directes en domaine. Des estimations selon la technique des petits domaines avec choix de modèles ("small area estimations") sont à évaluer aussi en parallèle avec la méthode directe.

[4] Note : les estimations sont effectuées annuellement dès 2004 par la section SUKO.

A Code NOGA au niveau 2

TAB. 6 Activités économiques NOGA niveau 2 de la WS.

noga2	
14	Extraction de la pierre et de la terre, autres industries extractives
15	Industries alimentaires et industrie des boissons
16	Industrie du tabac
17	Industrie textile
18	Industrie de l'habillement
19	Industrie du cuir et d'articles en cuir
20	Travail bois et fabrication articles en bois, liège, vannerie et sparterie (sans fabrication meubles)
21	Industrie du papier, du carton et de leurs dérivés
22	Edition, impression, reproduction d'enregistrements sonores, vidéo et informatiques
23	Cokéfaction ; raffinage ; traitement nucléaire
24	Industrie chimique
25	Fabr. de caoutchouc et matières plastiques
26	Fabr. d'autres prod. minéraux non métalliques
27	Métallurgie
28	Travail des métaux
29	Fabrication de machines et d'équipements
30	Fabr. de machines de bureau, ordinateurs
31	Fabr. de machines et appareils électriques
32	Fabr. d'équip. de radio, TV, communication
33	Instruments médicaux/précision, horlogerie
34	Industrie automobile
35	Fabrication d'autres moyens de transport
36	Fabrication meubles, bijoux, instruments musique, articles sport, jeux, jouets et autres produits
37	Récupération
40	Production et distribution d'électricité, de combustibles gazeux et de chaleur
41	Captage, traitement et distribution d'eau
45	Construction
50	Commerce et réparation de véhicules
51	Commerce de gros, intermédiaires du com.
52	Commerce de détail, rép. articles domestiques
55	Hôtellerie et restauration
56	Classe des hôtels (**spécial WS**)
60	Transports terrestres et par conduites
61	Transports par eau
62	Transports aériens
63	Auxiliaires des transp., agences de voyage
64	Postes et télécommunications
70	Activités immobilières
71	Location de machines et équipements
72	Activités informatiques
73	Recherche et développement
74	Autres services fournis aux entreprises
80	Enseignement
85	Santé et activités sociales
90	Assainissement, voirie et gestion des déchets
91	Activités associatives
92	Activités récréatives, culturelles et sportives
93	Services personnels

B Résultats TIC par classe NOGA niveau 2

Les tables 7 et 8, présentent les totaux TIC par classe NOGA au niveau 2. On note que les résultats à ce niveau détaillé sont liés à une grande variabilité. De plus, ils sont souvent basés sur un tout petit sous-échantillon (variable `nTIC`).

TAB. 7 Estimations détaillées dans les classes NOGA niveau 2 contenant des TIC, avec `nTot` le nombre total d'entreprises dans l'échantillon, `nTIC` le nombre d'entreprises TIC dans l'échantillon, les totaux (valeur en 10^9 CHF) et les CV (%) pour les trois variables dans le domaine des TIC. Première partie.

1997								
NOGA	nTot	nTIC	Tot017	cv017	Tot022	cv022	Tot053	cv053
30	13	10	1.4	31.3	1.0	31.7	0.3	11.5
31	99	7	1.0	22.2	0.6	52.0	0.3	8.0
32	51	51	5.1	6.0	3.2	6.6	1.5	2.9
33	179	46	4.3	18.4	2.4	25.7	1.5	7.3
51	749	41	3.8	97.5	1.2	75.1	1.7	45.7
64	14	7	1.0	44.1	0.5	65.2	0.3	36.5
72	48	48	7.8	6.5	2.8	10.5	3.7	8.5
Total	5626	210	24.3	15.8	11.7	10.9	9.3	9.0
1998								
NOGA	nTot	nTIC	Tot017	cv017	Tot022	cv022	Tot053	cv053
30	13	10	1.1	10.9	0.7	13.5	0.3	2.9
31	112	8	0.9	36.5	0.5	68.6	0.3	22.8
32	50	50	5.9	4.0	3.8	5.7	1.6	2.3
33	189	43	4.4	21.6	2.4	32.6	1.4	12.3
51	755	45	5.4	23.2	1.7	53.9	2.1	83.1
64	15	6	0.6	45.1	0.3	73.5	0.2	53.6
72	41	41	9.9	9.8	4.5	21.0	4.1	12.8
Total	5700	203	28.1	6.8	13.9	11.5	10.0	18.7
1999								
NOGA	nTot	nTIC	Tot017	cv017	Tot022	cv022	Tot053	cv053
30	12	9	0.5	35.0	0.3	44.9	0.1	5.9
31	105	10	1.2	28.1	0.7	48.8	0.4	10.0
32	67	67	5.8	7.2	3.6	9.7	1.6	2.6
33	203	51	4.2	18.8	2.3	31.9	1.4	6.3
51	790	45	4.8	72.5	1.4	102.1	2.2	48.4
64	20	11	2.6	45.1	1.9	54.9	0.6	23.3
72	62	62	11.4	7.9	5.3	14.9	5.0	5.4
Total	6116	255	30.5	12.7	15.5	13.7	11.3	10.0

TAB. 8 Estimations détaillées dans les classes NOGA niveau 2 contenant des TIC, avec nTot le nombre total d'entreprises dans l'échantillon, nTIC le nombre d'entreprises TIC dans l'échantillon, les totaux (valeur en 10^9 CHF) et les CV (%) pour les trois variables dans le domaine des TIC. Deuxième partie.

2000								
NOGA	nTot	nTIC	Tot017	cv017	Tot022	cv022	Tot053	cv053
30	15	10	0.7	18.8	0.5	25.9	0.2	16.1
31	109	9	0.9	64.6	0.6	112.1	0.2	41.2
32	67	67	6.0	4.0	3.8	5.0	1.7	1.8
33	213	57	5.3	12.6	3.0	19.4	1.7	5.9
51	792	42	3.9	54.5	1.3	40.7	1.8	36.0
64	26	14	12.6	7.7	6.0	14.5	2.7	4.6
72	60	60	11.2	4.9	4.4	9.9	5.8	6.1
Total	6150	259	40.6	6.3	19.5	7.3	14.2	5.4
2001								
NOGA	nTot	nTIC	Tot017	cv017	Tot022	cv022	Tot053	cv053
30	20	13	0.7	24.2	0.4	31.6	0.2	12.5
31	100	8	1.3	48.7	0.8	84.7	0.4	18.6
32	58	58	6.6	4.5	4.0	6.9	2.1	1.5
33	210	56	5.2	15.6	2.9	23.0	1.8	7.2
51	784	40	3.4	57.0	1.2	60.6	1.7	38.0
64	21	11	11.4	6.8	6.3	12.5	3.1	2.4
72	65	65	13.2	4.0	5.0	11.3	6.9	2.8
Total	6087	251	41.9	5.8	20.6	7.5	16.1	4.2
2002								
NOGA	nTot	nTIC	Tot017	cv017	Tot022	cv022	Tot053	cv053
30	43	33	0.8	11.6	0.5	15.1	0.3	7.1
31	156	13	1.2	24.8	0.8	36.5	0.4	6.7
32	108	108	6.8	2.4	4.3	2.9	2.2	2.3
33	168	48	6.1	18.4	3.5	31.1	2.2	8.0
51	455	33	3.3	59.9	1.2	76.1	1.6	43.0
64	61	42	18.8	5.5	10.8	8.0	3.0	1.9
71	39	1	0.0	0.0	0.0	0.0	0.0	0.0
72	219	219	14.8	5.0	6.9	8.2	7.3	2.4
Total	6685	497	51.9	5.0	27.9	6.3	17.0	4.4
2003								
NOGA	nTot	nTIC	Tot017	cv017	Tot022	cv022	Tot053	cv053
30	39	31	0.8	8.1	0.5	12.4	0.3	3.5
31	150	13	1.2	29.3	0.7	24.4	0.4	29.9
32	111	111	6.8	3.2	4.4	5.0	2.0	1.5
33	165	47	5.8	13.2	3.3	19.2	2.2	14.8
51	456	31	3.0	111.0	1.1	95.7	1.3	61.2
64	71	50	17.8	4.4	9.1	9.1	2.7	1.9
71	33	1	0.0	0.0	0.0	0.0	0.0	0.0
72	195	195	13.5	7.4	6.6	11.0	6.3	2.3
Total	6551	479	49.0	7.6	25.6	6.5	15.2	5.9

C Codes SAS

Cette section présente les macros construites pour le projet et, comme exemple, le programme utilisé pour l'application en 2003.

C.1 Macros

Macros de production WS (auteur : Markus Eichenberger, Office fédéral de l'informatique et de la télécommunication) :

1. `%estvar` : totaux estimés avec variance par `grou`
2. `%agg` : agrégat par NOGA au niveau 2 (output sous forme de PRINT).

Nouvelles macros (voir listing ci-dessous) :

1. `%outuigi` : détermination explicite des u_i, g_i et du poids final $wtot_i$ pour chaque unité de l'échantillon, avec appel de `%outui`.
2. `%EstVarDecDom` : estimation totaux avec variances pour `grou` pour domaine identifié par `dom=1`
3. `%aggEstNogaDom` : agrégat des résultats de `%EstVarDecDom` par NOGA au niveau 2.

```
* ListeMacrosTICAre_v1.sas (extrait)        ;
*                                           ;
* Macros pour les estimations TIC           ;
* à partir de la WS.                        ;
*                                           ;
* Note: appliqué pour les TIC des WS97-03   ;
*                                           ;
* Version V1                                ;
* Anne Renaud, janvier 2005                 ;
*********************************************;

/* [Are] pour modifications des macros de Markus */


/*****************************/
/* Restriction extraction ui*/
/*****************************/

%macro outui(data=,/*[Are] var1=,*/var2=,gewvar=,idkorr=,strata=,class=,robust=,c=,
             group=,uvar1=,uvar2=,uvar3=,robgew=,ident=,
             /*[Are] datpop=,clpop=,npop=,*/
             dataout=);

 /*      ------------------------------------------------------           */
 /*      Makro fuer robuste Varianzschaetzung eines Quotienten            */
 /*      ------------------------------------------------------           */
 /* Eingabe Paramter :                                                    */
 /*  data=    Name der Auswertungsdatei als SAS File                      */
 /*  var1=    1. Variable des Quotienten (Zaehler)                        */
 /*  var2=    2. Variable des Quotienten (Nenner)                         */
 /*  gewvar=  Gewichtungsvariable                                         */
 /*  idkorr=  Variable zur Identifikation stat. Einheiten, fuer die       */
 /*           das u Gewicht auf 1 bleibt (idkorr : 1)                     */
 /*  strata=  Schichtungs(-Stratifikations)-Variable(n)                   */
 /*  class=   Klassierungsvariable fuer Resultatsausgabe                  */
 /*  robust=  Robustifikation (1/0 - JA/NEIN)                             */
 /*  c=       Tuning Konstante fuer den univariaten one-step Schaetzer    */
 /*  group=   Klassierungsvariable fuer Schaetzer (im allg = class-Var)   */
 /*  uvar1=   Variable fuer den one-step-Algorithmus zur Berechnung der   */
 /*  uvar2=      u-Gewichte (bis zu max. 3 Variablen)                     */
 /*  uvar3=      u ist dann definiert durch das Minimum                   */
 /*  robgew=  1 oder 0 ob im one-step-Algorithmus der Median              */
 /*              mit/ohne Gewichtung berechnet wird                       */
 /*  ident=   Identifikationsvariable fuer das Outfile                    */
 /*  datpop=  SAS File fuer eine eventuelle Kalibrierung                  */
 /*  clpop=   Klassierungsvariable fuer Kalibrierung. Der Wertebereich    */
 /*             muss mit demjenigen der in class= angegebenen gleich sein */
 /*  npop=    Kalibrierungsvariable                                       */
 /*  Die Parameter : gewvar,idkorr,class,robust,c,group,uvar1-uvar3,      */
 /*                  ident,datpop,clpop,npop sind OPTIONAL                */
 /*                                                                       */
 /*  dataout: data avec ajout variables ui                          */
 /*-----------------------------------------------------------------------*/

   /* Initialisierungen */
%if &robust eq  %then %let robust=1;
%if &c eq  %then %let robust=0;
%if &robust = 0 %then %do;
    %let c=; %let uvar1=; %let uvar2=; %let uvar3=; %let group=;
%end;
%if /*[Are] &var1 eq or*/ &var2 eq %then %do;
    data _null_;
    file print;
    put #3 @2 "  ---->  Fehler : Keine Auswertungsvariablen";
    run;
   %goto ende;
%end;
%if &data eq  %then %do;
    data _null_;
    file print;
    put #3 @2 "  ---->  Fehler : Kein Auswertungsfile";
    run;
   %goto ende;
%end;
   /* Kalibrierungsfile (falls vorhanden)  */
/*[Are]%if &datpop ne %then %do;
 %let clren=;
 %if &clpop ne %then %do;
   %let clren=rename &clpop=&class;
 %end;
    data _pop_;
      set &datpop;
      _tot_=1;
      rename &npop=_npop_;
      &clren;
    run;
 %if &clpop ne %then %do;
    proc sort data=_pop_;
         by &class;
    run;
 %end;
%end;[Are]*/
   /*           Routine : one-step-Algorithmus                 */
   /*----------------------------------------------------------*/
/************************************************************/;
%macro coeff(datar=,var1=,var2=,w1=,w2=,by=,idkorr=,strata=,
             ind=,robgew=,ident=);
 %if &idkorr ne  %then %do;
  data tmpsub;
    set &datar;
    if  &idkorr ne 1;
  run;
 %end;
 %if &idkorr eq  %then %do;
  data tmpsub;
    set &datar;
  run;
 %end;
     /* Median (mit/ohne Gewichtung) */
  %if &robgew = 1  %then %let fr=freq &w2;
  %else %let fr=;
  proc univariate data=tmpsub noprint;
       var &var1 &var2;
       &fr;
       &by;
       output out=os1 pctlpts=50 pctlpre=ma mb;
  run;
     /* Quotientenschätzer */
  data estr;
   merge &datar os1;
   &by;
   t=ma50/mb50;
   res=abs(&var1-t*&var2);
   resb=res/(&var2**0.5);
  run;
 %if &idkorr ne  %then %do;
  data tmpsub;
    set estr;
```

```
    if  &idkorr ne 1;
  run;
 %end;
 %if &idkorr eq  %then %do;
  data tmpsub;
    set estr;
  run;
 %end;
  proc univariate data=tmpsub noprint;
       var resb;
       &fr;
       &by;
       output out=os2 pctlpts=50 pctlpre=sig;
  run;
      /* Berechnung der u-Gewichte mit tuning Konstante c */
  data estr;
   merge estr os2;
   &by;
   u&ind=1;
   sigro=sig50/0.67;
  %if &idkorr ne  %then %do;
   if &idkorr ne 1 and res  >  co*sigro*(&var2**0.5)  and res > 0   then do;
       u&ind=co*sigro*(&var2**0.5)/res;
      *u&ind=u&ind+(1-u&ind)/w;
   end;
  %end;
  %if &idkorr eq  %then %do;
   if  res  >  co*sigro*(&var2**0.5)  and res > 0   then do;
       u&ind=co*sigro*(&var2**0.5)/res;
      *u&ind=u&ind+(1-u&ind)/w;
   end;
  %end;
  run;
  data gewu&ind;
   set  estr;
   keep recn _tot_  u&ind  &group  &ident;
  run;
%mend;
/*************************************************************/;
         /* Einlesen und Aufbereitung der Auswertungdaten */
data tmp;
 set &data;
 %if &c eq %then %do;
   keep /*[Are] &var1*/ &var2 &gewvar &strata &class &group &idkorr
        _tot_  w wi   &uvar1 &uvar2 &uvar3 &ident noga2 burnr/*ajout noga2 et burnr [Are]*/;
 %end;
 %if &c ne %then %do;
   co=&c;
   keep /*[Are] &var1*/  &var2 &gewvar &strata &class &group &idkorr
        _tot_  w wi   &uvar1 &uvar2 &uvar3 &ident co noga2 burnr/*ajout noga2 et burnr[Are]*/;
 %end;
  _tot_=1;
 %if &gewvar ne  %then %do;
     %let flgw=1;
     %let gwa=Ja;
     w=&gewvar;
     wi=int(w*1000);
 %end;
 %if &gewvar eq  %then %do;
     %let flgw=0;
     %let gwa=Nein (w=1);
     %let gewvar=w;
     w=1;
     wi=1;
 %end;
 if /*[Are] &var1 ne .  and */ &var2 ne . and &gewvar ne . then do;
    output;
 end;
run;
data gewu;
  set tmp;
  keep _tot_  u;
  u=1;
run;
%let rba=NEIN;
%if &robust = 1 and &uvar1 ne  %then %do;
  %let rba=JA;
  %if &group eq  %then %let by=by _tot_;
  %if &group ne  %then  %do;
     %let by=by &group;
     proc sort data=tmp;
          &by;
     run;
     data tmp;
      set tmp;
      recn=_n_;
     run;
  %end;
       /* u1 - Gewichte */
 %coeff(datar=tmp,var1=&uvar1,var2=&var2,w1=&gewvar,w2=wi,by=&by,
        idkorr=&idkorr,strata=&strata,ind=1,robgew=&robgew,ident=&ident);
 data gewu;
   set gewu1;
   u=u1;
 run;
%end;
       /* u2 - Gewichte */
%if &robust = 1 and &uvar1 ne and &uvar2 ne   %then %do;
 %coeff(datar=tmp,var1=&uvar2,var2=&var2,w1=&gewvar,w2=wi,by=&by,
        idkorr=&idkorr,strata=&strata,ind=2,robgew=&robgew,ident=&ident);
 data gewu;
   merge gewu gewu2;
   by recn;
   u=min(u1,u2);
 run;
%end;
       /* u3 - Gewichte */
%if &robust = 1 and &uvar1 ne and &uvar2 ne and &uvar3 ne  %then %do;
 %coeff(datar=tmp,var1=&uvar3,var2=&var2,w1=&gewvar,w2=wi,by=&by,
        idkorr=&idkorr,strata=&strata,ind=3,robgew=&robgew,ident=&ident);
       /* u=min(u1,u2,u3) Gewicht */
 data gewu;
   merge gewu gewu3;
   by recn;
   u=min(u1,u2,u3);
 run;
%end;
%if &robust = 1  %then %do;
    data &dataout;/*[Are]data tmp;*/
     merge  tmp  gewu;
     by recn;
     mod=0;
     if u < 1  then mod=1;
     /*[Are] uw=u*w;*/
    run;
%end;
%if &robust = 0  %then %do;
    data &dataout;/*[Are]data tmp;*/
```

```
     set  tmp;
     u=1;
     /*[Are] uw=u*w;*/
    run;
%end;

* ordre;
proc datasets lib=work;
 delete os1 os2 gewu1-gewu3;
run;

%ende:
%mend outui;
*----------------------;
*----------------------;

*----------------------;
*----------------------;
* Macro appel outui, calcul gi et poids final;
*----------------------;
*----------------------;
%macro outuigi(dataIn=, var2=, datapop=, wtot=, dataOut=);

  %outui(data=&dataIn,
        /*var1=&var1,*/
        var2=&var2, gewvar=gew, idkorr=gros,
        strata=stra, class=grou,
        robust=1, c=10, group=grou,
        uvar1=T017, uvar2=T022, uvar3=T053, robgew=0,
        ident=, dataout=resM);

   data resM2;
    set resM;
    wux= gew*u*&var2;
   run;
   proc sort data=resM2; by grou;run;
   proc univariate data=resM2 noprint;
    var wux;
    by grou;
    output out=interM sum=Swux;
   run;
   data prepM;
    merge &datapop (rename=(X=Xg)) /*modif ws02.pro2ws-> datapop*/
       interM;
    by grou;
    gros=grou-10*floor(grou/10);
    if gros=0 then g=Xg/Swux; /*g-weight pour entreprises normales*/
    else if gros=1 then g=1; /*g-weight pour grandes entreprises*/
   run;
   proc sort data=resM2; by grou;run;
   proc sort data=prepM; by grou;run;
   data &dataOut;
    merge resM2 (drop=u1 u2 u3 wi _tot_ co mod wux in=a)
       prepM (keep=grou g);
    by grou;
    if a=1; /* restriction ech*/
    &wtot=gew*u*g;
   run;
%mend outuigi;
*----------------------;
*----------------------;

*----------------------;
*----------------------;
```

```
* macro construction estimation total ;
* et écart-type du total pour la pop  ;
* iddom=1 si dans domaine et 0 sinon  ;
*----------------------;
*----------------------;
%macro EstVarDecDom (dataIn=, varY=, iddom=,
                     class=, ncla=,
                     datapop=, varX=,
                     poidsW=, poidsU=, poidsG=,
                  stra=, dataOut=);

 * estim beta par classe (=grou);
 * x domaine                    ;
 *******************************;
 data interM;
  set &dataIn;
  * base pour b;
  bv1=&poidsW*&poidsU*&varY * &iddom;
  bv2=&poidsW*&poidsU*&varX  * &iddom;
  Ypond=&poidsW*&poidsU*&poidsG*&varY;
 run;
 proc sort data=interM; by &class; run;
 proc univariate data=interM noprint;
  var bv1 bv2;
  by &class;
  output out=interM2 sum=sbv1 sbv2;
 run;
 data listeB; set interM2;
  if sbv2 ne 0 then b=sbv1/sbv2;
  else b=.;
run;
 /*
 title 'calcul beta par grou x dom';
 proc print data=interM3;run; */

 * collage b sur ech, calcul résidus;
 ***********************************;
 data interM4;
  merge interM listeB;
  by &class;
  e=&varY-b*&varX;
  eb=&poidsW*&poidsU*e;
 run;
 /*
 title 'ech avec e et eb';
 proc print data=interM4;run; */

 *liste des classes (=grou);
 **************************;
 proc sort data=interM4;by &class;run;
 proc univariate data=interM4 noprint;
  var &iddom; by &class; output out=interL n=nb sum=nbdom;
 run;
 data liste; set interL; num=_N_;run;

 * boucle;
 ********;
 %do k=1 %to &ncla;

  * ident class (=grou), extrait données et extrait sbv2;
  data _null_; set liste; if num=&k;
   call symput('numero',&class);
   call symput('nb', nb);
   call symput('nbdom', nbdom);
```

```
 run;

 * prep jeu de données avec indicateurs (groupe et dom);
 data don&k;
  set interM4;
  indic=0;
  if &class=&numero then indic=1;
  YpondI=Ypond*indic*&iddom;
  ebI=eb*indic*&iddom;
 run;

 * total w*u*x dans class (= grou);
 data _null_; set listeB; if &class=&numero;
   call symput('sbv2',sbv2);
 run;

 *état des variables internes;
 * %put _user_;

 * estim var de eb par strate, fpc et varb dans strate;
 proc sort data=don&k; by &stra;run;
 proc univariate data=don&k noprint;
 var ebI &poidsW YpondI;
 by &stra;
 output out=tmp     n=   nh   i12  i13
                    var= dh2  i22  i23
                    sum= i31  i32  Ytoth
                    mean=i41  wmean i43;
 run;
 data tmp&k;
  set tmp (drop=i12 i13 i22 i23 i31 i32 i41 i43);
  fpch=(1- (1/wmean));
  if &sbv2 ne 0 then varbh=nh*fpch*dh2/(&sbv2**2);
  else varbh=.;
 run;

 * total des strates;
 proc univariate data=tmp&k noprint;
  var varbh Ytoth;
  output out=fin&k sum= varbtot Ytot;
 run;
 data fin&k;
  set fin&k;
  stdbtot=sqrt(varbtot);
  &class=&numero*1;
  nb=&nb;
  nbdom=&nbdom;
 run;

 * collage;
 %if &k=1 %then %do; data fin; set fin1;run; %end;
 %else %do; proc append base=fin data=fin&k;run;%end;

%end;

* reprise de Xref et calcul var total (par &class=grou et global);
************************************;
data out1;
 merge fin  &datapop (rename=(X=Xref grou=&class));
 by &class;
 stdtot=stdbtot*Xref;
 vartot=stdtot**2;
 if Ytot ne 0 then cvtot=stdtot/Ytot*100;
 else if Ytot=0 then cvtot=.;
run;
proc univariate data=out1 noprint;
 var nb nbdom vartot Ytot;
 output out=interM5 sum= nb nbdom vartot Ytot;
run;
data out2; set interM5;
  stdtot=sqrt(vartot);
  if Ytot ne 0 then cvtot=stdtot/Ytot*100;
  else if Ytot=0 then cvtot=.;
  grou=.;
run;
data &dataOut; set out1 out2;run;

 * ordre;
 proc datasets lib=work;
   delete out1 out2
          interM1-interM5
          don1-don&ncla
          tmp1-tmp&ncla
          fin1-fin&ncla;
 run;

%mend EstVarDecDom;
*----------------------;
*----------------------;


*----------------------;
*----------------------;
* Macro construction aggrégats des résultats;
* de la macro EstVarDecDom                   ;
*----------------------;
*----------------------;
%macro aggResNogaDom (dataIn=, class=, dataOut=);
 data interM;
  set &dataIn;
  noga2=floor(&class/10);
  keep &class noga2 Ytot nb nbdom vartot;
 run;
 proc sort data=interM; by noga2;run;
 proc univariate data=interM noprint;
  var nb nbdom Ytot vartot;
  by noga2;
  output out=interM2 sum=nb nbdom Ytot vartot;
 run;
 data &dataOut;
  set interM2;
  stdtot=sqrt(vartot);
  if Ytot>0 then cvtot=stdtot/Ytot*100;
  else cvtot=.;
 run;
 proc datasets lib=work;
   delete interM inter2;
 run;
%mend aggResNogaDom;
*----------------------;
*----------------------;
```

C.2 Exemple pour 2003

```
************************************;
* H:\SUKO\ProgSAS\lecBaseTICWS03.sas  ;
* (extrait)                          ;
*                                    ;
* Estimations WS03                   ;
* A. Renaud, mai 2005                ;
************************************;

* options;
options linesize=70 pagesize=50 pageno=1;

* librairies;
libname WS02 "D:\SUKO\DataPSA\WS_02";
libname WS03 "D:\SUKO\DataPSA\WS_03";
libname lib03 "D:\SUKO\libres03";

* macros;
%include "H:\SUKO\ProgSAS\Macros\ListeMacrosAre_2.sas";

*************************;
* Lecture données       ;
*************************;
proc contents data=WS03._all_;run;

************************;
* Définition échantillon;
************************;
data ech03;
 set WS03.ech1WS03;
 stra=noga2st*10+taille;
 * selection entreprises avec plus de 1 employés;
 if s14>1;
run;

* test poids 0 ou EPT<2;
proc means data=ech03; var gew; run;
data test;
 set WS03.ech1WS03;
 if gew=0 or s14<2;
run;
proc print;run;
data test; set ech03; if gew=0;run;

************************;
* Test valeurs de calage;
************************;
proc sort data=ech03; by grou;run;
proc univariate data=ech03 noprint;
 var BURNR;
 by grou;
 output out=inter n=n;
run;
data test;
 merge inter (in=a) WS03.PRO2WS03 (in=b);
 by grou;
 t=0;
 if a*b=0 then t=1;
run;
data inter; set test; if t=1;run;
proc print;run;

* liste pas besoin d'être modifiée;
data pro2ws03; set WS03.PRO2WS03; run;

**********************;
* Calculs wi, ui et gi ;
**********************;
%outuigi(dataIn=ech03, var2=s14, datapop=pro2ws03, wtot=wtot, dataOut=echWUG03);
proc contents data=echWUG03;run;
proc means data=echWUG03 n nmiss min max mean; var w g u wtot; run;

*********************************;
* Estimations pour éch complet    ;
*********************************;

* ----- macro;
%macro inter (dataIn=, var=, ncla=, datapop=, dom=);
 * global;
 %EstVarDecDom (dataIn=&dataIn, varY=&var, iddom=&dom, class=grou, ncla=&ncla,
        datapop=&datapop, varX=s14,
        poidsW=w, poidsU=u, poidsG=g,
        stra=stra, dataOut=res);
 %let nom=agg&var;
 %aggResNogaDom(dataIn=res, class=grou, dataOut=&nom);
 title '&var, domaine=&dom, strate=stra';
 proc print data=&nom (drop=vartot) noobs;run;
%mend inter;
*----------;
%macro appelinter (dataIn=, datapop=, ncla=, dom=, dataOut=);
 %inter(dataIn=&dataIn, var=T017, ncla=&ncla, datapop=&datapop, dom=&dom);
 %inter(dataIn=&dataIn, var=T022, ncla=&ncla, datapop=&datapop, dom=&dom);
 %inter(dataIn=&dataIn, var=T053, ncla=&ncla, datapop=&datapop, dom=&dom);
 data &dataOut;
  merge aggT017 (keep=noga2 nb Ytot cvtot
                 rename=(Ytot=Tot017 cvtot=cv017))
        aggT022 (keep=noga2  Ytot cvtot
                 rename=(Ytot=Tot022 cvtot=cv022))
        aggT053 (keep=noga2 Ytot cvtot
                 rename=(Ytot=Tot053 cvtot=cv053));
  by noga2;
 run;
 proc print data=&dataOut;run;
%mend appelinter;
 *----------;

* prep data;
data ech;
 set echWUG03;
 idall=1;
run;

* application (ncla=nombre d'éléments dans pro2ws);
%appelinter(dataIn=ech, datapop=pro2ws03, ncla=88,
            dom=idall, dataOut=FinWS03);
proc print data=FinWS03 noobs;run;

**********************;
* On garde des datasets;
**********************;
proc copy in=work out=lib03;
 select pro2ws03 echWUG03 FinWS03;
run;
proc contents data=lib03._all_;run;

*****************************;
* Définition TIC             ;
*****************************;
```

```
* modif burnr de ech pour lien avec sbur;
* voir programme pondération de PSA     ;
* annulation modif MON                  ;
****************************************;
data sbur03b; set WS03.sbur03;
  BURNR2=BURNR;
  /* 4 modifications de BURNR*/;
run;

* ajout s0007 (noga5 relevée) à l'échantillon avec WUG;
* et ajout ent_id pour lien avec reclassement         ;
******************************************************;
proc contents data=WS03.sbur03; run;
proc sort data=lib03.echWUG03 out=echWUG03; by burnr;run;
proc sort data=sbur03b; by burnr2;run;
proc sort data= WS02.Samplews2002_6fev2003 out=samp; by burnr;run;
data echwugNo03;
 merge echWUG03 (in=a )
       sbur03b (keep=burnr2 s0007 UNCDE rename=(burnr2=burnr))
       samp (keep=burnr entid);
 by burnr;
 if a=1;
 test=0;
 if entid=. or s0007='' then test=1;
run;
proc freq data=echwugNo03; tables test; run;

* identif TIC selon s0007;
*************************;
data pretic03;
 set echwugNo03 (drop=test);
 nogaTIC07=s0007;
 Idtic07=0;
 if s0007 in ('3002A', '3130A', '3210A', '3220A', '3230A',
              '3320A', '3330A',
              '5184A', '5184B', '5186A',
              '6420A', '6420B', '6420C',
              '7133A',
              '7210A', '7221A', '7222A', '7230A', '7240A', '7250A', '7260A')
  then Idtic07=1;
run;

* allure;
proc sort data=pretic03;by Idtic07;run;
proc freq data=pretic03; tables Idtic07;run;
proc freq data=pretic03;
 tables s0007*nogaTic07/list missing nopercent;
 by Idtic07;
run;

* ajout info sur reclassement;
* voir DefTicWS02.sas;
*********************;
proc contents data=WS02.recla5165A;run;
proc freq data=WS02.recla5165A;
 tables noga95*noga2002/list missing;
run;
proc sort data=pretic03; by entid;run;
proc sort data=WS02.recla5165A out=recla ; by ent_id;run;
data comp;
 merge pretic03 (in=a)
       recla (in=b
              keep=ent_id noga95 noga2002
              rename=(ent_id=entid noga95=reN95 noga2002=nogasep));
 by entid;
 if a=1;
 * identif classe selon s0007 (vide);
 test07=0;
 if s0007 in ('5186A', '5187A') then test07=1;
 * reclass;
 testre=0;
 if b=1 then testre=1;
run;

* sélection;
data test;
 set comp;
 if testre=1;
run;
proc freq data=test;
 tables s0007*reN95*nogasep/list missing;
run;
data test2;
 set test;
 if s0007 ne '5186A' and nogasep='5186A';
run;
proc print data=test2 (drop=nogaTic07 gros idtic07 test07 testre);run;

*****************************;
* Définition TIC selon choix ;
* explicité dans .out        ;
*****************************;

data echTIC03;
 set comp;
 * identification;
 Idtic=0;
 nogaTIC=s0007;
 if Idtic07=1 then Idtic=1;
 if s0007='5187A' and nogasep='5186A' then do;
       Idtic=1;
       nogaTIC='5186A';
    end;
 drop Idtic07 test07 nogaTIC07 testre reN95;
run;
proc contents data=echTic03;run;

* allure;
proc sort data=echTic03; by Idtic;run;
proc freq data=echTic03;
 tables Idtic*s0007*nogasep*nogatic/list missing nopercent;
 by Idtic;
run;

*******************;
* totaux (sans var);
*******************;

proc univariate data=echTic03 noprint;
 var T017 T022 T053;
 weight wtot;
 output out=inter1 n=n sum=ST017 ST022 ST053;
run;
proc sort data=echTic03;by idTIC;run;
proc univariate data=echTic03 noprint;
 var T017 T022 T053;
 by idTIC;
 weight wtot;
```

```
 output out=inter2 n=n sum=ST017 ST022 ST053;
run;
data inter03; set inter1 inter2;run;
proc print data=inter03 noobs;run;

* allure;
proc tabulate data=echTic03 formchar='     ' format=6. noseps missing;
 class nogatic idTIC;
 tables nogatic all, idTIC all /rts=10;
run;
proc freq data=echTic03; tables nogatic;by idTIC;run;

* allure ETP et nombres pond;
*****************************;
data inter; set echTic03; idall=1;run;
proc univariate data=inter noprint;
 var s14 idall;
 weight wtot;
 output out=inter1 n=n sum=s14p np;
run;
proc sort data=inter;by idtic; run;
proc univariate data=inter noprint;
 var s14 idall;
 by idtic;
 weight wtot;
 output out=inter2 n=n sum=s14p np;
run;
data inter03; set inter1 inter2;run;
proc print data=inter03;run;

************************;
* Allure stratification  ;
************************;
data dec03;
 set echTic03;
 noga2st=floor(stra/10);
 taille=stra-10*noga2st;
run;
proc tabulate data=dec03 formchar='  ' format=5. noseps missing;
 class noga2st taille idtic;
 tables noga2st all,  idtic * taille all /rts=10;
run;

* regroupements;
***************;
data decReg03;
 set dec03;
 straad=stra;
 if stra in (283) then straad=303;
 else if stra in (312,313) then straad=311;
 else if stra in (511) then straad=512;
 else if stra in (711) then straad=721;
 else if stra in (732) then straad=722;
 else if stra in (743) then straad=723;
run;

* test;
proc sort data=decReg03; by idtic;run;
proc freq data=decReg03;
 tables straad*stra /list missing;
 table straad /missing;
 by idtic;
run;
```

```
******************;
* On garde fichier;
******************;
data lib03.echTIC03; set decReg03;run;
proc contents data=lib03.echTIC03;run;

* allure;
proc freq data=lib03.echTIC03;
 tables s0007*nogaTIC*IdTic /list missing;
run;

************************;
* Estimations TIC        ;
************************;

* macro pour outputs;
********************;
* ----- macro;
%macro inter2 (dataIn=, var=, ncla=, datapop=, dom=, stra=);
 * global;
 %EstVarDecDom (dataIn=&dataIn, varY=&var, iddom=&dom, class=grou, ncla=&ncla,
        datapop=&datapop, varX=s14,
        poidsW=w, poidsU=u, poidsG=g,
        stra=&stra, dataOut=res);
 %let nom=agg&var;
 %aggResNogaDom(dataIn=res, class=grou, dataOut=&nom);
 title '&var, domaine=&dom, strate=&stra';
 proc print data=&nom (drop=vartot) noobs;run;
%mend inter2;
*----------;
%macro appelinter2 (dataIn=, datapop=, ncla=, dom=, stra=, dataOut=);
 %inter2(dataIn=&dataIn, var=T017, ncla=&ncla, datapop=&datapop, dom=&dom, stra=&stra);
 %inter2(dataIn=&dataIn, var=T022, ncla=&ncla, datapop=&datapop, dom=&dom, stra=&stra);
 %inter2(dataIn=&dataIn, var=T053, ncla=&ncla, datapop=&datapop, dom=&dom, stra=&stra);
 data &dataOut;
  merge aggT017 (keep=noga2 nb nbdom Ytot cvtot
                 rename=(Ytot=Tot017 cvtot=cv017))
        aggT022 (keep=noga2  Ytot cvtot
                 rename=(Ytot=Tot022 cvtot=cv022))
        aggT053 (keep=noga2 Ytot cvtot
                 rename=(Ytot=Tot053 cvtot=cv053));
  by noga2;
 run;
 proc print data=&dataOut;run;
%mend appelinter2;
 *----------;

*Application with straad (ncla=nombre d'éléments dans pro2ws) ;
****************;
%appelinter2 (dataIn=lib03.echTIC03, datapop=lib03.pro2ws03, ncla=88, dom=idtic,
              stra=straad, dataOut=FinWS03tic);
title 'résultats tic 2003 avec straad';
options linesize=75;
proc print data=FinWS03tic;run;

**********;
* on garde;
**********;
proc copy in=work out=lib03;
 select FinWS03tic;
run;
proc contents data=lib03.FinWS03tic;run;
```

Bibliographie

Cochran, W. G. (1977), *Sampling Techniques*. John Wiley & Sons.

Hulliger, B. (1999). Simple and robust estimators for sampling. In ASA : Proceedings of the section on Survey Research. American Statistical Association.

OFS (2003), Une introduction à la théorie et à la pratique des Comptes nationaux Méthodes et concepts du SEC, Office fédéral de la statistique, Neuchâtel.

Peters, R., J.-P. Renfer et B. Hulliger (2001), Statistique de la valeur ajoutée 1997-1998. Procédure d'exploitation des données, Rapport de méthodes, numéro de commande : 338-0005, Office fédéral de la statistique.

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

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). Bereinigung der Stichprobe, Ersatz fehlender Werte und Schätzverfahren. Erhebung über F+E in der schweizerischen Privatwirtschaft 2000. 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 é lèves (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.

Les données de la statistique de la valeur ajoutée (WS) sont utilisées pour la production d'estimations dans le domaine des technologies de l'information et de la communication (TIC). La méthode utilisée pour la WS est adaptée afin de produire des estimations du total de la valeur de production brute, du total de la consommation intermédiaire et du total de la charge de personnel dans les TIC. Il s'agit d'un domaine transversal aux résultats habituels de la WS. La nouvelle méthode est appliquée sur les données de la période 1997 à 2003 afin d'estimer les totaux des grandeurs économiques d'intérêt et leur variabilité.

N° de commande: 338-0049
Commandes: 032 713 60 60 Fax: 032 713 60 61 E-Mail: order@bfs.admin.ch
Prix: gratuit
ISBN 978-3-303-00401-2