Rapport de méthodes

Statistique suisse des bénéficiaires de l'aide sociale

0

Statistische Grundlagen und Übersichten
Bases statistiques et produits généraux
Basi statistiche e presentazioni generali

Pondération des communes 2004

Office fédéral de la statistique
Bundesamt für Statistik
Ufficio federale di statistica
Uffizi federal da statistica
Swiss Federal Statistical Office

OFS BFS UST

Neuchâtel, 2006

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 | **Statistique de la Suisse**

Methodenbericht | **Rapport de méthodes**

Statistique suisse des bénéficiaires de l'aide sociale

Pondération des communes 2004

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

Editeur **Office fédéral de la statistique**

OFS BFS UST

Office fédéral de la statistique
Bundesamt für Statistik
Ufficio federale di statistica
Uffizi federal da statistica
Swiss Federal Statistical Office

Neuchâtel, 2006

Préambule

La statistique suisse des bénéficiaires de l'aide sociale relève d'une collaboration entre la Confédération, les cantons et les communes. La section Sécurité sociale (SOZ) de l'Office fédéral de la statistique (OFS) a demandé au Service de méthodes statistiques (METH) de pondérer l'échantillon des communes ayant participé à l'enquête de 2004. Cet échantillon provient d'un plan et d'un tirage effectué en 1999 par METH. La pondération de 2004 a la particularité de devoir tenir compte de fusions de communes entre 1999 et 2004, en plus de la non-réponse observée en 2004.

Merci aux collaborateurs de la section SOZ et spécialement à Regula Häberli pour le suivi du projet et les nombreux travaux sur les listes de communes. Merci également à Yves Tillé, professeur au groupe de statistique de l'Université de Neuchâtel, pour les discussions sur le partage des poids et à Paul-André Salamin (METH) pour la relecture attentive du rapport.

Résumé

La statistique suisse des bénéficiaires de l'aide sociale fournit des résultats pour tous les cantons pour la première fois en 2004. Cette statistique est basée sur un échantillon de communes sélectionnées en 1999. L'échantillon est exhaustif dans certains cantons et non-exhaustif dans d'autres. Etant prévu pour durer plusieurs années, l'échantillon est touché par des morcellements et fusions de communes. Les données relevées en 2004 concernent donc des communes qui ne correspondent plus toujours exactement aux communes sélectionnées en 1999. Cette particularité est traitée à l'aide de la méthode généralisée du partage des poids. Les poids ainsi construits sont ensuite adaptés pour tenir compte de la non-réponse et d'un calage sur des tailles de populations connues. Des difficultés d'estimations sont détectées dans les cantons avec un faible taux de réponse, les cantons avec une grande modification de la liste des communes ou encore les cantons avec un bon taux de réponse global mais de la non-réponse dans quelques grandes communes.

Mots-clé

rapport de méthodes; pondération; aide sociale; partage des poids; non-réponse; post-stratification; calage.

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: 338-0039

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 2006
La reproduction est autorisée, sauf à des fins commerciales,
si la source est mentionnée.

ISBN: 3-303-00339-4

Table des matières

1 Introduction **5**

2 Plan d'échantillonnage **5**

3 Évolution de la liste des communes **7**

3.1 Morcellements de communes 8

3.2 Fusions de communes : généralités 9

3.3 Fusions de communes : estimation des poids inconnus 10

3.4 Fusions de communes : application 11

4 Un peu plus sur le partage des poids **14**

4.1 Formalisation 15

4.2 Généralisation à des liens pondérés 18

4.3 Choix de la variable de pondération 18

4.4 Post-stratification automatique sur la taille de la population ? 20

4.5 Modifications successives 21

4.6 Utilisation future du partage des poids sur les communes 23

5 Ajustement des poids **23**

5.1 Non-réponse 23

5.2 Calage 26

5.3 Un peu plus sur la pondération 28

6 Conclusion **31**

Références **31**

1 Introduction

La statistique suisse des bénéficiaires de l'aide sociale est organisée par la section de la Sécurité sociale (SOZ) de l'Office fédéral de la statistique (OFS). Les cantons, introduits par étapes dans la statistique, sont tous présents pour la première fois en 2004.

Les communes participant à l'enquête ont été sélectionnées en 1999. Dans la plupart des cantons, toutes les communes font partie de l'échantillon. Dans certains cantons, un échantillon de communes est sélectionné.

La liste des communes utilisée pour le tirage a subi des modifications au fil des ans. Les poids fixés au moment du plan d'échantillonnage ne sont donc plus valables pour les communes touchées par les modifications. De nouveaux poids doivent être déterminés afin de pouvoir extrapoler les résultats à la population d'intérêt de 2004. La méthode du partage des poids est appliquée à cet effet.

Certaines communes faisant partie de l'échantillon n'ont pas fourni les données pour 2004. Ces non-réponses sont traitées par le biais d'une adaptation des poids dans des cellules de traitement de non-réponse. De plus un calage est appliquée sur les communes ayant répondu à l'enquête afin d'ajuster les données sur des valeurs de référence connues (taille de population).

Le présent rapport présente la méthode de pondération choisie pour 2004 ainsi que les résultats d'alternatives envisagées puis laissées de côté. Il décrit également la méthode du partage des poids et quelques unes de ses particularités dans le cas particulier de notre application.

2 Plan d'échantillonnage

Le plan d'échantillonnage des communes de la statistique suisse des bénéficiaires de l'aide sociale est un plan stratifié selon le canton et trois classes de taille des communes ; voir détails dans Renaud (2001). Les communes sont tirées selon un algorithme aléatoire simple dans chaque strate.

L'échantillon des communes a été tiré en 1999 sur la base d'une liste de $N = 2903$ communes dans l'état de 1998. Le plan était construit sur la base d'informations sur la population des communes et de valeurs - relevées ou estimées - du nombre de bénéficiaires de l'aide sociale en 1993.

Les communes tirées en 1999 sont à la base de l'échantillon utilisé pour l'enquête de 2004. Cependant, 11 cantons ont décidé de faire un relevé exhaustif au lieu du relevé par échantillonnage prévu en 1999 ; voir Table 1. Par conséquent, l'échantillon adapté comporte 2272 communes au lieu des 1135 prévues initialement.

Les poids d'échantillonnage déterminés lors du plan d'échantillonnage sont adaptés pour tenir compte de l'adaptation de l'échantillon. Ainsi, le poids d'échantillonnage w_i de la commune $i = 1, .., N$ est définit par l'inverse de sa probabilité de sélection π_i. Le

poids est fixé à 0 si la commune n'est pas sélectionnée dans l'échantillon :

$$w_i = \begin{cases} 1 & \text{si } i \text{ est dans un canton exhaustif} \\ 1/\pi_i = N_h/n_h & \text{sinon, avec } i \text{ sélectionné dans strate } h \text{ de son canton} \\ 0 & \text{sinon (hors échantillon)} \end{cases} \quad (1)$$

avec N_h le nombre total de communes dans la strate h et n_h le nombre de communes tirées dans cette même strate.

TAB. 1: Liste des cantons avec le nombre total de commune (N), le plan d'origine (plan99, et taille n_99) et le plan adapté (plan99ad et taille n_99ad).

Kt	Canton	N	plan99	n_99	plan99ad	n_99ad
1	ZH	171	9 strates	87	idem	87
2	BE	400	3 strates	82	idem	82
3	LU	107	3 strates	44	*exhaustif*	107
4	UR	20	exhaustif	20	idem	20
5	SZ	30	exhaustif	30	idem	30
6	OW	7	exhaustif	7	idem	7
7	NW	11	exhaustif	11	idem	11
8	GL	29	exhaustif	29	idem	29
9	ZG	11	exhaustif	11	idem	11
10	FR	246	3 strates	66	*exhaustif*	246
11	SO	126	3 strates	44	idem	44
12	BS	3	exhaustif	3	idem	3
13	BL	86	3 strates	38	*exhaustif*	86
14	SH	34	exhaustif	34	idem	34
15	AR	20	exhaustif	20	idem	20
16	AI	6	exhaustif	6	idem	6
17	SG	90	3 strates	45	*exhaustif*	90
18	GR	213	3 strates	93	idem	93
19	AG	232	3 strates	71	*exhaustif*	232
20	TG	78	3 strates	51	idem	51
21	TI	245	3 strates	99	*exhaustif*	245
22	VD	385	3 strates	75	*exhaustif*	385
23	VS	163	3 strates	53	*exhaustif*	163
24	NE	62	3 strates	34	*exhaustif*	62
25	GE	45	3 strates	32	*exhaustif*	45
26	JU	83	3 strates	50	*exhaustif*	83
Total		2903		1135		2272

3 Évolution de la liste des communes

La liste des communes utilisée pour le tirage a subi des modifications au fil des ans. Certaines unités se sont partagées en sous-unités. D'autres unités ont fusionné pour créer des unités plus grandes. Ces changements ont pour conséquence que la liste des communes qui fournissent des données sur l'aide sociale en 2004 ne correspond plus à la liste des communes sélectionnées dans l'échantillon. Les poids fixés au moment du plan d'échantillonnage ne sont donc plus valables pour les communes touchées par les modifications.

Différents types de modifications ont eu lieu dans la liste des communes ; voir Figure 1. On appellera *morcellements* les cas de partage de communes en deux ou plus entités différentes (exemple : d partagée en $d1$ et $d2$) et *fusions* les cas de regroupements de deux ou plus communes en une entité (exemple : b et c groupées en b'). On notera que des morcellements et des fusions peuvent être combinés (exemple : la commune f est partagée en $f1$ et $f2$; cette dernière entité étant fusionnée avec la commune g pour donner g').

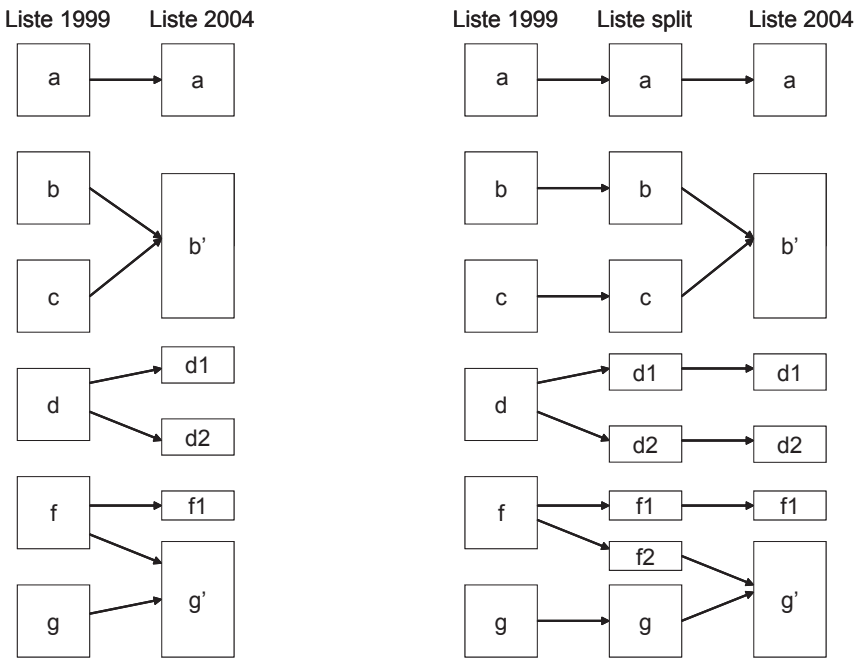


FIG. 1: Modification de la liste des communes. A gauche, modifications combinées. A droite, décomposition en morcellements (liste split) et fusions.

Les morcellements et fusions touchent parfois les mêmes communes. Ils peuvent donc difficilement être traités simultanément. Par conséquent, on traitera les deux cas l'un après l'autre, en commençant par les morcellements (liste split dans la Figure 1).

Une comparaison des listes de 1999 et de 2004, combinée avec une analyse des mo-

difications annoncées annuellement dans la liste des communes, a permis de décomposer les modifications en morcellements et fusions. Ce important travail a été effectué par la section SOZ. Il a permis de bien comprendre quelles entités ont effectivement subit un morcellement ou une fusion dans les années.

Le but du traitement des morcellements et des fusions est de créer des poids pour les communes de l'échantillon selon la liste de 2004. Ces poids pourront ainsi être utilisés pour extrapoler les résultats de l'enquête 2004 à la liste des communes de 2004.

3.1 Morcellements de communes

Des morcellements de communes sont observés uniquement dans le canton de Turgovie. Quatre communes sont partagées en 12 entités ; voir Table 2. Ces partages correspondent en fait à un passage d'une liste de "Munizipalgemeinde" à une liste de "Ortsgemeinde". Bien qu'ils ne correspondent pas à des partages réels de communes, ils peuvent être traités comme des morcellements.

La liste des $N = 2903$ communes du cadre de 1999 devient une liste de $N^{(s)} = 2911$ communes après morcellements (split).

TAB. 2: Morcellements de communes dans le canton de Turgovie, avec `Gnr`, `Gname` et la taille de la commune `Pop97` dans le cadre de 99, `GnrS` et `GnameS` les communes après morcellement, la taille `Pop90S` et la répartition `Prop90` selon le recensement 90, et la taille estimée `Pop97est` pour le cadre de référence de 99.

Gnr	Gname	Pop97	GnrS	GnameS	Pop90S	Prop90	Pop97est
4540	Basadingen	3039	4540	Basadingen	821	31.9%	969.3
			4533	Schlattingen	512	19.9%	604.5
			4532	Mett-Oberschlatt	326	12.7%	384.9
			4534	Unterschlatt	915	35.5%	1080.3
4773	Tobel	1951	4773	Tobel	739	45.1%	879.7
			4772	Tägerschen	389	23.7%	463.1
			4723	Braunau	511	31.2%	608.3
4801	Berlingen	1031	4801	Berlingen	784	76.0%	784.0
			4836	Dettighausen	247	24.0%	247.0
4870	Steckborn	4131	4864	Steckborn	3381	84.9%	3507.5
			4863	Salen-Reutenen	134	3.4%	139.0
			4861	Göndelhart-Hörhausen	467	11.7%	484.5

La population `Pop97` de 1997 est une information du cadre des communes de 1999. Cette variable sera utilisée pour le traitement des fusions des communes. Ainsi, nous estimons une population `Pop97est` pour les 12 entités désagrégées. Cette estimation est basée sur la répartition de la population selon le recensement de 90 (variable : `Pop90S`). Nous faisons donc l'hypothèse que les tailles relatives de 1990 sont valables pour 1997 pour les communes concernées.

La probabilité de sélection π_i de l'entité i provenant du partage de la commune j est identique à la probabilité de sélection π_j de la commune j. On a donc les poids après morcellement $w_i^{(s)}$, $i = 1, .., N^{(s)}$:

$$w_i^{(s)} = \begin{cases} w_i & \text{si } i \text{ n'est pas touché par un morcellement} \\ w_j & \text{si } i \text{ provient du morcellement de la commune } j \end{cases} \quad (2)$$

avec les poids w_i, respectivement w_j, de l'équation (1).

3.2 Fusions de communes : généralités

Différents cas de fusions sont observés, voir Figure 2 :

1. fusion de communes dans l'échantillon,
2. fusion de communes en partie dans et en partie hors de l'échantillon, et
3. fusion de communes hors de l'échantillon.

Dans tous les cas, la fusion peut toucher deux communes ou plus.

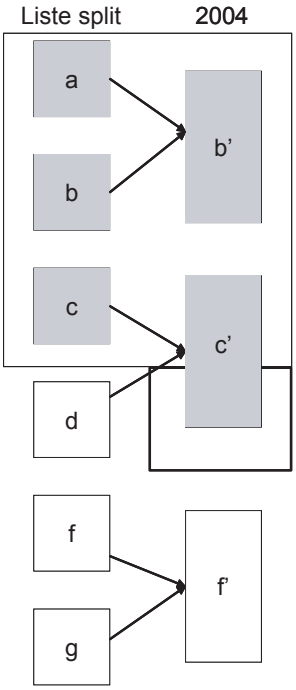


FIG. 2: Modification de la liste des communes par les fusions. Les communes a, b et c font partie de l'échantillon tiré (avec intégration des morcellements, "Liste split"). Les données sont récoltées en 2004 pour les communes b' et c'.

La particularité des fusions est que les données ne sont plus relevées sur les unités de base mais seul les agrégats, dans l'état fusionné, sont disponibles. Ainsi, dans

l'exemple de la Figure 2, le nombre de bénéficiaires de l'aide sociale est relevé pour les communes b' et c', sans avoir d'information sur la distribution de ce nombre dans les entités de base a, b et c. On note également que les bénéficiaires de la commune d ont été intégrés dans la fusion c' alors qu'il n'était pas prévu de relever ses informations.

3.3 Fusions de communes : estimation des poids inconnus

Supposons que les communes a et b, avec les poids d'échantillonnage connus w_a et w_b, fusionnent en une commune c. On relève le nombre de bénéficiaires y_c de c lors de l'enquête. Nous sommes donc dans la situation où les nombres de bénéficiaires y_a et y_b des communes a et b sont inconnus mais où la somme $y_a + y_b = y_c$ est connue. L'apport $\widehat{Y}_c$ de la commune c au total estimé des bénéficiaires du canton ne peut par conséquent pas être estimé par $\widehat{Y}_c = w_a y_a + w_b y_b$.

Supposons maintenant qu'il existe une variable auxiliaire x connue pour a, b et c, bien corrélée à la variable y, et satisfaisant l'égalité $x_c = x_a + x_b$. On peut alors estimer le nombre de bénéficiaires $\widehat{y}_a$ et $\widehat{y}_b$ dans les communes a et b comme suit :

$$\widehat{y}_a = y_c \frac{x_a}{x_c}, \tag{3}$$

$$\widehat{y}_b = y_c \frac{x_b}{x_c}, \tag{4}$$

avec $\widehat{y}_a + \widehat{y}_b = y_c$.

L'estimateur $\widehat{Y}_c$ devient alors :

$$\widehat{Y}_c = w_a \widehat{y}_a + w_b \widehat{y}_b = \left(\frac{w_a x_a + w_b x_b}{x_c} \right) y_c = \widehat{w}_c \, y_c \tag{5}$$

avec le poids $\widehat{w}_c$ estimé pour la nouvelle commune c :

$$\widehat{w}_c = \frac{w_a x_a + w_b x_b}{x_c}. \tag{6}$$

L'hypothèses sur le lien entre les variables y et x permet donc d'estimer un poids $\widehat{w}_c$ pour la commune fusionnée c. L'apport $\widehat{Y}_c$ de la commune c au total estimé des bénéficiaires du canton devient donc $\widehat{Y}_c = \widehat{w}_c \, y_c$.

L'estimateur du poids peut être généralisé à la commune j formée par la fusion de F communes de poids w_i, $i = 1, .., F$:

$$\widehat{w}_j = \frac{\sum_{i=1}^{F} w_i x_i}{\sum_{i=1}^{F} x_i}. \tag{7}$$

Le poids estimé $\widehat{w}_j$ correspond à une moyenne des poids w_i des communes fusionnées, pondérée par la variable x.

Si les communes fusionnées ont toutes le même poids w (même strate par exemple), l'estimateur $\widehat{w}_c$ est égal à ce poids w. La variable x n'intervient pas. Si une commune i

hors échantillon fait partie de la fusion, son poids est $w_i = 0$. Ainsi, elle intervient dans le calcul de $\widehat{w}_j$ uniquement dans la somme $\sum_i x_i$. De même dans le cas extrême ou la fusion ne touche que des communes hors échantillon $w_i = 0$, $i = 1, .., F$, le poids de la nouvelle commune reste égal à 0.

La détermination des poids $\widehat{w}_j$ peut être assimilée à l'application de la méthode généralisée du partage des poids (Lavallée, 2002). Cette technique garantit des estimations sans biais dans le cas des communes quelque soit la variable x ; voir Section 4.

La variable auxiliaire x peut-être choisie de différentes manières. Dans notre cas, nous choisissons la population résidante de la commune car cette variable est disponible pour toutes les communes et elle est bien corrélée avec le nombre de bénéficiaires ; voir plus d'information dans Section 3.4. D'autres possibilités sont décrites dans le cadre plus général de la Section 4.

3.4 Fusions de communes : application

De nombreuses fusions de communes ont eu lieu entre 1999 et 2004. Les fusions ont lieu à l'intérieur des cantons. Il n'y a pas de fusions à cheval sur deux cantons.

On considère différents types de cantons selon l'existence d'échantillon ou pas en plus des fusions ; voir Table 3 :

1. canton sans fusion (`TraitF`='-') : la liste des communes - après morcellements s'il y a lieu - reste fixe et les communes gardent leur poids $w_i^{(s)}$;
2. canton avec fusion et relevé exhaustif (`TraitF`='reprise') : la liste des communes de 2004 est prise comme liste de référence. Les poids sont fixés à 1.
3. canton avec fusion et échantillonnage (`TraitF`='special') : la liste des communes est adaptée et on estime les poids des nouvelles entités selon la méthode du partage des poids.

Les tailles du cadre indiquées dans la colonne N99s de la Table 3 sont identiques à la colonne N de la Table 1, excepté pour le canton de Turgovie qui a eu des morcellements.

La liste de 2004 reçue de la section SOZ est prise comme référence de 2004 dans les cas de `TraitF`='reprise'. Les différences avec la liste de 1999 sont considérées comme résultant de fusions.

Les non-réponses de communes seront traitées dans la Section 5.1. Elles sont indiquées dans la Table 3 car une attention soutenue est nécessaire dans l'identification des cas de non-réponse dans le cas de fusions. Il est en effet important de savoir quelle est exactement l'entité non-répondante ou répondante.

TAB. 3: Cantons avec taille N99s dans la liste adaptée pour les morcellements, existence de fusions `Fusion` et nouvelle taille Nfus, type d'échantillonnage plan99ad et existence `Non-rép` de non-réponse en 2004. La colonne `TraitF` indique s'il faut reprendre la liste de 2004 (reprise) ou faire un traitement spécial (special).

Kt	Canton	N99s	Fusion	Nfus	plan99ad	Non-rép	TraitF
1	ZH	171	-	171	*éch*	*oui*	-
2	BE	400	*oui*	398	*éch*	*oui*	*special*
3	LU	107	*oui*	103	exhaustif	-	reprise
4	UR	20	-	20	exhaustif	-	-
5	SZ	30	-	30	exhaustif	-	-
6	OW	7	-	7	exhaustif	-	-
7	NW	11	-	11	exhaustif	-	-
8	GL	29	*oui*	27	exhaustif	-	reprise
9	ZG	11	-	11	exhaustif	-	-
10	FR	246	*oui*	177	exhaustif	-	reprise
11	SO	126	-	126	*éch*	*oui*	-
12	BS	3	-	3	exhaustif	-	-
13	BL	86	-	86	exhaustif	*oui*	-
14	SH	34	*oui*	32	exhaustif	*oui*	reprise
15	AR	20	-	20	exhaustif	-	-
16	AI	6	-	6	exhaustif	-	-
17	SG	90	-	90	exhaustif	*oui*	-
18	GR	213	*oui*	208	*éch*	-	*special*
19	AG	232	*oui*	231	exhaustif	*oui*	reprise
20	TG	86	*oui*	80	*éch*	*oui*	*special*
21	TI	245	*oui*	204	exhaustif	-	reprise
22	VD	385	*oui*	381	exhaustif	-	reprise
23	VS	163	*oui*	152	exhaustif	-	reprise
24	NE	62	-	62	exhaustif	-	-
25	GE	45	-	45	exhaustif	-	-
26	JU	83	-	83	exhaustif	-	-
Total		2911		2764			

Les cantons de Berne, Grisons et Turgovie sont dans la situation de fusions combinées avec un échantillonnage (`TraitF`='special').

Les 2 fusions du canton de Berne ont lieu hors de l'échantillon ; voir Table 4. Il n'est donc pas nécessaire de faire des adaptations de poids pour les 82 communes de l'échantillon bernois.

TAB. 4: Communes touchées par des fusions dans les 3 cantons avec traitement spécial (BE, GR, TG), avec l'année de fusion, le nom et le numéro des communes en 1999, la présence ou non dans l'échantillon de 1999 (*Ech*), le poids d'échantillonnage $w_i^{(s)}$ après morcellement, la population relevée ou estimée pour 1997 (`Pop97`), puis le nom, le numéro et le poids estimé $\widehat{w}_i^{(f)}$ des communes de 2004.

Canton	année	Commune 1999	Gnr1999	*Ech*	$w_i^{(s)}$	`Pop97`	Commune 2004	Gnr2004	$w_i^{(f)}$
BE	2004	Englisberg	864	0	0	-	Wald	888	0
BE	2004	Zimmerwald	887	0	0	-			
BE	2004	Niederwichtrach	618	0	0	-	Wichtrach	632	0
BE	2004	Oberwichtrach	621	0	0	-			
GR	1998	Rona	3537	1	1.907	57	Tinizong-Rona	3541	0.299
GR	1998	Tinizong	3541	0	0	306			
GR	2002	Camuns	3591	0	0	55	Suraua	3599	1.544
GR	2002	Surcasti	3599	1	1.907	70			
GR	2002	Tersnaus	3601	1	1.907	72			
GR	2002	Uors-Peiden	3602	1	1.907	92			
GR	2004	Donath	3705	1	1.907	145	Donath	3705	1.907
GR	2004	Patzen-Fardün	3709	1	1.907	66			
TG	1999	Schlattingen	4533	1	1	604.5	Basadingen-Schlattingen	4536	1
TG	1999	Basadingen	4540	1	1	969.3			
TG	1999	Mett-Oberschlatt	4532	1	1	384.9	Schlatt	4546	1
TG	1999	Unterschlatt	4534	1	1	1080.3			
TG	1999	Tägerschen	4772	0	0	463.1	Tobel-Tägerschen	4776	0
TG	1999	Tobel	4773	0	0	879.7			
TG	1999	Homburg	4816	0	0	818	Homburg	4816	0.433
TG	1999	Göndelhart-Hörhausen	4861	1	1	484.5			
TG	1999	Salen-Reutenen	4863	1	1	139.0			
TG	1998	Dettighofen	4836	1	1.875	247	Pfyn	4841	0.254
TG	1998	Pfyn	4841	0	0	1578			

Par contre, les 3 fusions du canton des Grisons ont lieu entre communes tout ou partie dans l'échantillon. De même, 4 des 5 fusions du canton de Turgovie ont lieu entre communes tout ou partie dans l'échantillon. Les poids estimés $\widehat{w}_i^{(f)}$ de ces nouvelles communes sont déterminés selon le partage des poids ; voir équation (7).

La variable auxiliaire x est la population `Pop97` de 1997 ; estimée dans le cas des communes de Turgovie touchées par des morcellements. Elle est choisie pour des raisons pratiques et de disponibilité des données. Cette variable, plus très actuelle, pourrait être remplacée par la population durant l'année de la fusion. On note cependant que seules les valeurs relatives sont importantes dans le calcul de $\widehat{w}_i^{(f)}$. On fait donc l'hypothèse que les tailles relatives des communes avant fusion sont constantes entre 1997 et le moment de la fusion.

Nous notons que, dans deux cas de Turgovie, l'effet du partage des poids est indépendant de la variable x. Il s'agit des fusions de deux communes de poids égaux à 1.

Seules 5 communes - relativement petites - de la liste de 2004 ont un poids estimés $\widehat{w}_i^{(f)}$ qui dépend de la variable auxiliaire x (3 aux Grisons et 2 en Turgovie). Les autres communes gardent leur poids d'échantillonnage. L'influence du choix de x sur les estimations cantonales est donc faible dans notre cas.

Exemple de calcul : fusion de l'ensemble des communes $\Lambda =$ {Camuns, Surcasti, Tersnaus, Uors-Peiden} pour former la commune de Suraua aux Grisons. Les trois communes de Surcasti, Tersnaus et Uors-Peiden font partie de l'échantillon de 1999, avec un poids d'échantillonnage égal à 1.907. La commune de Camuns ne fait pas partie de l'échantillon de 1999 et a donc un poids d'échantillonnage égal à 0. Les 4 communes ont entre 55 et 92 habitants (variable x) ; voir Table 4.

Le poids $\widehat{w}_i^{(f)}$ de la commune $i =$ Suraua est estimé comme suit :

$$\widehat{w}_i^{(f)} = \frac{\sum_{j\in\Lambda} w_j^{(s)} x_i}{\sum_{j\in\Lambda} x_j} = \frac{0\times 55 + 1.907\times(70+72+92)}{55+70+72+92} = 1.544.$$

Il est inférieur au poids d'échantillonnage des trois communes faisant déjà partie de l'échantillon car une quatrième communes est maintenant prise en compte dans le relevé.

En résumé, les poids $w_i^{(f)}$ des communes de la liste de 2004 - avec morcellements et fusions - sont définis par :

$$w_i^{(f)} = \begin{cases} w_i^{(s)} & \text{si } i \text{ dans canton sans fusion ou avec fusions hors de l'échantillon} \\ \widehat{w}_i^{(f)} & \text{si } i \text{ dans canton avec échant et } i \text{ est touchée par une fusion} \\ 0 & \text{sinon (hors échantillon)} \end{cases} \tag{8}$$

On a $w_i^{(s)}$ le poids d'échantillonnage adapté pour le changement de plan et les morcellements, voir Section 3.1, et

$$\widehat{w}_i^{(f)} = \frac{\sum_{j\in\Lambda} w_j^{(s)} x_j}{\sum_{j\in\Lambda} x_j} \tag{9}$$

pour la commune i correspondant à la fusion des communes de l'ensemble Λ et x la population mesurée ou estimée pour 1997.

4 Un peu plus sur le partage des poids

La méthode du partage des poids trouve son application première dans les problèmes de sondages indirects (Lavallée, 2002).

Dans le sondage indirect, on a deux populations U^A et U^B, reliées entre elles. On tire dans U^A mais on désire produire des estimations pour U^B. Par exemple, on tire des parents pour estimer une statistique d'intérêt pour des enfants. L'estimation d'une statistique d'intérêt θ d'une population cible U^B en se servant d'un échantillon tiré dans U^A peut être très complexe, surtout si les liens entre les unités des deux populations ne

sont pas bijectifs. La difficulté se trouve principalement dans le choix de la probabilité d'inclusion des unités de U^B atteintes durant le relevé.

La méthode généralisée du partage de poids (MGPP) permet d'associer un poids pour chaque unité enquêtée dans U^B. Ce poids correspond à une certaine moyenne des poids de sondage des unités de U^A.

Dans notre cas de la fusion des communes, la liste des communes utilisées pour le tirage en 1999 forme U^A et la liste des communes au moment de l'enquête forme U^B. Le poids d'une commune après fusion dépendra donc des poids d'échantillonnage des communes avant fusion.

La méthode du partage des poids est résumée ci-dessous dans le cas général et en suivant les notations de Lavallée (2002).

4.1 Formalisation

Soient U^A la population des unités $j = 1, .., M^A$ et $s^A \subset U^A$ l'échantillon de taille m^A. Soit également U^B la population des grappes $i = 1, .., N^B$ contenant chacune $k = 1, .., M_i^B$ unités ; voir Figure 3.

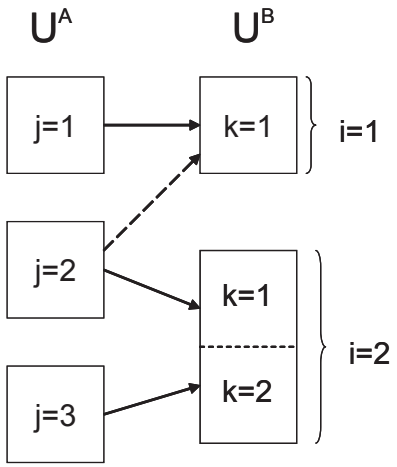


FIG. 3: Illustration des notations dans la théorie du partage des poids. La flèche en traitillé est possible dans le cas général mais pas dans le cas particulier de la fusion de communes.

Pour utiliser la MGPP et assurer son aspect sans biais, chaque grappe i de U^B doit posséder au moins un lien avec une unité j de U^A. Cette contrainte est satisfaite dans notre application car toute commune de la liste de 2004 existe déjà dans la liste de 1999 ou provient d'une fusion de communes existantes en 1999.

Dans la description de la MGPP, Lavallée fait l'hypothèse que l'on enquête auprès de chaque unité k des grappes i de U^B. De plus, les liens sont connus au niveau des unités k. Dans notre cas, le relevé se fait au niveau de la grappe i. Nous n'avons

pas le détail des parties de la commune, notamment pas les résultats au niveau des communes avant fusion.

Supposons que l'on désire estimer le total $Y^B = \sum_{i \in U^B} y_i$ de la variable d'intérêt y dans la population U^B. L'estimateur usuel a la forme :

$$\widehat{Y}^B = \sum_{i \in s^B} w_i y_i, \tag{10}$$

avec s^B l'échantillon de U^B construit à partir de l'échantillon s^A tiré dans U^A, et w_i un certain poids d'extrapolation à déterminer.

4.1.1 Les étapes de la MGPP

1. Pour chaque k de i, on calcule le poids initial w'_{ik} :

$$w'_{ik} = \sum_{j=1}^{M^A} I_{j,ik} \frac{t_j}{\pi_j^A}, \tag{11}$$

avec $I_{j,ik} = 1$ si l'unité j est liée à l'unité k de i et 0 sinon, $t_j = 1$ si $j \in s^A$ et 0 sinon, et π_j^A la probabilité d'inclusion de j dans s^A.

2. Pour chaque k de i, on calcule le nombre total de liens L_{ik}^B avec les unités de U^A :

$$L_{ik}^B = \sum_{j=1}^{M^A} I_{j,ik}. \tag{12}$$

3. On calcule le poids final w_i :

$$w_i = \frac{\sum_{k=1}^{M_i^B} w'_{ik}}{\sum_{k=1}^{M_i^B} L_{ik}^B}. \tag{13}$$

4. Le poids final w_{ik} est assigné aux parties k de i :

$$w_{ik} = w_i. \tag{14}$$

4.1.2 Adaptation à notre cas

Seul le poids final w_i est intéressant dans notre cas. Nous n'avons pas besoin des poids w_{ik}.

L'équation (11) peut être reécrite sous la forme :

$$w'_{ik} = \sum_{j \in U^A} I_{j,ik} \cdot w_j^A, \tag{15}$$

avec le poids d'échantillonnage w_j^A de j définit par :

$$w_j^A = \begin{cases} 1/\pi_j^A & \text{si } j \text{ est dans } s^A \\ 0 & \text{sinon} \end{cases} \tag{16}$$

De son côté, L_{ik}^B est donné par :

$$L_{ik}^B = \sum_{j \in U^A} I_{j,ik}. \tag{17}$$

La somme sur les k des w'_{ik} devient alors :

$$\sum_{k \in i} w'_{ik} = \sum_{j \in U^A} w_j^A \sum_{k \in i} I_{j,ik} = \sum_{j \in U^A} w_j^A \cdot L_{ji}, \tag{18}$$

avec L_{ji} le nombre de liens entre j et i. Notons que L_{ji} vaut 0 ou 1 dans notre application.

La somme sur les k des L_{ik}^B est L_i le nombre total de lien entre les unités de U^A et i :

$$\sum_{k \in i} L_{ik}^B = \sum_{j \in U^A} \sum_{k \in i} I_{j,ik} = \sum_{j \in U^A} L_{ji} = L_i. \tag{19}$$

Le poids w_i est donné par l'équation (13), ou encore :

$$w_i = \frac{\sum_{k \in i} w'_{ik}}{\sum_{k \in i} L_{ik}^B} = \frac{\sum_{j \in U^A} w_j^A \cdot L_{ji}}{\sum_{j \in U^A} L_{ji}} = \frac{\sum_{j \in U^A} w_j^A \cdot L_{ji}}{L_i}. \tag{20}$$

Le poids w_i a donc la forme d'une moyenne des w_j^A pondérée par le nombre de liens L_{ji}.

En résumé : Le poids w_i est donné par :

$$w_i = \frac{\sum_{j \in U^A} w_j^A \cdot L_{ji}}{\sum_{j \in U^A} L_{ji}},$$

avec

$$w_j^A = \begin{cases} 1/\pi_j^A & \text{si } j \text{ est dans } s^A \\ 0 & \text{sinon} \end{cases}$$

et L_{ji} le nombre de liens entre j et i.

4.2 Généralisation à des liens pondérés

La méthode du partage des poids peut être étendue à des liens pondérés. Ainsi, la variable indicatrice $I_{j,ik}$ est remplacée par une variable quantitative $\theta_{j,ik}$ représentant l'importance que l'on veut donner au lien $I_{j,ik}$.

Notons que, pour que la MGPP reste sans biais, il faut que la contrainte suivante soit respectée pour tout $i \in U^B$:

$$\theta_i = \sum_{j \in U^A} \sum_{k \in i} \theta_{j,ik} \neq 0. \tag{21}$$

Cette contrainte correspond à la contrainte dans le cas simple : chaque élément i est lié à au moins un élément de U^A.

Les étapes de la MGPP deviennent donc :

1. Poids initial de k dans i :

$$w'^{\alpha}_{ik} = \sum_{j \in U^A} \alpha_{j,ik} w^A_j, \tag{22}$$

avec

$$\alpha_{j,ik} = \frac{\theta_{j,ik}}{\sum_{j \in U^A} \sum_{k \in i} \theta_{j,ik}} = \frac{\theta_{j,ik}}{\theta_i}. \tag{23}$$

2. Poids final w^{α}_i :

$$w^{\alpha}_i = \sum_{k \in i} w'^{\alpha}_{ik} = \sum_{j \in U^A} w^A_j \sum_{k \in i} \alpha_{j,ik} = \sum_{j \in U^A} w^A_j \frac{\theta_{ji}}{\theta_i} = \frac{\sum_{j \in U^A} w^A_j \cdot \theta_{ji}}{\sum_{j \in U^A} \theta_{ji}}, \tag{24}$$

avec $\theta_{ji} = \sum_{k \in i} \theta_{j,ik}$.

On aboutit donc à nouveau sur une moyenne des poids w^A_j pondérée par une certaine valeur θ_{ji}. Il s'agit d'une extension de la moyenne pondérée par la variable binaire correspondant au lien L_{ji}.

4.3 Choix de la variable de pondération

La variable θ_{ji} peut-être choisie de différentes manières. Dans notre cas de liens entre unités j et i sans détails au niveau des k, on peut envisager par exemple trois types de poids :

1. MGPP non pondérée :

$$\theta^{(1)}_{ji} = I_{ij}. \tag{25}$$

2. MGPP pondérée par la probabilité d'inclusion π_j de j dans l'échantillon s^A :

$$\theta^{(2)}_{ji} = \pi_j \cdot I_{ij}. \tag{26}$$

3. MGPP pondérée par une variable auxiliaire x_j telle une certaine mesure de taille :

$$\theta^{(3)}_{ji} = x_j \cdot I_{ij}. \tag{27}$$

avec I_{ij}= 1 s'il y a un lien entre j et i, et 0 sinon.

Le choix de la variable utilisée dans la MGPP pondérée peut avoir une grande influence sur les poids estimés et donc également sur les résultats des estimations.

Dans l'exemple de la Table 5, les 5 communes de U^A fusionnent en deux communes. Les différents liens sont déterminés pour toutes les combinaisons de communes de U^A et de U^B (deuxième table). On observe alors dans la troisième table que les poids $w_i^{(1)}$, $w_i^{(2)}$ et $w_i^{(3)}$ correspondant respectivement aux pondérations $\theta_{ji}^{(1)}$, $\theta_{ji}^{(2)}$ et $\theta_{ji}^{(3)}$, sont sensiblement différents.

TAB. 5: Exemple de fusions de communes et de pondérations pour la MGPP.

Unité j	éch s^A	w_j	π_j	x_j	Unités i	éch s^B
1	oui	2	0.5	500	1	oui
2	oui	5	0.2	100	1	
3	oui	10	0.1	10	2	oui
4	oui	5	0.2	100	2	
5	non	0	0.5	1000	2	

Unité j	w_j	Unités i	$\theta_{ji}^{(1)}$	$\theta_{ji}^{(2)}$	$\theta_{ji}^{(3)}$
1	2	1	1	0.5	500
2	5	1	1	0.2	100
3	10	1	0	0	0
4	5	1	0	0	0
5	0	1	0	0	0
1	2	2	0	0	0
2	5	2	0	0	0
3	10	2	1	0.1	10
4	5	2	1	0.2	100
5	0	2	1	0.5	1000

Unité i	$w_i^{(1)}$	$w_i^{(2)}$	$w_i^{(3)}$
1	(2+5)/2=3.5	(1+1)/0.7=2.9	(1000+500)/600=2.5
2	(10+5)/3=5.0	(1+1)/0.8=2.5	(100+500)/1110=0.5

La différence entre les poids $w_i^{(1)}$, $w_i^{(2)}$ et $w_i^{(3)}$ est spécialement grande dans la deuxième fusion. Si x est la taille de la commune, on est dans la situation de la fusion de deux petites communes dans l'échantillon avec une grande commune hors échantillon. Si la variable d'intérêt est bien corrélée avec x, le poids $\theta_{ji}^{(3)}$ tenant compte de x est alors le plus cohérent. L'arrivée de la grande commune dans la fusion provoque en effet une augmentation substantielle de la variable y. Cette augmentation est alors attenuée par la réduction du poids.

L'estimateur basé sur les poids provenant du partage des poids est sans biais quelle que soit la variable θ_{ji} de pondération. Sa variance varie cependant en fonction de la

corrélation entre y et θ_{ji}[1].

4.4 Post-stratification automatique sur la taille de la population ?

Il est intéressant de noter que la somme des poids estimés par le partage des poids n'est plus égale à la taille M^A de la population d'origine U^A, ni à la taille N^B de U^B.

Cette propriété peut être un inconvénient dans les analyses, notamment dans les estimations de moyennes et proportions. Il s'agit en effet de choisir entre un dénominateur égal à la taille connue ou égal à la taille estimée par la somme des poids. Les résultats seront différents.

La première idée venant à l'esprit pour remédier à ce problème de taille variable est, si les informations sont connues, de faire une post-stratification sur la taille N^B de U^B ou sur la taille de post-strates N_h^B.

Une post-stratification sur les tailles connues n'a cependant pas que des avantages. Prenons l'exemple d'une population de $N^A = 8$ communes avec $y_j = 10$ pour $j = 1, .., 8$ dont on tire un échantillon aléatoire simple s de $n = 4$ communes. Les poids des communes tirées sont $w_j = 2$, le total des poids $\sum_{j \in s} w_j = 8 = N$ et le total estimé $\widehat{Y} = 80$.

Supposons que deux communes de l'échantillon fusionnent. La taille de U^B devient $N^B = 7$. Il reste 3 communes dans l'échantillon, chacune avec un poids $w_i = 2$ (partage des poids simple). Le total estimé $\widehat{Y}$ est toujours égal à 80 ($y_i = \{10, 10, 20\}, i \in s$) mais la somme des poids est égale à $\sum_{i \in s} w_i = 6 < N^B$. Une post-stratification sur N^B permet d'obtenir une somme des poids égale à N^B mais le total estimé devient alors $\widehat{Y} = \sum_{i \in s} w_i y_i \cdot N^B / \sum_i w_i = 80 \cdot 7/6 = 93.3 > Y$.

Supposons maintenant que deux communes hors échantillon fusionnent. La taille de U^B devient $N^B = 7$. Le total estimé est égal à 80 mais la somme des poids est égale à $\sum_{i \in s} w_i = 8 > N^B$. Une post-stratification sur N^B permet d'obtenir une somme des poids égale à N^B mais le total estimé devient alors $\widehat{Y} = \sum_{i \in s} w_i y_i \cdot N^B / \sum_i w_i = 80 \cdot 7/8 = 70 < Y$.

Une post-stratification sur des tailles de population peut être faite pour des besoins particuliers, mais le petit exemple ci-dessus indique qu'elle n'est pas recommandable comme solution passe-partout. Une étude plus approfondie de la situation serait nécessaire pour comprendre les phénomènes en présence, notamment le comportement de la variance de l'estimation de la taille de la population. Dans le cas des communes, on s'intéresse à des estimateurs de totaux et on choisit de ne pas post-stratifier sur la taille de la population.

[1] L'estimateur de la variance de la MGPP de Lavallée (2002, p. 46-47) est basé sur la variable $z_{ik} = Y_i / L_i^B$.

4.5 Modifications successives

Les modifications successives de la liste des unités doivent être traitées selon une méthode adaptée au partage des poids. Il n'est en effet pas toujours possible de reprendre les poids estimés par la MGPP comme s'ils étaient de vrais poids de sondage pour estimer des poids lors de nouvelles fusions. Les estimateurs peuvent en effet être biaisés (Lavallée et Deville, 2002).

Soient la population U^A dans laquelle on a tiré un échantillon s^A et l'ensemble des éléments atteints de U^B :

$$\Omega^B = \{i \in U^B \mid \text{ il existe } j \in s^A \text{ avec } \theta_{ji}^{AB} > 0\}.$$

On identifie alors l'ensemble des unités correspondantes de U^C :

$$\Omega^C = \{k \in U^C \mid \text{ il existe } i \in \Omega^B \text{ avec } \theta_{ik}^{BC} > 0\}.$$

Si la méthode des poids généralisés de U^A à U^B puis de U^B à U^C donne les mêmes résultats qu'une application de la méthode directement de U^A à U^C, on parle de transitivité.

Soient les matrices de transition $\Theta_{AB} = [\theta_{ji}^{AB}]$ de dimension $N^A \times N^B$ et $\Theta_{BC} = [\theta_{ik}^{BC}]$ de dimension $N^B \times N^C$. Les éléments des matrices sont positifs si un lien existe entre les deux unités en indice et nuls sinon. On définit également les matrices de liaison standardisées correspondantes :

$$\tilde{\Theta}_{AB} = \Theta_{AB}\, \text{diag}(\mathbf{1}'_A \Theta_{AB})^{-1}, \tag{28}$$

$$\tilde{\Theta}_{BC} = \Theta_{BC}\, \text{diag}(\mathbf{1}'_B \Theta_{BC})^{-1}. \tag{29}$$

Lavallée et Deville (2002) montrent que la transitivité est respectée si :

$$\tilde{\Theta}_{AC} = \tilde{\Theta}_{AB}\tilde{\Theta}_{BC}. \tag{30}$$

Prenons l'exemple de U^A formé de trois communes de taille 10, 100 et 200. Une première fusion a lieu entre les deux premières communes afin de former U^B avec 2 communes de tailles 110 et 200. Une deuxième fusion conduit à U^C avec une unique commune de taille 310.

Trois types de liens sont appliqués : les liens basés sur la taille, les liens binaires $I_{j,i}$ et les liens correspondants au nombre de communes déjà fusionnées.

(1) Soient les matrices Θ_{AB}, Θ_{BC} et Θ_{AC}, basées sur les liens pondérés par la taille :

$$\Theta_{AB} = \begin{pmatrix} \theta_{11}^{AB} & \theta_{12}^{AB} \\ \theta_{21}^{AB} & \theta_{22}^{AB} \\ \theta_{31}^{AB} & \theta_{32}^{AB} \end{pmatrix} = \begin{pmatrix} 10 & 0 \\ 100 & 0 \\ 0 & 200 \end{pmatrix}, \tag{31}$$

$$\Theta_{BC} = \begin{pmatrix} \theta_{11}^{BC} \\ \theta_{21}^{BC} \end{pmatrix} = \begin{pmatrix} 110 \\ 200 \end{pmatrix}, \tag{32}$$

$$\Theta_{AC} = \begin{pmatrix} \theta_{11}^{AC} \\ \theta_{21}^{AC} \\ \theta_{31}^{AC} \end{pmatrix} = \begin{pmatrix} 10 \\ 100 \\ 200 \end{pmatrix}. \tag{33}$$

Les matrices standardisées sont alors :

$$\tilde{\Theta}_{AB} = \begin{pmatrix} 10/110 & 0 \\ 100/110 & 0 \\ 0 & 1 \end{pmatrix}, \tag{34}$$

$$\tilde{\Theta}_{BC} = \begin{pmatrix} 110/310 \\ 200/310 \end{pmatrix}, \tag{35}$$

$$\tilde{\Theta}_{AC} = \begin{pmatrix} 10/310 \\ 100/310 \\ 200/310 \end{pmatrix}. \tag{36}$$

Et la contrainte de transitivité est respectée :

$$\tilde{\Theta}_{AB}\tilde{\Theta}_{BC} = \begin{pmatrix} 10/110 & 0 \\ 100/110 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 110/310 \\ 200/310 \end{pmatrix} = \begin{pmatrix} 10/310 \\ 100/310 \\ 200/310 \end{pmatrix} = \tilde{\Theta}_{AC}. \tag{37}$$

(2) Dans le cas de liens binaires :

$$\begin{aligned} \Theta_{AB} &= \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} & \tilde{\Theta}_{AB} &= \begin{pmatrix} 1/2 & 0 \\ 1/2 & 0 \\ 0 & 1 \end{pmatrix} \\ \Theta_{BC} &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} & \tilde{\Theta}_{BC} &= \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \\ \Theta_{AC} &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} & \tilde{\Theta}_{AC} &= \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}. \end{aligned}$$

Et donc :

$$\tilde{\Theta}_{AB}\tilde{\Theta}_{BC} = \begin{pmatrix} 1/2 & 0 \\ 1/2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} = \begin{pmatrix} 1/4 \\ 1/4 \\ 1/2 \end{pmatrix} \neq \tilde{\Theta}_{AC}. \tag{38}$$

(3) Dans le cas de liens correspondants au nombre de communes déjà fusionnées, la règle est à nouveau satisfaite :

$$\begin{aligned} \Theta_{AB} &= \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} & \tilde{\Theta}_{AB} &= \begin{pmatrix} 1/2 & 0 \\ 1/2 & 0 \\ 0 & 1 \end{pmatrix} \\ \Theta_{BC} &= \begin{pmatrix} \mathbf{2} \\ 1 \end{pmatrix} & \tilde{\Theta}_{BC} &= \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix} \\ \Theta_{AC} &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} & \tilde{\Theta}_{AC} &= \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}. \end{aligned}$$

Et :

$$\tilde{\Theta}_{AB}\tilde{\Theta}_{BC} = \begin{pmatrix} 1/2 & 0 \\ 1/2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix} = \tilde{\Theta}_{AC}. \quad (39)$$

Sur la base du petit exemple, on note que la contrainte de transitivité semble respectée dans le cas des communes si les poids θ_{ji}^{AB} s'additionnent pour donner les poids θ_{ik}^{BC} ($\theta_{ik}^{BC} = \sum_{j \in U^A} \theta_{ji}^{AB} L_{ji}$). Les liens binaires ne satisfont pas la contrainte. Une utilisation transitive du partage des poids dans notre exemple de communes est donc valable uniquement si les poids d'origine x_j utilisés lors de la première année sont utilisés lors de chaque modification.

4.6 Utilisation future du partage des poids sur les communes

Dans le cas des communes, le partage des poids a été effectué pour 2004 dans deux cantons : Turgovie et Grisons.

Il est prévu qu'un nouvel échantillon soit tiré pour 2007 en Turgovie. Le partage des poids effectué dans ce canton n'aura donc plus d'importance sur le futur.

Si de nouvelles fusions sont observées aux Grisons pour 2005 et les années suivantes, il faudra choisir entre reprendre les poids estimés par le partage des poids en 2004 ou repartir de 1999. Dans la mesure où la transitivité est respectée, les deux solutions sont possibles.

Les fusions apparaissant dans les autres cantons pourront être traitées à partir des listes de 2004 si le relevé est exhaustif. Le traitement des fusions dans des cantons avec échantillonnage devront reprendre les données de 1999.

5 Ajustement des poids

Les poids $w_i^{(f)}$ de la Section 3 tiennent compte de l'algorithme de tirage de l'échantillon et des modifications de listes (morcellement et fusions). Pour des estimations au niveau de la population des communes de chaque canton, il s'agit maintenant de faire des ajustements pour tenir compte de la non-réponse et d'utiliser les informations auxiliaires pouvant améliorer les estimations (calage sur des valeurs de référence connues).

Le canton de Neuchâtel n'est pas considéré dans cette section car les données 2004 de ce canton ne sont pas disponibles.

5.1 Non-réponse

Certaines communes de l'échantillon de 2004 n'ont pas fourni de données. Il s'agit de communes dans les cantons de Zurich, Berne, Soleure, Bâle campagne, Schaffouse,

St-Gall, Argovie et Turgovie ; voir Table 3.

Le nombre de communes sans réponse varie entre les cantons, voir Table 6. On note par exemple les faibles taux de réponse à St-Gall et Turgovie, et, dans une moindre mesure, à Soleure. Il faut aussi noter qu'un taux de réponse élevé peut en fait cacher le fait que certaines grandes communes manquent. C'est le cas par exemple de Pratteln (BL, 15'000 habitants, 4ème plus grande ville du canton) ou encore de Zuchwil (SO, 9'000 habitants, 4ème plus grande ville du canton) qui ont livré les données trop tard pour être intégrées dans les estimations. Les taux de réponse pondérés par la population $propP$ sont inférieurs aux taux de réponse non pondérés m/n si les communes manquantes sont en moyenne plus grandes que les communes ayant répondu.

TAB. 6: Nombre et taux de réponse dans les 8 cantons avec non-réponse, avec `Nfus` le nombre de communes dans le canton après morcellement et fusions, n la taille de l'échantillon brut selon la liste de 2004, m le nombre de répondants pour 2004 et m/n le taux de réponse. La colonne $propP$ indique le taux de réponse pondéré par la taille de la population.

Kt	Canton	Nfus	n	m	m/n [%]	$propP$ [%]
1	ZH	171	87	84	96.6	99.4
2	BE	398	82	80	97.6	99.1
11	SO	126	44	39	88.6	89.6
13	BL	86	86	85	98.8	94.3
14	SH	32	32	31	96.9	98.9
17	SG	90	90	65	72.2	75.4
19	AG	231	231	229	99.1	99.2
20	TG	80	54	43	79.6	87.6

Il est important de noter que nous faisons l'hypothèse, dans cette section, que les données disponibles correspondent exactement aux données des communes telles qu'elles sont listées en 2004. Les collaborateurs de la section SOZ ont ainsi contrôlé que les communes fusionnées ont bien fourni les données pour l'ensemble de la fusion et que les non-réponse concernaient également les entités dans l'état de 2004.

Nous faisons un ajustement des poids $w_i^{(f)}$ selon un modèle de classes de non-réponse homogène. Deux variables, disponibles pour l'ensemble des communes de la liste de 2004, sont conservées pour le choix des classes ou cellules d'ajustement de la non-réponse : le nombre d'habitants dans la commune (état : fin 2004) et le type de commune (1 : ville, 2 : agglomération et 4 : rurale).

Les cellules choisies pour l'ajustement des poids sont basées principalement sur les codes de types de communes car cette caractéristique est plus influente que la classe de taille sur le fait de répondre ou pas à l'enquête. Des classes de taille sont utilisées uniquement dans le canton de St-Gall. Il y a en effet beaucoup de non-réponse dans ce canton et l'idée est de séparer explicitement les plus grandes communes car elles ont toutes répondu ; voir Table 7. Le calage sur la taille de la population, prévu comme ajustement complémentaire, permettra de prendre en compte également la taille de la

population dans les estimations de tous les cantons.

Les poids $w_i^{(nr)}$ ajustés pour la non-réponse de la commune i ayant répondu et étant classifiée dans la cellule c sont définis comme suit :

$$w_i^{(nr)} = w_i^{(f)} \cdot \rho_c, \tag{40}$$

avec le facteur d'ajustement :

$$\rho_c = \begin{cases} n_c/m_c & \text{si } i \text{ dans canton avec non-réponse et dans cellule } c \\ 1 & \text{sinon} \end{cases} \tag{41}$$

avec n_c la taille de l'échantillon brut et m_c la taille de l'échantillon net dans la cellule c.

On note que la correction de non-réponse est - sans surprise - spécialement importante dans les cantons de SG et TG. Elle est aussi relativement grande pour SO ; voir Table 7.

TAB. 7: Cellules d'ajustement pour la non-réponse. Information sur le nombre total N_c, la taille des échantillons brut n_c et net m_c, le taux de réponse $r_c = m_c/n_c$ [%] et le facteur d'ajustement des poids $\rho_c = 1/r_c = n_c/m_c$.

Kt	Canton	Cellule	N_c	n_c	m_c	r_c	ρ_c
1	ZH	ville	5	5	5	100.0	1.00
		agglo	122	70	67	95.7	1.04
		rural	44	12	12	100.0	1.00
2	BE	ville	7	7	7	100.0	1.00
		agglo	80	33	33	100.0	1.00
		rural	311	42	40	95.2	1.05
11	SO	ville	3	3	3	100.0	1.00
		agglo	58	27	23	85.2	1.17
		rural	65	14	13	92.9	1.08
13	BL	agglo	52	52	51	98.1	1.02
		rural	34	34	34	100.0	1.00
14	SH	ville	1	1	1	100.0	1.00
		agglo	8	8	8	100.0	1.00
		rural	23	23	22	95.7	1.05
17	SG	ville_agglo_<10'000 hab	29	29	18	62.1	1.61
		ville_agglo_>=10'000 hab	7	7	7	100.0	1.00
		rural	54	54	40	74.1	1.35
19	AG	ville	6	6	6	100.0	1.00
		agglo	79	79	78	98.7	1.01
		rural	146	146	145	99.3	1.01
20	TG	ville_agglo	21	17	12	70.6	1.42
		rural	59	37	31	83.8	1.19

5.2 Calage

La caractéristique d'intérêt principale de l'enquête est le nombre total de bénéficiaires de l'aide sociale en Suisse et par canton. On observe que ce total est habituellement bien corrélé à la population de la commune. Il est alors intéressant de caler les résultats de l'enquête sur les totaux de population dans des classes de tailles de communes. Ce calage permet aussi potentiellement de redresser des biais dus à un ajustement non complet de l'effet de la non-réponse.

Le calage est appliqué pour tous les cantons. Il aura un effet dans les cantons avec échantillonnage et/ou non-réponse, *i.e.* dans les 9 cantons avec au moins une commune ayant un poids différent de 1 ; voir Table 8.

Trois groupes sont définis dans chaque canton c :

1. communes avec $\texttt{pop} < \texttt{lim1}_c$
2. communes avec $\texttt{lim1}_c \leq \texttt{pop} < \texttt{lim2}_c$
3. communes avec $\texttt{pop} \geq \texttt{lim2}_c$

avec pop la population de la commune à la fin 2004.

Les limites $\texttt{lim1}_c$ et $\texttt{lim2}_c$ sont différentes entre les cantons. Elles sont reprises des travaux effectués pour le plan d'échantillonnage ; voir la méthode de Dalenius dans Renaud (2001).

Soient x_i le nombre d'habitants dans la commune i à la fin 2004 et X_ℓ les valeurs de référence du total de la population dans les cellule de calage ℓ. Les poids ajustés pour le calage $w_i^{(c)}$ sont définis comme suit pour la commune i de la cellule ℓ :

$$w_i^{(c)} = w_i^{(nr)} \cdot \phi_\ell = w_i^{(nr)} \cdot \frac{X_\ell}{\widehat{X}_\ell}, \tag{42}$$

avec

$$\widehat{X}_\ell = \sum_{i \in S_\ell} w_i^{(nr)} \cdot x_i. \tag{43}$$

Le calage a un faible effet, inférieur à 5% dans les trois cellules des cantons GR et AG. L'effet reste inférieur à 10% pour les cantons SO, BL, SG et SH. Il est cependant important dans tout ou partie des cellules des cantons de ZH, BE et TG. L'effet le plus important est égal à 57%. Il est observé dans les petites communes du canton de Zurich, où l'échantillon contient des communes plutôt petites par rapport à l'entier de la cellule.

TAB. 8: Limite des cellules de calage, avec nombre de communes ayant répondu n_ℓ, nombre de commune total N_ℓ, totaux de référence X_ℓ, totaux estimés $\widehat{X}_\ell$ et facteur d'ajustement ϕ_ℓ pour chaque cellule (1 : petites communes, 2 : communes moyennes et 3 : grandes communes). Restriction aux cantons avec échantillonnage ou non-réponse.

	lim1	lim2	n_1	N_1	X_1	$\widehat{X}_1$	ϕ_1
ZH	2600	6800	14	75	91296	58124.0	1.571
BE	1000	3400	10	213	96219	94871.7	1.014
GR	200	1600	35	72	8199	8049.4	1.019
SO	800	2600	10	49	18870	19898.6	0.948
BL	1000	4600	40	40	21532	21690.2	0.993
SH	400	1200	9	9	2322	2411.8	0.963
SG	2600	5200	27	33	45641	48902.3	0.933
AG	1200	3600	98	98	65261	65770.6	0.992
TG	1400	3000	14	30	27977	28463.4	0.983

	n_2	N_2	X_2	$\widehat{X}_2$	ϕ_2	n_3	N_3	X_3	$\widehat{X}_3$	ϕ_3
ZH	36	58	246744	288184.2	0.856	34	38	894197	897475.6	0.996
BE	14	126	246523	217279.5	1.135	56	59	612179	605639.4	1.011
GR	30	110	65597	64535.7	1.016	25	26	116387	114755.0	1.014
SO	8	49	68201	72265.1	0.944	21	28	158926	147041.8	1.081
BL	27	27	49111	49915.8	0.984	18	19	190696	179115.5	1.065
SH	12	13	9032	8452.1	1.069	10	10	62588	63114.6	0.992
SG	21	32	123498	118107.8	1.046	17	25	285228	262887.8	1.085
AG	87	89	184333	181399.8	1.016	44	44	300933	303643.1	0.991
TG	12	30	62316	57586.0	1.082	17	20	139846	177522.8	0.788

5.3 Un peu plus sur la pondération

La pondération présentée dans les Sections 5.1 et 5.2 est choisie comme base pour les estimations officielles de l'enquête. Elle a été choisie après avoir calculé quelques variantes envisageables au niveau des ajustements de non-réponse et de calage. L'analyse de ces variantes permet notamment d'étudier la sensibilité des résultats aux modèles choisis. Les résultats présentés dans cette section illustrent les concepts décrits mais ne sont pas les résultats officiels de l'enquête.

Les variables d'intérêt pour les comparaisons sont le nombre de dossiers ($y^D = \texttt{Ndos}$) et le nombre de bénéficiaires ($y^P = \texttt{Npers}$) dans la commune. La caractéristique d'intérêt est le total Y^D, respectivement Y^P.

Les estimations de totaux ont été réalisées pour trois variantes dans les 9 cantons de la Table 8 :

1. $\widehat{Y}_{nr_c}$: ajustement non-réponse et calage tels que décrits dans les Sections 5.1 et 5.2.
2. $\widehat{Y}_{nr}$: ajustement non-réponse tel que décrit dans la Section 5.1 mais pas de calage.
3. $\widehat{Y}_c$: ajustement directement par calage dans les cellules de la Section 5.2 (classes de taille) sans passer explicitement par un ajustement pour la non-réponse. Cette variante est identique à un ajustement non-réponse et calage dans les classes de taille de la Section 5.2.

Le choix de la correction non-réponse n'a que peu d'influence sur les totaux ; voir Table 9. En effet, la différence relative entre $\widehat{Y}_c$, dont l'ajustement de non-réponse se base sur la typologie des communes, et $\widehat{Y}_{nr_c}$, dont l'ajustement de non-réponse repose sur des classes de taille, a un maximum de env. 2% pour SG et SO. Le calage a une influence non négligeable sur les totaux estimés, avec des différences relatives variant de -6.7% pour SG à +20.4% pour TG. Ces deux cantons sont cependant liés, respectivement à un mauvais taux de réponse et à une grande modification de la liste des communes combinée avec un bas taux de réponse. Si la précision visée lors de l'échantillonnage est atteinte (coefficient de variation de 5%), la différence entre les trois pondérations n'est probablement significative que pour TG.

La pondération choisie a l'avantage de satisfaire aux contraintes de calage tout en intégrant un ajustement non-réponse qui semble raisonnable. La figure 4 présente le lien entre les poids $w_i^{(f)}$ (échantillonnage, morcellement, fusion) et les poids finaux choisis pour les estimations $w_i^{(c)}$ (correction de la non-réponse et calage).

Avec les poids finaux choisis pour les estimations, on obtient une proportion de bénéficiaires de l'aide sociale correspondant à 1.3- 3.8% de la population ; voir Table 10. Le nombre moyen de personnes par dossier est très stable entre les cantons, avec des valeurs entre 1.6 et 1.9.

TAB. 9: Comparaison des totaux, avec n la taille de l'échantillon net, $\widehat{Y}_{nr_c}$ l'estimation avec poids ajustés non-réponse et calage, $\widehat{Y}_{nr}$ l'estimation avec poids ajustés pour la non-réponse, $\widehat{Y}_c$ l'estimation avec poids calés, puis différences relatives à $\widehat{Y}_{nr_c}$.

Ndos					Diff. rel. à $\widehat{Y}^D_{nr_c}$ [%]	
kt	n	$\widehat{Y}^D_{nr_c}$	$\widehat{Y}^D_{nr}$	$\widehat{Y}^D_c$	$\widehat{Y}^D_{nr}$	$\widehat{Y}^D_c$
AG	229	4557.7	4569.6	4559.3	0.3	0.0
BE	80	19481.7	18965.5	19518.4	-2.6	0.2
BL	85	4396.9	4182.2	4397.2	-4.9	0.0
GR	90	1394.3	1374.1	1394.3	-1.4	0.0
SG	65	6167.0	5756.9	6307.6	-6.7	2.3
SH	31	987.9	994.0	992.0	0.6	0.4
SO	39	3061.9	2922.0	3128.9	-4.6	2.2
TG	43	2885.1	3473.6	2852.9	20.4	-1.1
ZH	84	27542.6	27868.5	27691.1	1.2	0.5
Npers					Diff. rel. à $\widehat{Y}^P_{nr_c}$ [%]	
kt	n	$\widehat{Y}^P_{nr_c}$	$\widehat{Y}^P_{nr}$	$\widehat{Y}^P_c$	$\widehat{Y}^P_{nr}$	$\widehat{Y}^P_c$
AG	229	8555.1	8576.1	8554.4	0.2	0.0
BE	80	34893.4	33939.3	34946.8	-2.7	0.2
BL	85	7319.7	6965.7	7320.3	-4.8	0.0
GR	90	2546.1	2509.3	2546.1	-1.4	0.0
SG	65	11072.0	10343.2	11265.3	-6.6	1.7
SH	31	1847.4	1858.3	1854.4	0.6	0.4
SO	39	5151.7	4925.2	5253.9	-4.4	2.0
TG	43	4754.4	5711.1	4692.3	20.1	-1.3
ZH	84	47080.3	47675.1	47310.1	1.3	0.5

TAB. 10: Totaux estimés pour le nombre de dossiers $\widehat{Y}^D_{nr_c}$ et de personnes $\widehat{Y}^P_{nr_c}$ avec les poids finaux choisis pour les estimations. Information également sur le nombre moyen de personnes par dossier $Pers/Dos = \widehat{Y}^P_{nr_c}/\widehat{Y}^D_{nr_c}$, la population du canton $pop04$ et la part de bénéficiaires dans la population $PersP = \widehat{Y}^P_{nr_c}/Pop04$.

kt	n	$\widehat{Y}^D_{nr_c}$	$\widehat{Y}^P_{nr_c}$	$Pers/Dos$	$Pop04$	$PersP$
AG	229	4557.7	8555.1	1.9	550527	1.6%
BE	80	19481.7	34893.4	1.8	954921	3.7%
BL	85	4396.9	7319.7	1.7	261339	2.8%
GR	90	1394.3	2546.1	1.8	190183	1.3%
SG	65	6167.0	11072.0	1.8	454367	2.4%
SH	31	987.9	1847.4	1.9	73942	2.5%
SO	39	3061.9	5151.7	1.7	245997	2.1%
TG	43	2885.1	4754.4	1.6	230139	2.1%
ZH	84	27542.6	47080.3	1.7	1232237	3.8%

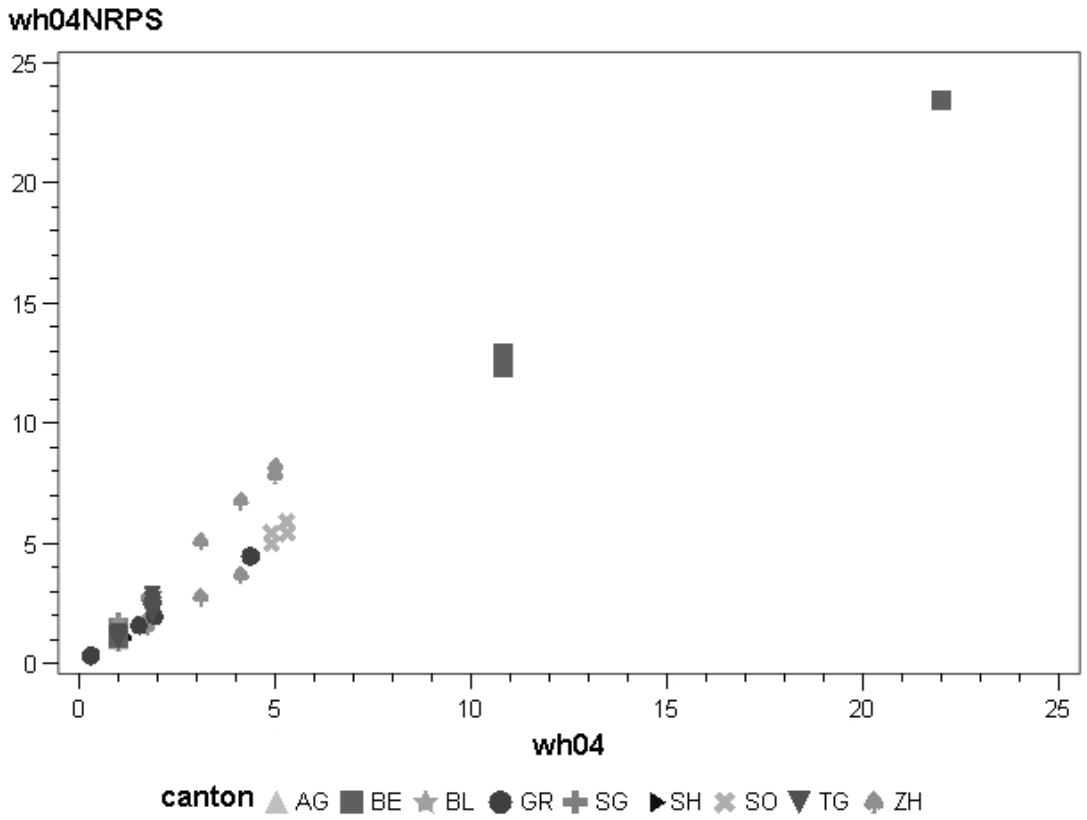


FIG. 4: Lien entre les poids $w_i^{(f)}$ (wh04) et les poids finaux $w_i^{(c)}$ (wh04NRPS).

6 Conclusion

La pondération 2004 proposée tient compte de la modification de la liste des communes par morcellement et fusion, ainsi que de la non-réponse et d'un calage sur des valeurs de référence connues.

Divers essais ont été faits avant de fixer la pondération. Ces essais ont montré que la pondération choisie pour 2004 fournit des estimations dont la qualité varie fortement entre les cantons.

La plus grande difficulté se trouve dans les cantons avec un mauvais taux de réponse. Il s'agit des cantons de SO (5 non-réponses sur 44 communes), TG (9 non-réponses sur 54 communes) et surtout SG (25 non-réponses sur 90 communes). Les estimations sont peu fiables dans ces cantons car le nombre de communes est faible et les résultats sont fortement dépendants du choix du modèle d'ajustement de la non-réponse et du calage.

Les grandes communes non-répondantes ont aussi un effet néfaste sur les estimations. L'effet est par exemple visible à BL.

Le partage des poids, effectif pour TG et GR, a permis de déterminer des poids pour des entités inexistantes au moment du tirage. Il n'a que peu d'influence sur la qualité des estimations car peu de communes sont concernées.

L'enquête de 2004 comporte pour la première fois tous les cantons[2]. Ces premiers résultats complets permettent de mettre en évidence l'impact très négatif de la non-réponse. Sur la base de ces résultats et des expériences lors du relevé, les prochaines enquêtes pourront être améliorées. Il est également prévu de refaire de tirer un nouvel échantillon dans le canton de Turgovie pour 2007 afin d'éviter de se retrouver confronté aux gros problèmes de listes. Des estimations de variance permettront également de compléter les résultats avec des précisions estimées.

Références

Lavallée, P., Deville, J.-C. (2002), Theoretical Foundations of the Generalised Weight Share Method, Proceedings of the International Conference on Recent Advances in Survey Sampling, 10-13 juillet 2002.

Lavallée, P. (2002) Le sondage indirect ou la méthode généralisée du partage des poids. Editions de l'Université de Bruxelles et Ellipses.

Renaud, A. (2001) Statistique suisse des bénéficiaires de l'aide sociale. Plan d'échantillonnage des communes. Office fédéral de la statistique. Rapport de méthodes 338-0001.

[2]Le canton de Neuchâtel n'est pas pris en compte dans ce rapport. Une solution ad hoc est développée dans la section SOZ en collaboration avec le canton.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Publikationsprogramm BFS

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

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

Programme des publications de l'OFS

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

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

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

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

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

Methodenberichte des Dienstes Statistische Methoden
Rapports de méthodes du Service de méthodes statistiques
Methodological reports of 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.

La statistique suisse des bénéficiaires de l'aide sociale fournit des résultats pour tous les cantons pour la première fois en 2004. Cette statistique est basée sur un échantillon de communes sélectionnées en 1999. L'échantillon est exhaustif dans certains cantons et non-exhaustif dans d'autres. Etant prévu pour durer plusieurs années, l'échantillon est touché par des morcellements et fusions de communes. Les données relevées en 2004 concernent donc des communes qui ne correspondent plus toujours exactement aux communes sélectionnées en 1999. Cette particularité est traitée à l'aide de la méthode généralisée du partage des poids. Les poids ainsi construits sont ensuite adaptés pour tenir compte de la non-réponse et d'un calage sur des tailles de populations connues. Des difficultés d'estimations sont détectées dans les cantons avec un faible taux de réponse, les cantons avec une grande modification de la liste des communes ou encore les cantons avec un bon taux de réponse global mais de la non-réponse dans quelques grandes communes.

N° de commande: 338-0039	Commandes: 032 713 60 60	Fax: 032 713 60 61 E-Mail: order@bfs.admin.ch	Prix: gratuit	ISBN 3-303-00339-4