

2018

00

Bases
statistiques
et généralités

Neuchâtel 2020

Enquête suisse sur la structure des salaires 2018

Révision du plan de sondage

Rapports de méthodes de la section Méthodes statistiques

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

Enquête suisse sur la structure des salaires 2018

Révision du plan de sondage

Rédaction	Lionel Qualité, OFS; Jann Potterat, OFS
Contenu	Section méthodes statistiques, OFS
Éditeur	Office fédéral de la statistique (OFS)

Neuchâtel 2020

Éditeur: Office fédéral de la statistique (OFS)

Renseignements: Lionel Qualité, tél. 058 467 24 54
lionel.qualite@bfs.admin.ch

Rédaction: Lionel Qualité, OFS; Jann Potterat, OFS

Contenu: Section méthodes statistiques, OFS

Série: Statistique de la Suisse

Domaine: 00 Bases statistiques et généralités

Langue du texte original: français

Mise en page: section DIAM, Prepress/Print
Ce document a été produit automatiquement à partir de banques de données. Il ne répond donc pas aux normes typographiques des publications de l'OFS.

En ligne: www.statistique.ch

Imprimés: www.statistique.ch
Office fédéral de la statistique, CH-2010 Neuchâtel,
order@bfs.admin.ch, tél. 058 463 60 60
Impression réalisée en Suisse

Copyright: OFS, Neuchâtel 2020
La reproduction est autorisée, sauf à des fins commerciales, si la source est mentionnée.

Numéro OFS: 338-0081

ISBN: 978-3-303-00657-3

Table des matières

Préambule et Résumé	5
1 Introduction	7
1.1 Situation initiale	7
1.2 Description des spécificités de l'enquête	7
1.3 Brève description	7
2 Population cible, données auxiliaires, plan de sondage et cadres de sondage	8
2.1 Population cible	8
2.2 Données auxiliaires AVS	8
2.3 Plan de sondage	8
2.4 Cadres de sondage	9
2.4.1 Cadre des administrations publiques	9
2.4.2 Cadre du secteur privé et des entreprises publiques	10
3 Stratégie pour le calcul des taux de sondage	13
3.1 Calcul de la précision des statistiques d'intérêt	14
3.1.1 Variance de l'estimateur de Horvitz-Thompson	14
3.1.2 Variance dans un plan à deux degrés	15
3.1.3 Non-réponse	16
3.1.4 Calage	17
3.1.5 Linéarisation d'un estimateur complexe	18
3.1.6 Mise en évidence de l'allocation de l'ESS dans le calcul de variance	19
3.2 Hypothèses et données utilisées pour le calcul des variances	19
3.2.1 Revenus AVS	20
3.2.2 Proportion de salaires livrés	20
3.2.3 Probabilités de réponse	20
3.2.4 Linéarisation et variance approchée	21
3.3 Meilleure précision possible et objectifs retenus	22
3.4 Optimisation de la taille d'échantillon	22
3.4.1 Optimisation pour une statistique d'intérêt	23
3.4.2 Optimisation pour plusieurs statistiques d'intérêt	23
3.4.3 Travaux ultérieurs	24
4 Résultat des calculs d'allocation	26
5 Sélection de l'échantillon	30
6 Possibilités d'amélioration et évolutions	33
7 Conclusion	35
Bibliographie	36
Annexes	37
A Coefficients de variation par catégorie d'activité et par grande région	37
B Procédure d'optimisation pour plusieurs contraintes de précision	39
C Allocation optimale	40
Rapports de méthodes de la section méthodes statistiques de l'OFS	41

Préambule

L'enquête suisse sur la structure des salaires (ESS) est réalisée tous les deux ans par l'Office fédéral de la statistique (OFS) auprès d'un échantillon d'entreprises et d'administrations publiques. Elle relève des informations sur les salaires versés, la nature des emplois correspondants et les caractéristiques des employés qui perçoivent ces salaires. La section des méthodes statistiques (METH) de l'OFS a reçu le mandat de réviser le plan de sondage de l'enquête suisse sur la structure des salaires 2018 (ESS2018). Il s'agissait d'utiliser les données administratives disponibles sur les revenus des personnes salariées en Suisse afin d'améliorer l'efficacité de la stratégie d'estimation. Une nouvelle procédure d'allocation de l'échantillon a été développée pour cela. Elle est basée sur la spécification de statistiques principales d'intérêt (salaires médians notamment) et d'objectifs de précision à respecter pour leur estimation. Ce rapport décrit les cadres de sondage, le plan de sondage utilisé, les hypothèses retenues pour le calcul d'allocation et les objectifs de précision définis. Enfin, le calcul effectif du plan et le résultat du tirage sont fournis. Les auteurs remercient la section Salaires et conditions de travail (LOHN), qui est responsable de l'enquête suisse sur la structure des salaires et en particulier D. Droz, D. Froidevaux, J. Ignaczewska, A. Rouvinez et F. Santi pour le soutien qu'ils ont apporté à ce travail. Les auteurs remercient également D. Assoulin, D. Kilchmann et J.-P. Renfer de la section METH pour leur relecture attentive de ce rapport.

Résumé

Le plan de sondage de l'ESS a fait l'objet d'une révision approfondie pour l'enquête de l'année 2018. Le principal objectif de cette révision est d'améliorer l'efficacité de l'enquête en utilisant les données administratives disponibles sur les revenus des personnes salariées. À précision fixée, cela doit permettre de réduire la charge d'enquête qui pèse sur les entreprises du pays. Tous les deux ans, un échantillon d'entreprises est sélectionné pour participer à l'ESS. L'échantillon est sélectionné aléatoirement dans le registre des entreprises et des établissements (REE) de l'office fédéral de la statistique (OFS). Le champ de l'ESS est formé par les salaires des personnes employées au mois d'octobre 2018 dans les entreprises et administrations des secteurs secondaires et tertiaires qui emploient au moins trois personnes. Les données administratives sur les revenus contiennent, entre autres variables, un identifiant de l'entreprise et de l'employé concerné, un revenu annuel et une indicatrice de cotisation pour chaque mois de l'année. Le plan de sondage, comme pour les enquêtes précédentes, assigne à chaque entreprise ou administration une probabilité de sélection déterminée en fonction de sa grande région, de son canton ou de sa commune de densification, de son activité économique et de son nombre d'employés selon le REE. Toutes les entreprises de 50 employés ou plus, administrations de 800 employés ou plus, ainsi que toutes celles du programme profiling sont incluses dans l'échantillon. La méthode utilisée pour calculer les taux de sondage de l'ESS2018 repose sur le choix de paramètres d'intérêt principaux et de la précision avec laquelle ces paramètres doivent être estimés. L'allocation obtenue pour l'échantillon des administrations publiques (communes, districts, cantons et Confédération) prévoit la sélection de 323 unités en espérance, dont la Confédération, les 26 cantons et demi-cantons, et 296 communes. Le résultat de l'allocation pour le cadre du secteur privé et des entreprises publiques prévoit un échantillon de 45'663 entreprises en espérance. Cette allocation se décompose en un échantillon de 39'163 entreprises pour les besoins de l'OFS et 6500 entreprises supplémentaires pour les cantons et communes qui ont décidé de densifier leur échantillon. L'échantillon complet du secteur public et du secteur privé effectivement tiré pour l'ESS2018 compte 45'877 unités. Cela représente une diminution de 4691 unités, soit 9.3% par rapport à celui de l'ESS2016 qui en comptait 50'568.

1 Introduction

1.1 Situation initiale

Le plan de sondage de l'enquête suisse sur la structure des salaires a été mis à jour pour l'enquête portant sur l'année 2018. Le principal objectif de cette révision est d'utiliser de nouvelles données administratives pour améliorer l'efficacité de l'enquête, c'est à dire pour améliorer la relation entre précision des estimations et taille de l'échantillon. À précision fixée, cette stratégie doit ainsi permettre de réduire la charge d'enquête qui pèse sur les entreprises du pays. Pour la première fois, les données des caisses de compensation AVS (Assurance-vieillesse et survivants) concernant les revenus des personnes salariées ont été utilisées lors du développement du plan de sondage de l'ESS. Ces données sont mises à disposition de l'Office fédéral de la statistique (OFS) par les caisses de compensation AVS et sont traitées et gérées par la section URD (Unternehmensregisterdaten) de l'OFS. Elles permettent de connaître, pour chaque employé d'une entreprise, le revenu annuel entraînant une obligation de cotiser au titre du premier pilier de la prévoyance vieillesse et le nombre de mois d'activité concernés.

1.2 Description des spécificités de l'enquête

Tous les deux ans, un échantillon d'entreprises est sélectionné pour participer à l'ESS. Cette enquête permet d'obtenir des informations sur les niveaux des salaires versés par les entreprises en Suisse lors de l'année de référence. Les données relevées permettent d'estimer les salaires médians dans des domaines d'activité choisis, le coût horaire de la main d'œuvre dans différentes branches d'activité, et de mettre en regard ces salaires avec les caractéristiques socio-démographiques des personnes qui les perçoivent. Les résultats sont également utilisés pour l'élaboration du calculateur de salaires « Salarium » de l'OFS et pour l'étude des inégalités salariales entre les femmes et les hommes. Certains cantons et la ville de Zurich financent une densification de leur échantillon pour permettre des exploitations locales. L'ESS est une enquête structurelle qui a pour objectif d'informer sur les salaires à un moment donné. Elle n'est pas optimisée pour estimer les évolutions des salaires.

1.3 Brève description

L'échantillon de l'ESS est sélectionné aléatoirement dans le registre des entreprises et des établissements (REE) de l'OFS. En 2018, cet échantillon contenait 45'877 unités du secteur public et du secteur privé. Selon leur taille, les entreprises et administrations sélectionnées sont invitées à renseigner l'enquête pour une partie de leurs employés salariés en octobre : tous les employés pour les *petites* entreprises (3-19 employés), au moins la moitié d'entre eux pour les entreprises de taille *moyenne* (20-49 employés), et au moins le tiers pour les *grandes* entreprises (50 employés ou plus).

La population cible, les cadres de sondage, les données disponibles et le plan de sondage sont décrits dans la section 2. La méthode utilisée pour calculer l'allocation de l'échantillon fait l'objet de la section 3. Le résultat de ces calculs figure dans la section 4 et le résultat du tirage dans la section 5. Enfin, des possibilités de travaux ultérieurs et une conclusion sont donnés dans les sections 6 et 7.

2 Population cible, données auxiliaires, plan de sondage et cadres de sondage

2.1 Population cible

Le champ de l'ESS est formé par les salaires des personnes employées au mois d'octobre 2018 dans les entreprises et administrations des secteurs secondaires et tertiaires (codes 05 à 96 de la nomenclature générale des activités économiques - NOGA) qui emploient au moins trois personnes. Dans le secteur public, les communes et districts peuvent être enquêtés dès lors qu'ils emploient au moins une personne. Ces salaires sont observés grâce à un échantillon d'entreprises et d'administrations qui a quant à lui été sélectionné dans un extrait du REE daté du 06.08.2018. La population sélectionnable a été délimitée en utilisant le statut d'activité, le type d'entreprise, la catégorie d'activité et le nombre d'emplois renseignés dans le registre. Outre ces caractéristiques, le plan de sondage a été défini en prenant en considération le canton et la grande région des entreprises ainsi que les données des caisses de compensation AVS sur les revenus 2016. Ces données ont été utilisées dans leur état du 19.07.2018.

2.2 Données auxiliaires AVS

Le fichier des revenus AVS contient, entre autres variables, des identifiants de l'entreprise et de l'employé (numéro AVS) concerné, un revenu annuel et une indicatrice de cotisation pour chaque mois de l'année. On obtient un *revenu mensuel moyen* AVS en divisant le revenu annuel par le nombre de mois de cotisation. Ce revenu diffère du salaire standardisé calculé par l'ESS. En effet, il ne tient par exemple pas compte du taux d'occupation ou du nombre d'heures travaillées. Toutefois, les revenus AVS sont corrélés aux salaires standardisés. Ils constituent une information auxiliaire utile dont on peut tirer parti pour améliorer l'efficacité de la procédure d'enquête. Pour les calculs d'allocation, seuls les revenus mensuels moyens AVS supérieurs ou égaux à 200CHF des personnes actives en octobre ou en décembre ont été retenus. Un revenu annuel inférieur à 2400CHF n'entraîne pas de cotisation obligatoire à l'AVS. Les revenus annuels inférieurs à 2400CHF enregistrés dans les données des caisses de compensation sont uniquement ceux qui ont fait l'objet de cotisations volontaires et sont donc incomplets.

Les revenus AVS étaient déjà utilisés lors de la préparation des données relevées par l'ESS2016 : la comparaison des revenus AVS avec les résultats livrés par les entreprises répondantes permet d'identifier des possibles erreurs de réponse.

2.3 Plan de sondage

Le plan de sondage de l'ESS2018, comme pour les ESS précédentes, assigne à chaque entreprise ou administration du domaine d'étude présente dans les cadres de sondage une probabilité de sélection déterminée en fonction de sa grande région ou de sa zone de densification (canton ou commune), de son activité économique et de sa classe de taille. Les classes de taille sont définies à partir du nombre d'employés selon le REE :

- 3 à 19 employés : *petites* entreprises,
- 20 à 49 employés : *moyennes* entreprises,
- 50 employés ou plus : *grandes* entreprises.

Une quatrième pseudo-classe de taille, séparée, regroupe les entreprises qui font partie du programme « profiling » de partenariat entre les entreprises et l'OFS. Les 39 catégories d'activité économique (OFS39) considérées pour le plan de sondage sont issues de la nomenclature générale des activités économiques (NOGA). Elles sont énumérées plus loin dans le tableau 2. Ces catégories d'activité sont également utilisées pour la diffusion des résultats de l'ESS.

Toutes les entreprises de 50 employés ou plus dans les catégories d'activité concernées, les administrations publiques de 800 employés ou plus, ainsi que toutes les entreprises et administrations du programme profiling sont incluses dans l'échantillon. Les taux de sélection des autres entreprises et administrations font l'objet des calculs présentés dans les sections suivantes. La sélection de l'échantillon est opérée en utilisant un plan de Poisson : chaque unité est sélectionnée ou non indépendamment des autres unités. L'échantillon est tiré de manière coordonnée avec les autres enquêtes de l'OFS de manière à réduire autant que possible le nombre d'unités sélectionnées dans plusieurs enquêtes rapprochées dans le temps.

Les entreprises et administrations sélectionnées dans l'échantillon définissent les unités primaires d'un plan à deux degrés dont les unités secondaires sont les salaires versés par ces entreprises et administrations à leurs salariés pour le mois d'octobre 2018. Les unités primaires sélectionnent elles-mêmes les unités secondaires qu'elles livrent, en respectant les exigences de l'OFS :

- les petites unités primaires doivent livrer tous les salaires,
- les unités primaires de taille moyenne doivent livrer au moins la moitié des salaires,
- les grandes unités primaires et celles du profiling doivent livrer au moins le tiers des salaires.

Les unités primaires reçoivent des recommandations de l'OFS pour la sélection des salaires à livrer. Si ces recommandations sont suivies, l'échantillon du second degré résulte d'un sondage aléatoire simple dans chaque unité de l'échantillon du premier degré.

Certaines statistiques, à l'instar des salaires médians estimés, portent sur la population des heures de travail réalisées en octobre. Dans ce cas, les heures travaillées par les employés ou définies par leur contrat de travail peuvent être vues comme les unités finales d'un troisième degré d'échantillonnage. Ce troisième degré est exhaustif : la durée totale de travail et le salaire total associé sont entièrement renseignés pour chaque salaire d'octobre observé. Un exemple est le salaire médian dans un domaine selon l'ESS, qui est calculé comme suit : dans un premier temps, pour chaque employé du domaine concerné, un salaire mensuel brut standardisé est calculé pour le mois d'octobre. Le salaire mensuel brut standardisé est ramené à un temps plein de 40 heures par semaine, soit $173.\bar{3}$ heures en moyenne pour le mois d'octobre. Il comprend les compléments de salaire, $1/12^{eme}$ du 13^{e} salaire et des bonus le cas échéant, mais ne comprend pas des éventuelles heures supplémentaires payées.

Le salaire médian mensuel est alors défini comme étant la valeur m telle que la moitié des heures contractuelles ou travaillées par l'ensemble des employés du domaine l'ont été avec un salaire mensuel brut standardisé inférieur à m et l'autre moitié l'ont été avec un salaire mensuel brut standardisé supérieur à m . Ce sont ainsi les heures contractuelles ou travaillées qui sont partagées par la médiane en deux ensembles de tailles égales.

Cette définition de la médiane ne coïncide pas avec celle d'un salaire médian qui partagerait la population des employés en deux groupes de tailles égales, selon leur salaire mensuel brut standardisé ou encore selon leur salaire perçu.

2.4 Cadres de sondage

2.4.1 Cadre des administrations publiques

Le secteur public est scindé en deux. D'une part les administrations publiques : communes, districts, cantons et Confédération qui appartiennent par définition à la NOGA 84 *Administration publique et défense; sécurité sociale obligatoire*. D'autre part, les entreprises publiques qui sont regroupées avec le secteur privé dans leurs catégories d'activité NOGA respectives. Les cantons et la Confédération sont placés dans une classe de taille séparée qui est entièrement sélectionnée.

Le cadre des administrations publiques contient 2217 unités qui représentent 433'845 emplois

selon le REE. Ces unités sont réparties dans 34 strates de tirage définies par la grande région et la classe de taille de l'unité.

La ventilation des unités du cadre des administrations publiques dans les différentes grandes régions et dans les différentes classes de taille, avec les nombres d'emplois correspondants selon le REE, est donnée dans le tableau 1.

TABLE 1 Cadre des administrations publiques ventilé par grande région et par classe de taille.

		Nb. Unités	Nb. Emplois selon REE
Grande région	Région lémanique	483	78'133
	Espace Mittelland	669	132'304
	Suisse du Nord-ouest	290	48'778
	Zurich	167	75'703
	Suisse orientale	325	43'619
	Suisse centrale	167	37'889
	Tessin	116	17'419
Taille	3-19	1'279	10'761
	20-49	429	13'703
	50 ou +	435	56'877
	Profiling	47	65'315
	Cantons et Conféd.	27	287'189
Total		2'217	433'845

2.4.2 Cadre du secteur privé et des entreprises publiques

Le cadre du secteur privé et des entreprises publiques est constitué de 186'744 unités qui regroupent 4'112'483 emplois selon le REE. Il s'agit d'entreprises publiques ou privées qui sont en activité selon le REE, classées dans le secteur secondaire ou dans le secteur tertiaire (NOGA 05 à 96) et avec au minimum trois employés. De plus, dans le secteur privé, 34 entreprises qui ne disposent pas d'un identifiant UID et pour lesquelles il n'était pas certain qu'elles soient indépendantes et en capacité d'être enquêtées ont été exclues du cadre de sondage. Les unités sélectionnables sont réparties dans 2050 strates de tirage, déterminées par 39 catégories d'activité, 16 grandes régions ou zones de densification (GRS) et 4 classes de taille (certains croisements de ces variables sont vides). La localisation de l'entreprise est déterminée par le canton ou la ville dans lequel elle compte, au travers de ses établissements, le plus d'employés en équivalents plein temps selon le REE. Cette définition ne correspond pas nécessairement à la localisation du siège officiel de l'entreprise. La même définition est utilisée pour la sélection de l'échantillon et pour la production des résultats de l'enquête. La ventilation des unités et des emplois par catégorie d'activité apparaît dans le tableau 2. La ventilation par grande région ou zone de densification et la ventilation par classe de taille sont données dans le tableau 3.

TABLE 2 Cadre du secteur privé et des entreprises publiques ventilé en 39 catégories d'activité.

		Nb. Unités	Nb. Emplois selon REE
NOGA	SECTEUR 2 - PRODUCTION	42'939	1'017'825
05-09	Industries extractives	212	4'439
10-11	Industries alimentaires ; fabr. de boissons	2'887	93'946
12	Fabrication de produits à base de tabac	10	2'114
13-15	Industries du textile et de l'habillement	602	12'185
16-18	Industries du bois et du papier ; imprimerie	4'122	59'866
19-20	Cokéfaction ; industrie chimique	420	29'699
21	Industrie pharmaceutique	184	46'118
22-23	Industries du caoutchouc et du plastique	1'145	40'234
24-25	Métallurgie ; fabr. produits métalliques	4'174	90'780
26	Fabr. prod. informatiques, électroniques et optiques ; horlogerie	1'328	106'027
27	Fabrication d'équipements électriques	539	31'267
28	Fabr. de machines et équipements n.c.a	1'419	75'939
29-30	Fabrication de matériels de transport	247	16'727
31-33	Fabr. meubles ; autres ind. manufact. ; rép. et inst. machines	2'758	46'519
35	Production et distribution d'énergie	468	30'834
36-39	Prod. et distr. d'eau ; gestion déchets	923	16'708
41-43	Construction	21'501	314'423
	SECTEUR 3 - SERVICES	143'805	3'094'658
45-46	Commerce de gros ; com. et rép. d'automobiles	17'177	285'635
47	Commerce de détail	14'915	316'647
49-52	Transp. terrestres, par eau, aériens ; entreposage	4'586	178'213
53	Activités de poste et de courrier	239	51'503
55-56	Hébergement et restauration	17'249	228'860
58-60	Édition, audiovisuel et diffusion	1'243	31'597
61	Télécommunications	229	27'018
62-63	Activ. informatiques et services d'information	5'102	86'571
64, 66	Services financiers ; activ. auxiliaires de serv. fin. et d'ass.	5'328	167'067
65	Assurance	408	53'175
68	Activités immobilières	4'860	50'084
69-71	Activ. jur., comptables, de gestion, d'architecture, d'ingénierie	19'742	250'134
72	Recherche-développement scientifique	652	19'834
73-75	Autres activités spéc., scient. et techn.	3'780	41'484
77,79-82	Activités de services admin. (sans 78)	7'810	174'264
78	Activités liées à l'emploi	1'169	127'955
84	Admin. publ.	266	6'177
85	Enseignement	4'015	166'294
86-88	Santé humaine et action sociale	20'592	646'428
90-93	Arts, spectacles et activités récréatives	4'317	64'046
94-95	Activ. org. associatives et religieuses ; réparation biens pers.	6'005	90'759
96	Autres services personnels	4'121	30'913
Total		186'744	4'112'483

TABLE 3 Cadre du secteur privé et des entreprises publiques ventilé par grande région ou zone de densification et par classe de taille.

Grande région	GRS	Nb. Unités	Nb. Emplois selon REE
1- Région lémanique	GE	11'489	263'302
	VD	15'134	341'505
	Reste (VS)	7'994	117'748
2- Espace Mittelland	FR	5'755	101'862
	NE	3'617	78'897
	JU	1'647	26'973
	Reste (BE, SO)	26'579	632'200
3- Suisse du Nord-ouest	AG, BL, BS	22'616	532'862
4- Zurich	Ville de Zurich	13'118	466'808
	Reste ZH	20'163	517'376
5- Suisse orientale	TG	5'678	100'160
	Reste (AR, AI, GL, GR, SG, SH)	21'365	399'215
6- Suisse centrale	LU	8'822	186'617
	ZG	5'012	83'834
	Reste (LU, NW, OW, SZ, UR)	6'938	102'855
7- Tessin	TI	10'817	160'269
Taille	3 à 19	157'111	1'045'735
	20 à 49	17'698	529'500
	50 ou plus	10'233	1'582'020
	Profiling	1'702	955'228
Total		186'744	4'112'483

3 Stratégie pour le calcul des taux de sondage

La méthode utilisée pour calculer les taux de sondage de l'ESS2018 repose sur le choix de paramètres d'intérêt principaux et de la précision avec laquelle ces paramètres doivent être estimés. Plus précisément, les coefficients de variation associés à l'estimation de ces paramètres d'intérêt ne doit pas dépasser des valeurs choisies.

Les paramètres d'intérêt principaux retenus pour le calcul de l'allocation dans le cadre du secteur privé et des entreprises publiques sont :

1. les salaires médians par grande région, par catégorie d'activité OFS39 et par croisement de la grande région et de la catégorie d'activité OFS39,
2. les masses salariales dans les grandes régions, sections NOGA et dans les classes de taille de diffusion (3-9, 10-49, 50-249, 250-499, 500-999, 1000 employés et plus),
3. les salaires médians par zone de densification et par croisement du canton ou de la commune de densification et de la catégorie d'activité,
4. la masse salariale par zone de densification.

Les coefficients de variation (CV) visés ne doivent, si possible, pas excéder 3% pour chacun des estimateurs (statistiques d'intérêt) de ces paramètres. Du fait de la non-réponse et de la sélection des salaires livrés par les entreprises, ces objectifs ne sont pas tous réalisables, même en sélectionnant toutes les entreprises dans l'échantillon. De plus, lorsque le CV minimum, meilleur CV réalisable en sélectionnant toutes les entreprises, est très proche de l'objectif de précision visé, respecter strictement cet objectif de précision peut nécessiter de sélectionner un grand nombre d'entreprises. Les objectifs finalement retenus correspondent, pour l'essentiel des statistiques d'intérêt, aux règles suivantes :

- si le CV minimum est inférieur ou égal à 2.5% alors le CV visé est de 3%,
- si le CV minimum est compris entre 2.5 et 3.5% alors le CV visé est de 4%,
- si le CV minimum est compris entre 3.5 et 4.5% alors le CV visé est de 5%,
- si le CV minimum est compris entre 4.5 et 5% alors le CV visé est égal au CV minimum arrondi au pour mille supérieur et augmenté de 0.5%,
- si le CV minimum est supérieur à 5% alors le CV visé est égal au CV minimum arrondi au pour mille supérieur et augmenté de 1%.

Des exemples sont donnés plus loin dans le tableau 5.

Au total, 687 objectifs de précision ont été retenus pour le cadre du secteur privé et des entreprises publiques. Ces 687 objectifs se répartissent entre 342 objectifs définis pour la production de l'OFS et 345 objectifs définis pour les cantons et communes qui densifient leur échantillon. Le but de l'allocation est d'obtenir un échantillon de taille minimale, ou aussi petite que possible, tout en respectant les objectifs de précision. Un premier calcul d'allocation ne prend en compte que les 342 objectifs définis pour l'OFS. Il en résulte un plan de base auquel vont s'ajouter les densifications financées par certains cantons ou communes. Pour chaque canton ou commune qui le décide, est calculée une allocation supplémentaire qui permet d'atteindre les objectifs de précision qui le concernent. Cette allocation ne porte que sur l'échantillon du canton ou de la commune qui densifie : il n'y a pas d'augmentation de l'échantillon d'entreprises en dehors de la zone de densification. La densification vient toujours en supplément de l'allocation initiale de l'OFS. L'allocation finale est obtenue par l'addition de l'allocation OFS et des différentes densifications cantonales et communales.

Pour l'échantillon sélectionné dans le cadre des administrations publiques, les 14 objectifs de précision sont :

- un CV de 3% pour l'estimation de la médiane des salaires par grande région (sans la Confédération),

- un CV de 5% pour l'estimation de la médiane des salaires dans les quatre domaines suivants : Confédération, ensemble des cantons, communes et districts avec 50 employés ou plus et profiling, ,
- un CV de 5% pour l'estimation de la masse salariale totale dans les trois domaines suivants : Confédération, ensemble des cantons, communes et districts.

La Confédération, les cantons, les communes avec plus de 800 employés et celles qui participent au profiling sont sélectionnées d'office.

3.1 Calcul de la précision des statistiques d'intérêt

Un calcul de variance de chacune des statistiques d'intérêt principales est nécessaire pour appliquer le programme d'optimisation. Ces variances sont des fonctions des taux de sondage utilisés pour le tirage de l'échantillon d'unités primaires. Le calcul de variance est adapté à la méthode de sélection de l'échantillon, à la méthode de pondération et à la forme fonctionnelle des statistiques utilisées. Il est souvent impossible d'obtenir un calcul exact mais seulement des approximations de variance. C'est le cas par exemple lorsque les poids d'extrapolation de l'enquête sont calés, et lorsque les statistiques d'intérêt sont non-linéaires, à l'instar des estimateurs de médianes. De plus, le calcul repose sur des hypothèses concernant les paramètres qui ne sont pas entièrement contrôlées, telles que les probabilités de réponse des entreprises, ou la sélection des salaires livrés par les entreprises répondantes. Des formules approchées de ces variances dans le cas de l'enquête suisse sur la structure des salaires sont données dans la suite de cette section. Le cas simple de l'estimation d'un total si toutes les entreprises de l'échantillon répondent et livrent tous leurs salaires est considéré en premier. Sont intégrés progressivement dans le calcul les facteurs de complication que sont la sélection des salaires par les entreprises, la non-réponse, le calage des poids et la non-linéarité des estimateurs considérés.

3.1.1 Variance de l'estimateur de Horvitz-Thompson

Un échantillon s désigne une partie d'une population d'intérêt $U = \{1, \dots, N\}$. Un plan de sondage P est une loi de probabilité sur les échantillons s , c'est à dire une fonction qui associe une valeur dans $[0, 1]$ à chaque échantillon s et dont la somme sur tous les échantillons vaut 1 (voir par exemple Tillé, 2019). Pour chaque unité i de la population, on définit l'indicatrice d'appartenance $I_i(s)$ qui vaut 1 lorsque i est dans l'échantillon s et 0 sinon. Les probabilités d'inclusion d'ordre 1 et 2 sont définies respectivement par :

$$\pi_i = E(I_i) = \sum_{s \ni i} P(s) \text{ et } \pi_{ij} = E(I_i I_j) = \sum_{s \ni i, j} P(s), \quad i, j \in U,$$

où $E(\cdot)$ désigne l'espérance sous la loi P .

Un plan de Poisson est un plan de sondage tel que les variables aléatoires $(I_i)_{i \in U}$ sont indépendantes : la sélection d'un groupe d'unités dans l'échantillon est indépendante de la sélection d'un autre groupe disjoint d'unités dans l'échantillon. On obtient un plan de Poisson à probabilités d'inclusion $(\pi_i)_{i \in U}$ en réalisant N tirages de Bernoulli indépendants de paramètres π_i , $i = 1, \dots, N$, pour décider de la sélection ou non de chaque unité dans l'échantillon. Les échantillons d'unités primaires des ESS2012, 2014, 2016 et 2018 ont été sélectionnés selon des plans de Poisson.

Soit $(y_i)_{i \in U}$ des valeurs associées aux unités $i \in U$ de la population, et s un échantillon pour lequel les valeurs y_i ont été observées. L'estimateur de Horvitz & Thompson (1952) du total

$T = \sum_{i \in U} y_i$ est défini par :

$$\hat{T}_{HT} = \sum_{i \in U} d_i y_i I_i(s),$$

où $d_i = 1/\pi_i$ si $i \in s$. Si les π_i sont tous strictement positifs, alors l'estimateur $\hat{T}_{HT}$ est sans biais sous le plan : $E(\hat{T}_{HT}) = T$. Sa variance sous la loi P vaut

$$\text{var}(\hat{T}_{HT}) = \sum_{i,j \in U} d_i y_i d_j y_j (\pi_{ij} - \pi_i \pi_j).$$

Elle peut être estimée sans biais lorsque les π_{ij} sont strictement positifs par l'estimateur

$$\widehat{\text{var}}(\hat{T}_{HT}) = \sum_{i,j \in s} d_i y_i d_j y_j (\pi_{ij} - \pi_i \pi_j) / \pi_{ij}.$$

Pour un plan de Poisson, l'indépendance des sélections implique que $\pi_{ij} = \pi_i \pi_j$ si $i \neq j$. La variance de l'estimateur du total vaut alors

$$\text{var}(\hat{T}_{HT}) = \sum_{i \in U} d_i^2 y_i^2 \pi_i (1 - \pi_i) = \sum_{i \in U} y_i^2 (d_i - 1).$$

Elle peut être estimée sans biais par

$$\widehat{\text{var}}(\hat{T}_{HT}) = \sum_{i \in s} y_i^2 d_i (d_i - 1).$$

3.1.2 Variance dans un plan à deux degrés

Dans un plan à deux degrés, avec un échantillon s d'unités primaires sélectionné dans une population $U = \{1, \dots, N\}$, on peut définir l'estimateur par dilatation

$$\hat{T} = \sum_{i \in s} \frac{\hat{T}_i}{\pi_i},$$

où $\hat{T}_i$ est un estimateur du total de la variable d'intérêt dans l'unité primaire i et π_i est la probabilité d'observer i .

Sa variance peut s'écrire comme

$$\text{var}(\hat{T}) = \sum_{i=1}^N \sum_{j=1}^N \frac{T_i T_j}{\pi_i \pi_j} \Delta_{ij} + \sum_{i=1}^N \frac{\text{var}_2(\hat{T}_i)}{\pi_i}, \quad (1)$$

où $\Delta_{ij} = \pi_{ij} - \pi_i \pi_j$ et $\text{var}_2(\hat{T}_i)$ est la variance de $\hat{T}_i$ conditionnellement à la sélection de i dans l'échantillon de premier degré (voir par exemple Tillé, 2019).

Si $\widehat{\text{var}}_2(\hat{T}_i)$ est un estimateur sans biais de $\text{var}_2(\hat{T}_i)$ pour tout i , alors on peut estimer la variance (1), sans biais si les probabilités d'inclusion jointes π_{ij} sont strictement positives, par

$$\widehat{\text{var}}(\hat{T}) = \sum_{i \in s} \sum_{j \in s} \frac{\hat{T}_i \hat{T}_j}{\pi_i \pi_j} \frac{\Delta_{ij}}{\pi_{ij}} + \sum_{i \in s} \frac{\widehat{\text{var}}_2(\hat{T}_i)}{\pi_i}.$$

Dans le cas de l'ESS, le plan de premier degré est un plan de Poisson tandis que le plan de second degré est considéré comme un sondage aléatoire simple de taille fixe dans chaque unité

primaire. Ce dernier point constitue nécessairement une approximation car on ne contrôle en réalité pas ce mécanisme de sélection. Toutefois, les instructions fournies aux entreprises ainsi que la méthode de pondération qu'il est prévu d'utiliser justifient ce choix.

Soit $\mu_i > 0$ le nombre total d'unités secondaires observables dans l'unité primaire i , $m_i > 0$ le nombre d'unités secondaires que l'unité primaire i a sélectionnées dans l'échantillon final, $y_k, k \in V$ une variable d'intérêt observable sur la population V des unités secondaires, S_i^2 la dispersion des y_k dans l'unité i lorsque $\mu_i > 1$,

$$S_i^2 = \frac{1}{\mu_i - 1} \sum_{k \in i} (y_k - \bar{y}_i)^2, \text{ où } \bar{y}_i = \frac{1}{\mu_i} \sum_{k \in i} y_k,$$

et $S_i^2 = 0$ lorsque $\mu_i = 1$. On a alors

$$\text{var}_2(\hat{T}_i) = \mu_i^2 \left(1 - \frac{m_i}{\mu_i}\right) \frac{S_i^2}{m_i},$$

si $\hat{T}_i$ est l'estimateur de Horvitz-Thompson du total des y_k dans i :

$$\hat{T}_i = \frac{\mu_i}{m_i} \sum_{k \in s_i} y_k,$$

où s_i est l'échantillon d'unités secondaires de i dans l'échantillon final.

La variance de l'estimateur par dilatation d'un total dans l'ESS (qui est aussi l'estimateur de Horvitz-Thompson) peut donc s'écrire comme suit :

$$\text{var}(\hat{T}) = \sum_{i=1}^N T_i^2 (d_i - 1) + \sum_{i=1}^N d_i \mu_i^2 \left(1 - \frac{m_i}{\mu_i}\right) \frac{S_i^2}{m_i}.$$

Lorsque $m_i > 1$, S_i^2 peut être estimé sans biais par

$$s_i^2 = \frac{1}{m_i - 1} \sum_{k \in s_i} (y_k - \hat{\bar{y}}_i)^2, \text{ où } \hat{\bar{y}}_i = \frac{1}{m_i} \sum_{k \in s_i} y_k.$$

Le plan spécifié pour l'ESS devrait assurer que $m_i = 1$ uniquement lorsque $\mu_i = 1$. Dans ce cas il n'y a pas de problème d'estimation de S_i^2 et de $\text{var}_2(\hat{T}_i)$. On peut ainsi définir l'estimateur de variance sans biais sous le plan :

$$\widehat{\text{var}}(\hat{T}) = \sum_{i \in s} \hat{T}_i^2 d_i (d_i - 1) + \sum_{i \in s} d_i \mu_i^2 \left(1 - \frac{m_i}{\mu_i}\right) \frac{s_i^2}{m_i}.$$

3.1.3 Non-réponse

Les unités primaires sélectionnées dans l'échantillon de premier degré ne répondent pas toutes. On modélise cette non-réponse par un processus aléatoire : chaque unité i choisit de répondre ou non avec une probabilité p_i et les décisions sont prises de manière indépendante entre les unités primaires. Il s'agit d'un échantillonnage de Poisson au sein de l'échantillon sélectionné. Les probabilités p_i sont inconnues. Toutefois, on dispose de prédictions $\hat{p}_i$. Celles-ci sont par exemple obtenues à l'aide d'une régression logistique en utilisant des caractéristiques connues pour toutes les entreprises de l'échantillon, répondantes ou non répondantes. Une autre possibilité est de constituer des groupes de réponse homogènes à l'aide de techniques de segmentation pour estimer les probabilités de réponse.

Si les p_i étaient connues, on pourrait estimer un total en utilisant l'estimateur de Horvitz-Thompson avec les probabilités $\pi_i p_i$ d'observer les unités $i \in U$:

$$\hat{T}_{NR} = \sum_{i \in s} \frac{\hat{T}_i}{\pi_i p_i}.$$

La variance de $\hat{T}_{NR}$ vaut

$$\text{var}(\hat{T}_{NR}) = \sum_{i=1}^N \frac{1 - \pi_i p_i}{\pi_i p_i} T_i^2 + \sum_{i=1}^N \frac{1}{\pi_i p_i} \mu_i^2 \left(1 - \frac{m_i}{\mu_i}\right) \frac{S_i^2}{m_i}.$$

En pratique, les p_i sont inconnues et on utilise les prédictions $\hat{p}_i$ pour définir l'estimateur par dilatation :

$$\hat{T}_{nr} = \sum_{i \in s} \frac{\hat{T}_i}{\pi_i \hat{p}_i},$$

qui est bien défini mais n'est pas exactement sans biais. Selon [Kim & Kim \(2007\)](#), la variance de cet estimateur est, en général et sous conditions sur les propriétés des prédictions $\hat{p}_i$, inférieure à la variance de $\hat{T}_{NR}$. Ainsi, $\text{var}(\hat{T}_{NR})$ est habituellement un majorant de $\text{var}(\hat{T}_{nr})$ et l'estimateur

$$\widehat{\text{var}}(\hat{T}_{nr}) = \sum_{i \in s} \frac{1}{\pi_i \hat{p}_i} \left(\frac{1}{\pi_i \hat{p}_i} - 1 \right) \hat{T}_i^2 + \sum_{i \in s} \frac{1}{\pi_i \hat{p}_i} \mu_i^2 \left(1 - \frac{m_i}{\mu_i}\right) \frac{s_i^2}{m_i},$$

est a priori un estimateur conservateur de la variance de $\hat{T}_{nr}$.

3.1.4 Calage

Les poids d'extrapolation finaux w_i sont obtenus par calage des poids $1/\pi_i \hat{p}_i$ sur les totaux de certaines variables auxiliaires x_i observées pour chaque unité de l'échantillon (voir [Deville & Särndal, 1992](#)). Ce traitement nécessite de connaître les totaux des x_i dans la population. Dans l'ESS, seul un calage des poids des unités primaires est effectué.

Soit

$$\hat{T} = \sum_{i \in s} w_i \hat{T}_i,$$

un estimateur calé, où

$$\hat{T}_i = \frac{\mu_i}{m_i} \sum_{k \in s_i} y_k.$$

Une approximation de la variance de $\hat{T}$ est obtenue en remplaçant, dans les formules de variance précédentes, les totaux T_i par les résidus de régression E_i de ces totaux sur les variables de calage (voir [Deville & Särndal, 1992](#)). On a alors la variance approchée définie par :

$$\text{var}(\hat{T}) \approx \sum_{i=1}^N \left(\frac{1}{\pi_i p_i} - 1 \right) E_i^2 + \sum_{i=1}^N \frac{1}{\pi_i p_i} \mu_i^2 \left(1 - \frac{m_i}{\mu_i}\right) \frac{S_i^2}{m_i}. \quad (2)$$

Un estimateur usuellement conservateur de cette variance approchée est donné par :

$$\widehat{\text{var}}(\hat{T}) = \sum_{i \in s} \frac{1}{\pi_i \hat{p}_i} \left(\frac{1}{\pi_i \hat{p}_i} - 1 \right) \hat{E}_i^2 + \sum_{i \in s} \frac{1}{\pi_i \hat{p}_i} \mu_i^2 \left(1 - \frac{m_i}{\mu_i}\right) \frac{s_i^2}{m_i},$$

où les $\hat{E}_i$ sont les résidus de la régression pondérée des $\hat{T}_i$ sur les variables de calage, avec les poids $1/\pi_i \hat{p}_i$. Les poids calés peuvent également être utilisés pour définir l'estimateur

$$\widehat{\text{var}}(\hat{T}) = \sum_{i \in s} w_i \left(\frac{1}{\pi_i \hat{p}_i} - 1 \right) \hat{E}_i^2 + \sum_{i \in s} w_i \mu_i^2 \left(1 - \frac{m_i}{\mu_i}\right) \frac{s_i^2}{m_i}. \quad (3)$$

3.1.5 Linéarisation d'un estimateur complexe

Dans le cas d'estimateurs d'une fonction non linéaire des valeurs observées, telle la médiane de salaires y_k , des techniques de linéarisation permettent d'obtenir une approximation de la variance (voir par exemple [Deville, 1999](#); [Demnati & Rao, 2004](#); [Osier, 2009](#); [Graf & Tillé, 2014](#); [Vallée & Tillé, 2019](#)). À l'instar de la solution proposée dans la section 3.1.4 pour le calage, il s'agit de remplacer la variable d'intérêt y_k par une variable dite linéarisée z_k dont l'estimateur du total a approximativement la même variance que l'estimateur du paramètre considéré. Une variable linéarisée différente doit être calculée pour chaque paramètre d'intérêt. Il existe de multiples manières de calculer des variables linéarisées, qui permettent de proposer différentes approximations de variance.

Pour un estimateur de quantile, on peut citer par exemple [Osier \(2009\)](#) et [Vallée & Tillé \(2019\)](#) avec les résultats suivants :

1. Suivant [Deville \(1999\)](#) et [Osier \(2009\)](#),

$$z_k = -\frac{1}{Mf(q_\alpha)} [I(y_k \leq q_\alpha) - \alpha],$$

où q_α est le quantile d'ordre α dans la population V , M la taille de la population, I la fonction indicatrice, et f est une « estimation de densité » de y par la méthode des noyaux, c'est à dire

$$f(y) = \frac{1}{hM} \sum_{k \in V} \phi\left(\frac{y - y_k}{h}\right), \quad (4)$$

où ϕ est une fonction de noyau (par exemple la densité de la loi normale centrée réduite, aussi appelée noyau gaussien), et la fenêtre h est un nombre positif qui permet de régler le degré de lissage de l'estimation. Une petite valeur de h produit une estimation de densité qui lisse peu les données et une grande valeur de h produit à l'opposé une densité estimée très lisse. [Osier \(2009\)](#) propose d'utiliser le noyau gaussien et le choix de fenêtre classique $h = (4/3)^{1/5} M^{-1/5} \sigma$, où σ est l'écart-type de la variable d'intérêt.

2. Suivant [Vallée & Tillé \(2019\)](#),

$$z_k = -\frac{1}{Mf(q_\alpha)} \left[\Phi\left(\frac{q_\alpha - y_k}{h}\right) - F(q_\alpha) \right],$$

où Φ et F sont des primitives de ϕ et f . Les auteurs ne donnent pas de recommandations quant au choix du noyau et de la fenêtre.

Pour l'ESS, une variance approchée d'un estimateur non-linéaire est encore donnée par l'équation (2), où les E_i^2 sont les carrés des résidus de régression du total de la variable linéarisée z_k dans l'unité primaire i sur les variables de calage x_i et les S_i^2 sont les variances corrigées des z_k dans l'unité i .

Les z_k dépendent de valeurs inobservées de la population et font elles-mêmes, après enquête, l'objet d'une estimation à partir des données relevées :

$$\hat{z}_k = -\frac{1}{\widehat{M}\widehat{f}(\hat{q}_\alpha)} \left[\Phi\left(\frac{\hat{q}_\alpha - y_k}{h}\right) - \widehat{F}(\hat{q}_\alpha) \right],$$

où $\widehat{M}$, $\widehat{f}$ et $\widehat{F}$ sont obtenus comme sommes pondérées sur l'échantillon.

Différents choix de la méthode de linéarisation, ainsi que différents choix de la méthode d'estimation de densité, de la fonction de noyau et de la fenêtre conduisent à des approximations de variance différentes et aussi à des estimateurs du quantile différents. En effet, le quantile estimé $\hat{q}_\alpha$ est la, ou une, solution de l'équation $\widehat{F}(\hat{q}_\alpha) = \alpha$.

L'estimateur de variance (3) est obtenu en calculant, pour chaque unité primaire i , le total estimé $\hat{E}_i$ des résidus $\hat{e}_k$, $k \in i$ de la régression des $\hat{z}_k$ sur les variables de calage :

$$\hat{E}_i = \frac{\mu_i}{m_i} \sum_{k \in s_i} \hat{e}_k.$$

Le calage est réalisé au niveau des poids des unités primaires seulement. On peut montrer que les $\hat{E}_i$ sont aussi les résidus de régression des totaux estimés $\hat{Z}_i$

$$\hat{Z}_i = \frac{\mu_i}{m_i} \sum_{k \in s_i} \hat{z}_k,$$

sur les variables de calage. Les poids utilisés pour calculer les résidus dans ces deux régressions dépendent de la méthode de calage utilisée lors du calcul de la pondération.

Les s_i^2 sont les dispersions estimées des $\hat{z}_k$ dans les unités i :

$$s_i^2 = \frac{1}{m_i - 1} \sum_{k \in s_i} \left(\hat{z}_k - \frac{1}{m_i} \sum_{k \in s_i} \hat{z}_k \right)^2,$$

lorsque $m_i > 1$. S'il arrive lors de l'enquête que $m_i = 1$ alors que $\mu_i > 1$, il est par exemple possible de calculer un estimateur de variance en définissant, pour chaque unité primaire i qui se trouve dans ce cas, une autre unité primaire j qui servira de donneur pour imputer les valeurs s_i^2 .

3.1.6 Mise en évidence de l'allocation de l'ESS dans le calcul de variance

L'équation (2) donne la variance approchée de l'estimateur d'un total. Elle fait apparaître les probabilités d'inclusion utilisées pour sélectionner l'échantillon d'unités primaires. Soit $h = 1, \dots, H$ les strates de tirage définies dans la Section 2.4, par exemple pour le secteur privé, N_h le nombre d'entreprises dans la strate h et

$$n_h = \sum_{i \in h} \pi_i = N_h \pi_i \text{ si } i \in h.$$

Le paramètre n_h est la taille espérée de l'échantillon sélectionné dans la strate h sous le plan de sondage. Ce n'est pas nécessairement un nombre entier.

Dans le cas de l'ESS, on peut alors réécrire :

$$\text{var}(\hat{T}) \approx \sum_{h=1}^H \frac{d_h^2}{n_h} - \sum_{h=1}^H f_h, \quad (5)$$

où

$$d_h^2 = N_h \sum_{i \in h} \frac{1}{p_i} \left[E_i^2 + \mu_i^2 \left(1 - \frac{m_i}{\mu_i} \right) \frac{S_i^2}{m_i} \right] \text{ et } f_h = \sum_{i \in h} E_i^2.$$

Les paramètres d_h^2 et f_h doivent être calculés pour chaque statistique d'intérêt et chaque domaine d'estimation. Ils dépendent des variables de calage x_k qui sont connues mais aussi des valeurs inconnues des probabilités de réponse p_i et des variables y_k ainsi que des linéarisées z_k .

3.2 Hypothèses et données utilisées pour le calcul des variances

Les paramètres d_h^2 et f_h de la formule de variance (5) ont été calculés de manière approchée en faisant les hypothèses et approximations suivantes.

3.2.1 Revenus AVS

Les revenus moyens AVS sont utilisés dans les calculs en remplacement des salaires qui sont inconnus. Les revenus moyens AVS sont, par définition, égaux au revenu total annuel enregistré divisé par le nombre de mois d'activité indiqué dans les fichiers AVS. Pour les calculs, les dernières données disponibles ont été utilisées. Elles concernent les revenus de l'année 2016. La taille de l'entreprise μ_i est égale au nombre de revenus AVS disponibles.

Les revenus AVS sont la seule information quasiment exhaustive disponibles sur les revenus des employés des entreprises. Les salaires sur lesquels sont calculées les médianes dans l'ESS sont les salaires standardisés : il s'agit du salaire horaire effectif rapporté à 40h de travail hebdomadaire. Une comparaison entre les revenus moyens AVS et les salaires observés dans l'ESS2016 montre une bonne corrélation entre les deux données. Le coefficient de corrélation de Pearson entre le revenu moyen AVS pour 2016 et le salaire standardisé, calculé sur les données appariées avec l'ESS2016, vaut 0.84. Le coefficient de Spearman, qui n'est pas sensible à la présence de grandes valeurs, vaut 0.78. La corrélation est encore beaucoup plus forte entre le revenu moyen AVS et le salaire mensuel (non-standardisé) déclaré dans l'ESS2016, avec un coefficient de corrélation de Pearson de 0.92 et un coefficient de Spearman de 0.95.

Comme évoqué dans la section 2.3, les médianes calculées dans l'ESS portent sur les salaires mensuels bruts standardisés des heures contractuelles ou travaillées. Le temps de travail hebdomadaire ou mensuel n'étant pas renseigné dans les données AVS, les calculs d'allocation ont été effectués en se basant sur les médianes des revenus AVS mensuels moyens.

Pour certains groupes d'entreprises, les données AVS ne contiennent qu'un identifiant d'entreprise pour tous les employés du groupe. Dans ce cas, les revenus AVS ont été répartis aléatoirement entre les entreprises du groupe en respectant les proportions données par les nombres d'employés respectifs selon le cadre de sondage.

Enfin, certaines entreprises du cadre de sondage de l'ESS2018 n'apparaissent pas dans les données AVS. Les calculs sont réalisés en excluant ces entreprises. Elles recevront, pour le tirage de l'échantillon, la probabilité de sélection calculée pour la strate à laquelle elles appartiennent.

3.2.2 Proportion de salaires livrés

Les entreprises et administrations répondantes livrent effectivement la proportion de salaires demandée. Cette hypothèse est certainement conservatrice car il arrive fréquemment qu'une grande entreprise livre les données pour l'ensemble de ses employés plutôt que d'en faire une sélection. En 2016, 35% des grandes entreprises répondantes ont livré plus de 90% de leurs salaires. Au total, environ 1.8 million de salaires ont été livrés alors que le strict respect des consignes d'enquête n'imposait d'en livrer que 945'000.

3.2.3 Probabilités de réponse

Les probabilités de réponse dépendent de la taille de l'entreprise ou administration. Dans le cadre des administrations publiques, les probabilités de réponse utilisées étaient de 1 pour la Confédération et les cantons et 0.8 pour les districts et les communes. Dans le cadre du secteur privé et des autres entreprises publiques, les probabilités utilisées étaient de 0.8 pour les grandes entreprises et celles du profiling, 0.7 pour les entreprises de taille moyenne et 0.6 pour les petites entreprises. Ces taux correspondent à des valeurs moyennes observées dans les enquêtes précédentes.

3.2.4 Linéarisation et variance approchée

Des simulations à l'aide des données AVS ont permis de valider l'approximation de variance (2) et son estimateur (3). Différentes linéarisées pour l'estimation de la médiane ont été considérées. Le tableau 4 donne le résultat de ces simulations pour :

- (m0) la méthode de Deville (1999) et Osier (2009) avec le noyau gaussien,
- (m1) la méthode de Vallée & Tillé (2019) avec le noyau gaussien,
- (m2) la méthode de Vallée & Tillé (2019) avec le noyau dont la fonction de répartition est $x \mapsto 1/[1 + \exp(-x)]$.

Ce tableau est le résultat de 10'000 répétitions de la sélection d'un échantillon de 5000 « revenus AVS » par sondage aléatoire simple. Les échantillons sélectionnés sont calés au niveau des entreprises sur le nombre d'entreprises et le nombre d'employés par classe de taille et au niveau des employés sur des classes de revenus. Les colonnes du tableau 4 sont respectivement la méthode utilisée, la médiane calculée des revenus AVS annuels, la moyenne empirique des médianes estimées $\hat{q}_{.5}$ sur les 10'000 échantillons, la racine S de la variance approchée (2), la moyenne des écarts-types $\hat{S}$ estimés à l'aide de l'estimateur (3), l'écart-type empirique des médianes estimées sur les 10'000 échantillons et le « taux de couverture » : proportion des intervalles de confiance approchés $(\hat{q}_{.5} \pm 1.96 \cdot \hat{S})$ qui contiennent la médiane.

TABLE 4 Résultats de 10'000 simulations, échantillon simple de taille 5000, poids calés.

Méthode	Médiane	$E(\hat{q}_{.5})$	S calculé	$E(\hat{S})$	E.T. simul.	Tx. Couv.
m0	55'948	55'974	350.92	351.08	347.80	0.9473
m1	55'958	55'963	314.88	315.60	317.72	0.9470
m2	55'956	55'959	315.86	316.50	315.40	0.9483

Les médianes calculées diffèrent selon la méthode choisie : pour les méthodes (m1) et (m2), elles correspondent aux médianes des distributions estimées par l'estimateur à noyau. Les résultats produits par l'OFS utilisent la méthode (m0). En effet, il s'agit des médianes calculées par défaut par la procédure FREQ du logiciel statistique SAS.

Les résultats du tableau 4 montrent une bonne concordance entre les calculs et estimations de variance et les résultats empiriques. Le taux de couverture des intervalles de confiance simulé est proche du niveau choisi à 95%. On observe une différence entre les différentes méthodes d'estimation, tant sur la valeur de la médiane estimée que sur la précision obtenue. Toutefois, des simulations supplémentaires laissent penser que les différences de précision entre les méthodes s'estompent avec des échantillons de plus grandes tailles.

Un enseignement important des simulations porte sur le choix du paramètre de lissage h de l'estimateur à noyau (4). En effet, la formule standard préconisée par Osier (2009) fait intervenir le paramètre de dispersion σ qui est sensible aux valeurs extrêmes de la variable d'intérêt. Dans le cas de données de revenu, il faut s'attendre à observer de telles valeurs. C'est effectivement le cas dans les données AVS.

La présence de grands revenus dans certains échantillons simulés entraînait, pour ceux là, une valeur élevée de h et un fort lissage de la distribution estimée. Cela conduisait pour ces échantillons à une mauvaise estimation de la variance qui était éloignée à la fois de la variance « théorique » et de la dispersion empirique observée sur les simulations.

Une solution classique à ce problème consiste à utiliser la définition suivante :

$$h = (4/3)^{1/5} M^{-1/5} \min[\sigma, (Q_3 - Q_1)/1.34],$$

où $\min[\sigma, (Q_3 - Q_1)/1.34]$ est le minimum entre σ et l'écart inter-quartiles de la variable d'intérêt divisé par 1.34 (voir Silverman, 1986, p.47). Les résultats du tableau 4 ont été obtenus avec cette

définition, qui a été retenue pour la suite des calculs.

3.3 Meilleure précision possible et objectifs retenus

Les meilleures précisions atteignables sous les hypothèses retenues ont été calculées en incluant toutes les entreprises dans l'échantillon. Les CV minimaux et les CV visés selon les règles définies dans la section 3 sont donnés dans le tableau 5. On y trouve les résultats pour l'estimation de la médiane par catégorie d'activité, par grande région et pour la Suisse, ainsi que pour l'estimation du total par section NOGA et par classe de taille. Les CV minimaux et les CV visés pour l'estimation de la médiane par croisement de la catégorie d'activité et de la grande région sont donnés dans le tableau 15 en annexe A.

Les CV visés dans les tableaux 5 et 15 doivent être considérés comme des maxima à ne pas dépasser et non des cibles à obtenir. Pour une large proportion des objectifs, les CV effectivement réalisés devraient être bien plus faibles. En effet, les objectifs sont interdépendants : obtenir un CV de 3% dans chaque croisement de catégorie d'activité et de grande région implique un CV inférieur à 3% pour la catégorie d'activité ou pour la grande région. De plus, les entreprises de plus de 50 employés sont entièrement recensées, ce qui peut aussi entraîner des CV inférieurs aux CV visés pour certaines statistiques d'intérêt.

3.4 Optimisation de la taille d'échantillon

L'objectif est d'obtenir un échantillon de taille minimale ou aussi petite que possible mais qui permet d'atteindre les précisions voulues. La taille espérée n de l'échantillon total en fonction des tailles espérées n_h de l'échantillon dans chaque strate, vaut :

$$n = \sum_{h=1}^H n_h.$$

D'après l'équation (5), les contraintes de précision peuvent s'écrire :

$$\sum_{h=1}^H \frac{d_{h,g}^2}{n_h} - \sum_{h=1}^H f_{h,g} \leq \alpha_g, \quad g = 1, \dots, G,$$

où

- G est le nombre de statistiques d'intérêt considérées : 14 pour l'échantillon des administrations publiques des communes, districts, cantons et de la Confédération, et 687 pour l'échantillon du secteur privé et reste du secteur public.
- $d_{h,g}^2$ et $f_{h,g}$ sont les coefficients de la formule de variance (5) calculés pour chacune de ces statistiques d'intérêt,
- α_g est la variance visée, égale au carré du résultat de la multiplication du coefficient de variation visé par le paramètre d'intérêt calculé sur les données AVS.

Par ailleurs, les n_h satisfont aux contraintes :

$$0 \leq n_h \leq N_H, \quad h = 1, \dots, H,$$

soit 68 contraintes correspondant aux 34 strates de tirage pour l'allocation de l'échantillon des administrations publiques, et 4100 contraintes correspondant aux 2050 strates de tirage pour l'allocation de l'échantillon du secteur privé et des entreprises publiques.

Plus précisément, le calcul d'allocation se décompose en une collection de problèmes d'optimisation convexe :

- le calcul du plan de l'échantillon OFS des entreprises publiques et privées, avec au total 996 paramètres n_h (croisement des grandes régions, catégories NOGA et classes de taille) et 2334 contraintes,
- le calcul du plan de l'échantillon des administrations publiques, avec 34 paramètres n_h et 82 contraintes,
- les 9 calculs de densification pour les cantons et communes qui ont décidé de densifier, avec un nombre de paramètres et de contraintes qui varie selon le canton ou la commune (environ 130 paramètres n_h et 300 contraintes pour chaque zone de densification).

Le premier de ces problèmes est celui qui détermine essentiellement la composition de l'échantillon.

3.4.1 Optimisation pour une statistique d'intérêt

Le problème de minimisation de la taille d'échantillon peut être résolu exactement lorsqu'il y a une seule contrainte de précision :

$$\sum_{h=1}^H \frac{d_h^2}{n_h} - \sum_{h=1}^H f_h \leq \alpha,$$

et des tailles d'échantillon minimales et maximales dans les strates

$$0 \leq a_h \leq n_h \leq b_h \leq N_h.$$

Schématiquement, la solution est obtenue en déterminant l'ordre dans lequel les contraintes $a_h \leq n_h$ sont dé-saturées et les contraintes $n_h \leq b_h$ sont saturées pour obtenir l'allocation de variance minimale, lorsque la taille d'échantillon varie de $n_{min} = \sum_{h=1}^H a_h$ à $n_{max} = \sum_{h=1}^H b_h$. On peut montrer que cet ordre dépend uniquement des valeurs de a_h , b_h et d_h^2 , $h = 1, \dots, H$.

Pour chaque taille d'échantillon comprise entre ces deux bornes, on détermine l'allocation qui fournit la meilleure précision et la variance correspondante. Si la variance cible à ne pas dépasser α est comprise entre la variance optimale obtenue avec la taille n_{min} et celle obtenue avec la taille n_{max} , on trouve alors une allocation $n_\alpha = (n_{1,\alpha}, \dots, n_{H,\alpha})$ qui permet d'obtenir exactement la variance α . On peut de plus montrer que chaque $n_{h,\alpha}$ est une fonction décroissante de α .

3.4.2 Optimisation pour plusieurs statistiques d'intérêt

La solution retenue dans le cas de l'ESS2018, avec de nombreuses contraintes de précision, consiste à traiter les contraintes séquentiellement : l'allocation optimale est cherchée pour une première contrainte, puis une sur-allocation est calculée pour la deuxième contrainte, et le processus est poursuivi jusqu'à épuisement des contraintes de précision. Toutefois, cette démarche appliquée telle quelle conduit à obtenir des variances v_g , $g = 1, \dots, G$, inférieures ou égales à celles visées pour chacune des statistiques d'intérêt :

$$v_g \leq \alpha_g, \quad g = 1, \dots, G.$$

Il en résulte une taille d'échantillon plus grande que strictement nécessaire pour satisfaire chacune des contraintes. L'ordre dans lequel les contraintes de précision sont considérées (la numérotation $g = 1, \dots, G$) a une influence sur le résultat.

Une première étape d'optimisation de la procédure consiste à chercher un ordre qui conduit à une taille d'échantillon aussi réduite que possible. Une exploration exhaustive est impossible dans le cas des 687 objectifs concernant les entreprises publiques et privées. En pratique, ces 687 objectifs ont été divisés en blocs qui reflètent les objectifs 1, 2, 3 et 4 de la section 3. Par exemple, l'objectif 1 est divisé en trois blocs :

1. estimation des médianes par croisement de la catégorie d'activité et de la grande région,
2. estimation des médianes par catégorie d'activité,
3. estimation des médianes par grande région.

On obtient finalement six blocs pour les objectifs 1 et 2 propres à l'OFS, et trois blocs pour les objectifs 3 et 4 propres aux zones de densification. Au sein d'un bloc, l'ordre des contraintes n'a que peu d'influence : les estimateurs portent sur des domaines disjoints de la population. Toutes les permutations des six blocs OFS puis des trois blocs pour l'échantillon densifié ont pu être considérées et l'ordre des blocs qui donne la plus petite taille d'échantillon a été retenu. Une seconde étape d'optimisation consiste à adapter les contraintes de variance entrées dans le calcul séquentiel de manière à obtenir des variances v_g aussi proches que possible des valeurs α_g , $g = 1, \dots, G$ spécifiées, sans les dépasser. Pour ce faire, des nouvelles contraintes initiales de variance ont été recherchées à l'aide d'une procédure itérative. Cette procédure est décrite en annexe B. Pour l'échantillon d'entreprises publiques et privées de l'ESS2018, des contraintes jugées satisfaisantes ont été trouvées au bout de quinze itérations. Pour l'allocation des échantillons densifiés, cinq itérations ont été suffisantes. Pour l'échantillon d'administrations publiques, cette étape d'optimisation n'a pas montré d'effet notable et n'a pas été utilisée.

3.4.3 Travaux ultérieurs

La procédure présentée ci-dessus et mise en œuvre pour l'ESS2018 ne permet généralement pas de trouver une solution optimale, c'est-à-dire une allocation de taille minimale qui satisfait les contraintes. L'utilisation d'algorithmes d'optimisation convexe existants permettrait de trouver un tel optimum. Elle peut être envisagée pour de futures enquêtes mais leur application risque d'être compliquée par le fait que les solutions de ces problèmes d'allocation d'échantillons ne sont en général pas à l'intérieur du domaine des paramètres $(n_h)_{h=1, \dots, H}$ admissibles. Cette possibilité a été poursuivie dans un travail ultérieur qui a permis de trouver l'allocation de taille minimale sous les contraintes données pour l'ESS2018. Le résultat, présenté en annexe C dans le tableau 17, est très proche de la solution utilisée en production.

TABLE 5 CV minimaux et CV visés pour l'estimation de la médiane par catégorie d'activité et par grande région, pour l'estimation du total par section NOGA et par classe de taille.

Médiane par		CV min.	CV visé	Total par		CV min.	CV visé
NOGA	05-09	0.8	3.0	Section	B	2.4	3.0
	10-11	0.9	3.0	NOGA	C	0.5	3.0
	12	9.2	10.3		D	1.6	3.0
	13-15	3.0	4.0		E	1.4	3.0
	16-18	0.3	3.0		F	0.3	3.0
	19-20	1.1	3.0		G	0.9	3.0
	21	1.6	3.0		H	1.8	3.0
	22-23	0.5	3.0		I	0.7	3.0
	24-25	0.4	3.0		J	1.4	3.0
	26	1.2	3.0		K	1.2	3.0
	27	2.1	3.0		L	1.4	3.0
	28	0.5	3.0		M	0.9	3.0
	29-30	1.6	3.0		N	1.2	3.0
	31-33	0.7	3.0		O	3.4	4.0
	35	1.5	3.0		P	2.1	3.0
	36-39	0.8	3.0		Q	0.7	3.0
	41-43	0.1	3.0		R	2.3	3.0
	45-46	0.5	3.0		S	1.0	3.0
	47	0.5	3.0				
	49-52	0.9	3.0				
	53	3.0	4.0				
	55-56	0.5	3.0				
	58-60	1.6	3.0				
	61	1.4	3.0				
	62-63	0.8	3.0				
	64, 66	1.6	3.0				
	65	2.2	3.0				
	68	2.6	4.0				
	69-71	0.4	3.0				
	72	1.5	3.0				
	73-75	2.8	4.0				
	77,79-82	3.7	5.0				
	78	1.6	3.0				
	84	2.7	4.0				
	85	1.8	3.0				
	86-88	0.6	3.0				
	90-93	4.1	5.0				
	94-95	2.7	4.0				
	96	0.7	3.0				
Grande région	Région lémanique	0.7	3.0	Taille	3 à 9	0.2	3.0
	Espace Mittelland	0.6	3.0		10 à 49	0.3	3.0
	Suisse du Nord-ouest	0.7	3.0		50 à 249	0.5	3.0
	Zurich	0.7	3.0		250 à 499	1.8	3.0
	Suisse orientale	0.6	3.0		500 à 999	2.8	3.0
	Suisse centrale	0.7	3.0		1'000 et +	1.3	3.0
	Tessin	0.7	3.0				
Suisse		0.15	-			0.29	-

4 Résultat des calculs d'allocation

L'allocation obtenue pour l'échantillon des administrations publiques prévoit la sélection de 323 unités en espérance, dont la Confédération, les 26 cantons et demi-cantons, et 296 communes. Avec cette taille d'échantillon, tous les objectifs de précision sont atteints. Le détail par grande région de cette allocation et les CV correspondants pour l'estimation de la médiane des salaires sont donnés dans le tableau 6.

TABLE 6 Taille de la population, de l'échantillon alloué, CV minimum pour l'estimation de la médiane par grande région, CV visé et CV obtenu après allocation.

Grande région	Nb. unit.	Ech. Total	CV min.	CV visé	CV alloc.
Région lémanique	483	56	0.6	3.0	2.4
Espace Mittelland	668	68	0.4	3.0	2.4
Suisse du Nord-ouest	290	45	0.8	3.0	2.8
Zurich	167	38	0.8	3.0	2.8
Suisse orientale	325	48	0.8	3.0	3.0
Suisse centrale	167	45	1.0	3.0	3.0
Tessin	116	22	1.2	3.0	3.0
Confédération	1	1	0.3	3.0	0.3
Total	2'217	323	0.23	-	0.70

Le résultat de l'allocation pour le cadre du secteur privé et des entreprises publiques est un plan de sondage avec un échantillon de 45'663 entreprises en espérance. Cette allocation se décompose en un échantillon de 39'163 entreprises pour les besoins de l'OFS et une densification de 6500 entreprises pour les cantons et communes qui ont décidé de densifier. Les résultats détaillés sont donnés dans le tableau 7.

TABLE 7 Taille d'échantillon allouée par grande région ou zone de densification.

Grande région	GRS	Nb. Ents.	Alloc. OFS	densif.	Total
Région lémanique	GE	11'489	2'205	553	2'758
	VD	15'134	2'735	653	3'388
	Reste (VS)	7'994	1'145	0	1'145
Espace Mittelland	FR	5'755	1'060	1'136	2'196
	NE	3'617	655	844	1'499
	JU	1'647	294	634	928
	Reste (BE, SO)	26'579	5'317	0	5'317
Suisse du Nord-ouest	AG, BL, BS	22'616	5'171	0	5'171
Zurich	Ville de Zurich	13'118	2'945	336	3'281
	Reste ZH	20'163	3'650	0	3'650
Suisse orientale	TG	5'678	1'155	982	2'137
	Reste (AR, AI, GL, GR, SG, SH)	21'365	4'418	0	4'418
Suisse centrale	LU	8'822	2'386	485	2'871
	ZG	5'012	1'418	876	2'294
	Reste (LU, NW, OW, SZ, UR)	6'938	1'712	0	1'712
Tessin	TI	10'817	2'898	0	2'898
Total		186'744	39'163	6'500	45'663

Le critère de précision déterminant pour l'allocation de l'échantillon OFS est l'estimation des salaires médians par croisement de la catégorie d'activité et de la grande région. On peut le

constater d'après le tableau 16 en annexe A : la plupart des CV obtenus après allocation correspondent aux objectifs fixés pour ces statistiques. Pour les densifications, le critère déterminant est également l'estimation des salaires médians par catégorie d'activité dans les zones de densification.

Les CV obtenus pour les autres statistiques sont quant à eux généralement inférieurs aux CV visés, ce qui veut dire que la précision attendue avec l'échantillon alloué est meilleure que l'objectif qui avait été fixé. Ceci est d'autant plus vrai que les densifications contribuent à améliorer encore cette précision. Les CV obtenus pour les différents objectifs sont donnés dans les tableaux 8, 9, 10 et 11. Une étude ultérieure a révélé une erreur dans la programmation de l'estimateur de précision concernant les médianes. Les CV obtenus pour l'échantillon total après correction sont donnés dans la colonne « Corr. » des tableaux 8 et 9. Ils respectent eux aussi les objectifs de précision qui ont été définis.

TABLE 8 Taille de la population, de l'échantillon alloué selon le plan, CV visé pour l'estimation de la médiane par grande région, CV obtenu avec l'échantillon OFS alloué, l'échantillon total et CV corrigé.

Grande région	Nb. Ents.	Ech. Total	CV : visé	OFS	Total	Corr.
Région lémanique	34'617	7'290	3.0	1.0	0.9	1.0
Espace Mittelland	37'598	9'940	3.0	0.9	0.8	0.8
Suisse du Nord-ouest	22'616	5'171	3.0	1.1	1.1	1.1
Zurich	33'281	6'931	3.0	1.1	1.0	0.9
Suisse orientale	27'043	6'556	3.0	1.1	1.0	1.0
Suisse centrale	20'772	6'876	3.0	1.1	1.0	1.0
Tessin	10'817	2'898	3.0	1.4	1.4	1.4
Total	186'744	45'663	-	0.29	0.28	0.27

Remarque : du fait du calage, la précision d'un estimateur dans un domaine dépend également de l'échantillon qui est en dehors du domaine. En effet, les poids des observations dans un domaine sont parfois modifiés par le calage afin de respecter des totaux qui ne concernent pas uniquement ce domaine. Ainsi, le fait de densifier un échantillon en dehors d'un domaine donné peut aussi augmenter la précision des estimations dans ce domaine. Ceci est la raison pour laquelle les CV « OFS » et « Total » calculés dans le tableau 11 pour les entreprises de plus de 50 employés ne sont pas égaux entre eux ni égaux aux CV minimaux calculés dans le tableau 5, bien que ces domaines sont entièrement recensés.

TABLE 9 Taille de la population, de l'échantillon alloué selon le plan, CV visé pour l'estimation de la médiane par catégorie d'activité, CV obtenu avec l'échantillon OFS alloué, l'échantillon total et CV corrigé.

NOGA	Nb. Ents.	Ech. Total	CV : visé	OFS	Total	Corr.
05-09	212	107	3.0	1.4	1.3	1.2
10-11	2'887	798	3.0	1.4	1.4	1.3
12	10	9	10.3	10.0	10.0	8.9
13-15	602	298	4.0	3.7	3.6	3.3
16-18	4'122	563	3.0	1.2	1.1	1.1
19-20	420	176	3.0	2.1	1.9	2.0
21	184	145	3.0	1.9	1.8	1.8
22-23	1'145	373	3.0	1.0	1.0	1.0
24-25	4'174	695	3.0	1.0	0.9	0.9
26	1'328	429	3.0	1.7	1.6	1.5
27	539	190	3.0	3.0	2.9	2.8
28	1'419	422	3.0	1.2	1.1	1.1
29-30	247	93	3.0	2.0	2.0	1.8
31-33	2'758	664	3.0	1.3	1.2	1.1
35	468	180	3.0	2.0	2.0	2.0
36-39	923	377	3.0	1.3	1.2	1.2
41-43	21'501	1'353	3.0	0.9	0.9	0.8
45-46	17'177	2'615	3.0	1.1	1.0	0.9
47	14'915	1'992	3.0	1.0	0.9	1.0
49-52	4'586	876	3.0	1.9	1.7	1.9
53	239	113	4.0	3.6	3.5	3.8
55-56	17'249	4'122	3.0	1.3	1.1	1.1
58-60	1'243	601	3.0	3.0	3.0	3.0
61	229	117	3.0	1.9	1.8	2.1
62-63	5'102	1'496	3.0	1.5	1.4	1.2
64, 66	5'328	1'262	3.0	2.0	1.9	1.8
65	408	204	3.0	2.5	2.4	2.4
68	4'860	3'418	4.0	3.2	3.1	2.6
69-71	19'742	3'300	3.0	1.3	1.2	1.0
72	652	320	3.0	2.2	2.1	2.2
73-75	3'780	2'790	4.0	3.3	3.2	3.2
77,79-82	7'810	2'474	5.0	4.2	4.1	3.8
78	1'169	636	3.0	2.2	2.1	1.2
84	266	202	4.0	3.3	3.1	3.0
85	4'015	1'513	3.0	2.4	2.4	2.6
86-88	20'592	2'732	3.0	1.3	1.3	1.2
90-93	4'317	2'686	5.0	4.6	4.6	3.3
94-95	6'005	3'602	4.0	3.2	3.1	2.7
96	4'121	1'718	3.0	1.4	1.2	1.4
Total	186'744	45'663	-	0.29	0.28	0.27

TABLE 10 Taille de la population, de l'échantillon alloué, CV visé pour l'estimation du total par section NOGA, CV obtenu avec l'échantillon OFS et l'échantillon total.

Sect. NOGA	Nb. Ents.	Ech. Total	CV : visé	OFS	Total
B	212	107	3.0	3.0	2.8
C	19'835	4'858	3.0	0.7	0.7
D	468	180	3.0	2.0	2.0
E	923	377	3.0	2.1	2.1
F	21'501	1'353	3.0	1.5	1.4
G	32'092	4'607	3.0	1.3	1.2
H	4'825	988	3.0	2.5	2.4
I	17'249	4'122	3.0	1.8	1.7
J	6'574	2'214	3.0	2.0	1.9
K	5'736	1'467	3.0	1.8	1.8
L	4'860	3'418	3.0	1.7	1.6
M	24'174	6'410	3.0	1.7	1.5
N	8'979	3'110	3.0	1.5	1.4
O	266	202	4.0	3.9	3.7
P	4'015	1'513	3.0	2.3	2.3
Q	20'592	2'732	3.0	1.5	1.5
R	4'317	2'686	3.0	2.4	2.4
S	10'126	5'319	3.0	1.2	1.2
Total	186'744	45'663	-	0.43	0.41

TABLE 11 Taille de la population, de l'échantillon alloué, CV visé pour l'estimation du total par classe de taille, CV obtenu avec l'échantillon OFS et l'échantillon total.

Taille	Nb. Ents.	Ech. Total	CV : visé	OFS	Total
3 à 9	125'110	22'461	3.0	1.8	1.6
10 à 49	50'391	11'958	3.0	1.4	1.3
50 à 249	9'460	9'460	3.0	0.7	0.7
250 à 499	1'044	1'044	3.0	1.9	1.9
500 à 999	415	415	3.0	3.0	3.0
1'000 et +	324	324	3.0	1.6	1.5
Total	186'744	45'663	-	0.43	0.41

5 Sélection de l'échantillon

L'échantillon a été sélectionné de manière coordonnée avec les autres enquêtes de l'OFS, dont l'ESS2016, l'ESS2014 et l'ESS2012. L'objectif de cette coordination est d'éviter de solliciter de manière répétée les mêmes entreprises lorsque ce n'est pas indispensable à la qualité de la production statistique. Les entreprises de 50 employés ou plus sont entièrement recensées et ne bénéficient pas de cette coordination dans l'ESS.

La méthode de coordination utilisée depuis 2009 à l'OFS est décrite dans [Qualité \(2009\)](#) et [Graf & Qualité \(2014\)](#). Elle repose sur l'utilisation de nombres aléatoires permanents et de tirages d'échantillons selon un plan de Poisson. La taille des échantillons sélectionnés est aléatoire. Le résultat du tirage de l'échantillon d'administrations publiques est donné dans le tableau 12. Le résultat du tirage de l'échantillon dans le cadre du secteur privé et des entreprises publiques figure dans les tableaux 13 et 14.

TABLE 12 Échantillon alloué et échantillon sélectionné dans le cadre des administrations publiques.

		Nb. unités	Alloc.	Tiré
Grande région	Région lémanique	483	56	54
	Espace Mittelland	669	69	64
	Suisse du Nord-ouest	290	45	33
	Zurich	167	38	35
	Suisse orientale	325	48	42
	Suisse centrale	167	45	51
	Tessin	116	22	31
Taille	3 à 19	1'279	102	97
	20 à 49	429	84	75
	50 ou plus	435	62	64
	Profiling	47	47	47
	Cantons et Conf.	27	27	27
Total		2'217	323	310

TABLE 13 Échantillon alloué et échantillon sélectionné par catégorie d'activité NOGA dans le cadre du secteur privé et des entreprises publiques.

NOGA	Nb. unités	Alloc.	Tiré
05-09	212	107	110
10-11	2'887	798	798
12	10	9	10
13-15	602	298	309
16-18	4'122	563	584
19-20	420	176	171
21	184	145	150
22-23	1'145	373	374
24-25	4'174	695	665
26	1'328	429	418
27	539	190	195
28	1'419	422	422
29-30	247	93	80
31-33	2'758	664	690
35	468	180	179
36-39	923	377	389
41-43	21'501	1'353	1'344
45-46	17'177	2'615	2'624
47	14'915	1'992	2'000
49-52	4'586	876	901
53	239	113	109
55-56	17'249	4'122	4'095
58-60	1'243	601	603
61	229	117	123
62-63	5'102	1'496	1'467
64, 66	5'328	1'262	1'262
65	408	204	198
68	4'860	3'418	3'397
69-71	19'742	3'300	3'301
72	652	320	320
73-75	3'780	2'790	2'790
77,79-82	7'810	2'474	2'426
78	1'169	636	651
84	266	202	196
85	4'015	1'513	1'516
86-88	20'592	2'732	2'722
90-93	4'317	2'686	2'670
94-95	6'005	3'602	3'580
96	4'121	1'718	1'728
Total	186'744	45'663	45'567

TABLE 14 Échantillon alloué et échantillon sélectionné par grande région ou zone de densification dans le cadre du secteur privé et des entreprises publiques.

Grande région	GRS	Nb. unités	Alloc.	Tiré
Région lémanique	GE	11'489	2'758	2'712
	VD	15'134	3'388	3'270
	Reste (VS)	7'994	1'145	1'167
Espace Mittelland	FR	5'755	2'196	2'224
	NE	3'617	1'499	1'492
	JU	1'647	928	927
	Reste (BE, SO)	26'579	5'317	5'307
Suisse du Nord-ouest	AG, BL, BS	22'616	5'171	5'145
Zurich	Ville de Zurich	13'118	3'281	3'278
	Reste ZH	20'163	3'650	3'654
Suisse orientale	TG	5'678	2'137	2'110
	Reste (AR, AI, GL, GR, SG, SH)	21'365	4'418	4'482
Suisse centrale	LU	8'822	2'871	2'848
	ZG	5'012	2'294	2'316
	Reste (LU, NW, OW, SZ, UR)	6'938	1'712	1'716
Tessin	TI	10'817	2'898	2'919
Total		186'744	45'663	45'567

6 Possibilités d'amélioration et évolutions

Des gains en efficacité supplémentaires peuvent encore être envisagés lors d'une révision ultérieure. On peut citer :

1. *Le découpage de la population en strates et son optimisation à l'aide des données AVS* : pour l'ESS2018 les strates, dans lesquelles un taux de sélection uniforme est appliqué, sont constituées de la même manière que pour l'ESS2016. Il s'agit de croisements de grandes régions, de catégories d'activité et de classes de taille avec les mêmes définitions que dans l'ESS2016. En outre, à l'instar de l'ESS2016, un taux de sélection de 100% est utilisé pour les grandes entreprises et les entreprises du profiling. Il est certainement possible de revoir ce découpage et ces contraintes de manière à obtenir une plus grande efficacité. C'est toutefois un travail difficile qui demande normalement d'anticiper la possibilité que certaines entreprises sont mal classées (par exemple lorsque leur taille dans le registre ne correspond pas du tout aux réponses qu'elles fournissent). Une exploration des possibilités de stratification nécessite de calculer un grand nombre de paramètres. Souvent, la seule information disponible pour calculer ces paramètres provient d'autres enquêtes ou d'enquêtes passées. Les estimations de ces paramètres sont alors soumises à un aléa d'échantillonnage parfois important lorsque les strates envisagées sont de petite taille ou mal couvertes par les enquêtes à disposition. Les données exhaustives de l'AVS permettent de mener cette exploration sans problème d'estimation.
2. *La révision des probabilités de réponse utilisées pour les calculs* : en lien avec le point précédent, on s'aperçoit que les taux de réponse agrégés selon les classes de taille du tirage rendent mal compte de leur variabilité dans la population. Par exemple, les très grandes entreprises répondent quasiment toutes. L'utilisation des taux de réponse observés lors des enquêtes précédentes pourrait permettre de revoir ces classes de taille. Parallèlement, il serait possible d'imposer un taux de sondage minimal dans les strates pour se protéger contre la possibilité que des probabilités de réponse très faibles entraînent la création de poids d'extrapolation élevés.
3. *L'anticipation de la proportion de salaires renseignés par les entreprises répondantes* : certaines entreprises répondantes fournissent l'intégralité des données concernant les salaires qu'elles versent, ce qui conduit à une précision d'estimation supérieure à celle qui est visée. On pourrait envisager d'intégrer une anticipation de ce comportement dans les calculs d'allocation.
4. *L'utilisation des données AVS lors de la pondération et des estimations* : il est prévu d'utiliser les données sur les revenus, fournies par les caisses de compensation AVS, lors du calcul de la pondération de l'ESS2018 pour améliorer la précision des résultats. Ces gains espérés n'ont pas été intégrés dans les calculs. L'allocation a été réalisée en fonction des procédures de pondération et d'estimation utilisées dans l'ESS2016.
5. *L'utilisation d'autres sources d'information auxiliaire* : l'intégration d'information supplémentaire dans le plan de sondage et dans les calculs de pondération permet d'augmenter la précision des estimations, ou de réduire la taille des échantillons à précision égale. D'autres données administratives pourraient être mises à profit dans le futur.
6. *L'application de méthodes d'apprentissage automatisé pour la préparation statistique des données* : les données relevées sont soumises à un processus de préparation statistique. Une partie de ce processus vise à détecter des éventuelles réponses erronées. Ce traitement repose habituellement sur la postulation de modèles et la confrontation des réponses avec ce modèle. Avec la disponibilité de données auxiliaires riches sur les entreprises répondantes et leurs employés, des méthodes complexes de sélection de modèles peuvent trouver leur utilité pour améliorer la qualité des estimations.

7. *L'anticipation des effets de la procédure de préparation statistique des données* : lors de l'exploitation des données collectées, des données aberrantes sont détectées et parfois écartées des résultats. Lorsque cela se produit dans des entreprises sélectionnées avec des faibles taux de sondage et lorsqu'une proportion importante des salaires de ces entreprises est écartée, les salaires conservés peuvent recevoir des poids d'extrapolation élevés. Ces poids élevés peuvent entraîner une instabilité pour certains résultats qui n'a pas été anticipée dans les calculs d'allocation. Définir un taux de sondage minimum dans les strates permettrait en partie de se protéger contre ce risque.
8. *L'utilisation d'un algorithme d'optimisation convexe* : la méthode de recherche d'allocation utilisée ne permet en général pas de trouver la solution optimale lorsque plusieurs contraintes de précision doivent être respectées. Dans un travail ultérieur, l'allocation optimale a pu être trouvée pour l'échantillon du secteur privé et des entreprises publiques en utilisant le package CVXR (ver. 1.0-1) du logiciel statistique R. Le résultat est très proche de celui de la section 4, avec une taille d'échantillon de 45'380 unités primaires contre 45'663. Le résultat de l'allocation optimale est donné en annexe C dans le tableau 17. Il peut être comparé à l'allocation utilisée pour le tirage qui figure dans le tableau 7.

En outre, la structure du registre des entreprises et des établissements a changé en août 2019. Les communes, districts, cantons et la Confédération font depuis lors partie du même cadre de sondage que les entreprises publiques et privées. Les calculs d'allocation pour l'ESS2020 pourront se baser plus fidèlement sur les résultats principaux de l'ESS qui concernent l'ensemble de l'économie ou bien le secteur privé et le secteur public séparément.

7 Conclusion

La révision du plan de sondage de l'enquête suisse sur la structure des salaires a permis de fonder celui-ci sur des objectifs de précision définis pour une collection de statistiques d'intérêt. Le choix des précisions visées est basé sur des hypothèses de calcul qui peuvent être révisées après chaque enquête. Il est réalisé en fonction des précisions maximales que l'on peut atteindre sous ces hypothèses et en tenant compte de la taille d'échantillon nécessaire pour réaliser ces objectifs. L'échantillon total du secteur public et du secteur privé de l'ESS2018 compte 45'877 unités. Cela représente une diminution de 4691 unités, soit 9.3%, par rapport à celui de l'ESS2016 qui en comptait 50'568. Une démarche similaire à celle utilisée pour mener à bien cette révision pourrait être envisagée pour d'autres enquêtes de l'OFS. L'explicitation des statistiques d'intérêt et des objectifs de précision souhaités permet de dimensionner au mieux les échantillons de ces enquêtes. L'utilisation accrue des données auxiliaires disponibles permet d'améliorer l'efficacité des stratégies d'estimation et la qualité des résultats publiés.

Bibliographie

- DEMNATI, A. & RAO, J. N. K. (2004). Linearization variance estimators for survey data (with discussion). *Survey Methodology* **30**, 17–34.
- DEVILLE, J.-C. (1999). Variance estimation for complex statistics and estimators: linearization and residual techniques. *Survey Methodology* **25**, 193–204.
- DEVILLE, J.-C. & SÄRNDAL, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association* **87**, 376–382.
- GRAF, É. & QUALITÉ, L. (2014). Sondage dans des registres de population et de ménages en Suisse : coordination d'échantillons, pondération et imputation. *Journal de la Société Française de Statistique* **155**, 95–133.
- GRAF, É. & TILLÉ, Y. (2014). Variance estimation using linearization for poverty and social exclusion indicators. *Survey Methodology* **40**, 61–79.
- HORVITZ, D. G. & THOMPSON, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* **47**, 663–685.
- KIM, J. K. & KIM, J. J. (2007). Nonresponse weighting adjustment using estimated response probability. *Canadian Journal of Statistics* **35**, 501–514.
- OSIER, G. (2009). Variance estimation for complex indicators of poverty and inequality using linearization techniques. *Survey Research Methods* **3**, 167–195.
- QUALITÉ, L. (2009). *Unequal probability sampling and repeated surveys*. Thèse de doctorat, Université de Neuchâtel, Neuchâtel.
- SILVERMAN, B. W. (1986). *Density estimation for statistics and data analysis*. London: Chapman and Hall.
- TILLÉ, Y. (2019). *Théorie des sondages: échantillonnage et estimation en populations finies*. Paris: Dunod, 2nd ed.
- VALLÉE, A.-A. & TILLÉ, Y. (2019). Linearisation for variance estimation by means of sampling indicators: Application to non-response. *International Statistical Review* **87**, 347–367.

Annexes

A Coefficients de variation par catégorie d'activité et par grande région

TABLE 15 CV minimaux et CV visés pour l'estimation de la médiane par croisement de la catégorie d'activité et de la grande région.

Noga	GR1		GR2		GR3		GR4		GR5		GR6		GR7	
	min.	visé	min.	visé	min.	visé	min.	visé	min.	visé	min.	visé	min.	visé
05-09	3.0	4.0	1.4	3.0	2.7	4.0	2.1	3.0	1.7	3.0	1.7	3.0	1.3	3.0
10-11	7.2	8.3	1.6	3.0	1.5	3.0	2.1	3.0	2.3	3.0	3.4	4.0	2.3	3.0
12	27.8	28.8	11.4	12.4	6.4	7.5	0.0	3.0	-	-	5.3	6.4	14.0	15.0
13-15	2.7	4.0	2.7	4.0	3.1	4.0	3.2	4.0	2.5	4.0	5.4	6.4	3.3	4.0
16-18	0.8	3.0	0.8	3.0	0.9	3.0	1.1	3.0	0.8	3.0	1.1	3.0	1.4	3.0
19-20	1.6	3.0	2.2	3.0	2.1	3.0	2.1	3.0	2.0	3.0	2.4	3.0	7.3	8.4
21	2.0	3.0	1.8	3.0	1.5	3.0	6.9	7.9	4.8	5.4	10.6	11.6	2.1	3.0
22-23	1.6	3.0	0.9	3.0	1.4	3.0	1.5	3.0	1.0	3.0	0.9	3.0	2.5	3.0
24-25	2.1	3.0	0.6	3.0	0.7	3.0	1.4	3.0	0.6	3.0	1.0	3.0	1.6	3.0
26	2.1	3.0	1.2	3.0	3.8	5.0	1.9	3.0	2.6	4.0	2.6	4.0	1.3	3.0
27	3.4	4.0	1.5	3.0	2.7	4.0	2.0	3.0	2.9	4.0	2.7	4.0	4.5	5.0
28	2.8	4.0	0.6	3.0	0.9	3.0	1.0	3.0	1.0	3.0	1.6	3.0	3.6	5.0
29-30	3.2	4.0	3.0	4.0	3.8	5.0	3.8	5.0	2.0	3.0	2.3	3.0	4.2	5.0
31-33	1.2	3.0	1.2	3.0	2.3	3.0	2.5	3.0	1.2	3.0	2.1	3.0	3.6	5.0
35	3.7	5.0	2.2	3.0	3.9	5.0	3.7	5.0	2.0	3.0	3.9	5.0	1.9	3.0
36-39	1.6	3.0	1.8	3.0	2.6	4.0	1.7	3.0	1.8	3.0	3.4	4.0	3.9	5.0
41-43	0.3	3.0	0.3	3.0	0.4	3.0	0.5	3.0	0.3	3.0	0.5	3.0	0.5	3.0
45-46	2.3	3.0	0.8	3.0	0.9	3.0	1.1	3.0	1.0	3.0	2.3	3.0	1.4	3.0
47	1.0	3.0	1.2	3.0	1.3	3.0	0.6	3.0	2.7	4.0	2.8	4.0	1.7	3.0
49-52	2.1	3.0	1.4	3.0	1.1	3.0	1.9	3.0	2.7	4.0	1.6	3.0	2.6	4.0
53	18.5	19.6	0.7	3.0	22.4	23.4	2.3	3.0	8.4	9.5	21.4	22.4	20.7	21.7
55-56	0.9	3.0	1.2	3.0	4.6	5.2	0.9	3.0	1.4	3.0	1.3	3.0	1.5	3.0
58-60	7.8	8.8	6.3	7.4	7.8	8.8	1.4	3.0	8.6	9.7	7.9	9.0	6.4	7.4
61	5.9	7.0	1.4	3.0	5.1	6.2	3.8	5.0	7.5	8.6	9.0	10.1	7.2	8.2
62-63	1.5	3.0	1.6	3.0	1.7	3.0	2.0	3.0	2.1	3.0	1.8	3.0	11.2	12.3
64, 66	2.8	4.0	3.1	4.0	3.0	4.0	1.2	3.0	1.9	3.0	2.3	3.0	4.3	5.0
65	5.5	6.6	8.5	9.5	2.1	3.0	3.6	5.0	8.1	9.1	3.2	4.0	10.4	11.5
68	3.2	4.0	8.6	9.6	7.9	9.0	4.4	5.0	6.4	7.5	6.5	7.5	7.7	8.7
69-71	1.8	3.0	1.1	3.0	1.1	3.0	1.2	3.0	1.3	3.0	1.3	3.0	1.6	3.0
72	2.9	4.0	2.2	3.0	2.6	4.0	2.8	4.0	4.5	5.0	9.5	10.5	5.4	6.4
73-75	4.7	5.2	3.3	4.0	5.1	6.1	9.4	10.5	4.2	5.0	14.0	15.0	4.2	5.0
77,79-82	3.4	4.0	13.9	14.9	5.4	6.5	7.6	8.7	3.7	5.0	9.5	10.6	6.7	7.7
78	2.9	4.0	3.0	4.0	3.1	4.0	7.3	8.3	3.8	5.0	4.8	5.4	9.7	10.7
84	3.8	5.0	6.9	8.0	10.7	11.7	13.5	14.5	10.3	11.3	9.3	10.4	17.0	18.0
85	3.8	5.0	6.3	7.4	6.8	7.8	4.0	5.0	3.0	4.0	6.4	7.5	5.9	6.9
86-88	2.2	3.0	1.5	3.0	1.7	3.0	1.6	3.0	1.6	3.0	2.4	3.0	1.9	3.0
90-93	6.7	7.8	6.9	7.9	9.9	11.0	11.6	12.7	10.0	11.1	8.3	9.3	12.2	13.3
94-95	4.3	5.0	6.7	7.8	5.9	6.9	5.0	6.1	4.9	5.5	6.3	7.3	8.3	9.4
96	1.1	3.0	2.5	4.0	2.7	4.0	2.2	3.0	3.1	4.0	4.3	5.0	2.8	4.0
total	0.7	3.0	0.6	3.0	0.7	3.0	0.7	3.0	0.6	3.0	0.7	3.0	0.7	3.0

TABLE 16 CV visé et CV obtenu avec l'allocation OFS pour l'estimation de la médiane par catégorie d'activité et par grande région.

Noga	GR1		GR2		GR3		GR4		GR5		GR6		GR7	
	visé	ofs	visé	ofs	visé	ofs	visé	ofs	visé	ofs	visé	ofs	visé	ofs
05-09	4.0	4.0	3.0	2.0	4.0	4.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
10-11	8.3	8.3	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	4.0	4.0	3.0	3.0
12	28.8	28.8	12.4	11.8	7.5	7.5	3.0	0.0	-	-	6.4	5.4	15.0	14.0
13-15	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	6.4	6.4	4.0	4.0
16-18	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
19-20	3.0	2.6	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	8.4	8.4
21	3.0	2.1	3.0	2.0	3.0	1.6	7.9	7.9	5.4	4.9	11.6	10.7	3.0	2.1
22-23	3.0	3.0	3.0	2.1	3.0	2.8	3.0	3.0	3.0	2.0	3.0	3.0	3.0	3.0
24-25	3.0	3.0	3.0	1.7	3.0	2.6	3.0	3.0	3.0	2.2	3.0	3.0	3.0	3.0
26	3.0	2.9	3.0	1.7	5.0	4.8	3.0	2.8	4.0	4.0	4.0	4.0	3.0	2.0
27	4.0	4.0	3.0	3.0	4.0	4.0	3.0	3.0	4.0	3.7	4.0	4.0	5.0	5.0
28	4.0	4.0	3.0	1.6	3.0	3.0	3.0	2.5	3.0	3.0	3.0	3.0	5.0	5.0
29-30	4.0	4.0	4.0	4.0	5.0	5.0	5.0	5.0	3.0	2.3	3.0	3.0	5.0	5.0
31-33	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	5.0	5.0
35	5.0	5.0	3.0	3.0	5.0	5.0	5.0	4.8	3.0	3.0	5.0	5.0	3.0	3.0
36-39	3.0	3.0	3.0	3.0	4.0	4.0	3.0	3.0	3.0	3.0	4.0	4.0	5.0	5.0
41-43	3.0	2.1	3.0	1.6	3.0	2.2	3.0	1.7	3.0	2.7	3.0	2.7	3.0	3.0
45-46	3.0	3.0	3.0	2.5	3.0	2.9	3.0	2.4	3.0	3.0	3.0	3.0	3.0	3.0
47	3.0	3.0	3.0	3.0	3.0	3.0	3.0	1.2	4.0	4.0	4.0	4.0	3.0	3.0
49-52	3.0	3.0	3.0	2.9	3.0	3.0	3.0	2.7	4.0	4.0	3.0	3.0	4.0	4.0
53	19.6	19.6	3.0	0.9	23.4	23.4	3.0	3.0	9.5	9.5	22.4	22.4	21.7	21.7
55-56	3.0	2.8	3.0	3.0	5.2	5.2	3.0	2.6	3.0	3.0	3.0	3.0	3.0	3.0
58-60	8.8	8.8	7.4	7.4	8.8	8.8	3.0	2.8	9.7	9.7	9.0	9.0	7.4	7.4
61	7.0	7.0	3.0	1.9	6.2	6.2	5.0	5.0	8.6	8.6	10.1	10.1	8.2	8.2
62-63	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	12.3	12.3
64, 66	4.0	4.0	4.0	4.0	4.0	4.0	3.0	1.8	3.0	3.0	3.0	3.0	5.0	5.0
65	6.6	6.0	9.5	9.5	3.0	2.6	5.0	3.9	9.1	9.1	4.0	3.5	11.5	11.5
68	4.0	4.0	9.6	9.6	9.0	9.0	5.0	5.0	7.5	7.5	7.5	7.5	8.7	8.7
69-71	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
72	4.0	4.0	3.0	3.0	4.0	3.9	4.0	4.0	5.0	5.0	10.5	10.5	6.4	6.4
73-75	5.2	5.0	4.0	3.6	6.1	5.6	10.5	10.5	5.0	4.6	15.0	14.3	5.0	5.0
77,79-82	4.0	4.0	14.9	14.9	6.5	6.5	8.7	8.7	5.0	5.0	10.6	10.6	7.7	7.7
78	4.0	4.0	4.0	4.0	4.0	4.0	8.3	8.3	5.0	5.0	5.4	5.4	10.7	10.7
84	5.0	5.0	8.0	8.0	11.7	11.7	14.5	14.5	11.3	11.3	10.4	10.4	18.0	18.0
85	5.0	5.0	7.4	7.4	7.8	7.8	5.0	5.0	4.0	4.0	7.5	7.5	6.9	6.9
86-88	3.0	3.0	3.0	2.5	3.0	2.9	3.0	2.4	3.0	3.0	3.0	3.0	3.0	3.0
90-93	7.8	7.8	7.9	7.9	11.0	11.0	12.7	12.7	11.1	11.1	9.3	9.3	13.3	13.3
94-95	5.0	5.0	7.8	7.8	6.9	6.9	6.1	6.1	5.5	5.5	7.3	7.3	9.4	9.4
96	3.0	3.0	4.0	4.0	4.0	4.0	3.0	3.0	4.0	4.0	5.0	5.0	4.0	4.0
Total	3.0	1.0	3.0	0.9	3.0	1.1	3.0	1.1	3.0	1.1	3.0	1.1	3.0	1.4

B Procédure d'optimisation pour plusieurs contraintes de précision

Comme décrit dans la section 3.4.2, la démarche d'allocation séquentielle appliquée directement aux contraintes de variance α_g , $g = 1, \dots, G$ conduit à obtenir des variances v_g , $g = 1, \dots, G$, inférieures ou égales à celles visées pour chacune des statistiques d'intérêt :

$$v_g \leq \alpha_g, \quad g = 1, \dots, G.$$

Il en résulte une taille d'échantillon plus grande que strictement nécessaire pour satisfaire chacune des contraintes. Il est possible d'optimiser la taille d'échantillon en adaptant les contraintes de variance entrées dans le calcul séquentiel de manière à approcher autant que possible les valeurs α_g , $g = 1, \dots, G$ spécifiées, tout en ne dépassant pas ces variances.

Notons $\alpha = (\alpha_1, \dots, \alpha_G)$ les variances cibles calculées dans la section 3.3, $\beta = (\beta_1, \dots, \beta_G)$ les contraintes de variance utilisées dans le calcul séquentiel, $\mathbf{n}\beta = (n_{1,\beta}, \dots, n_{H,\beta})$ l'allocation obtenue avec ce calcul et $\mathbf{v} = (v_1, \dots, v_G)$ les variances correspondantes à cette allocation. On peut définir les fonctions

$$\mathcal{A} = \beta \mapsto \mathbf{n}\beta \text{ et } \psi = \beta \mapsto \mathbf{v}.$$

Notons encore $\preceq$, la comparaison terme à terme des éléments de deux vecteurs dans un espace vectoriel réel de dimension finie :

$$(a_1, \dots, a_p) \preceq (b_1, \dots, b_p) \text{ si et seulement si } a_i \leq b_i, \text{ pour tout } i = 1, \dots, p, \quad p \geq 1.$$

On peut démontrer que $\mathcal{A}$ et ψ sont monotones :

$$\beta \preceq \gamma \Rightarrow \mathcal{A}(\gamma) \preceq \mathcal{A}(\beta),$$

et

$$\beta \preceq \gamma \Rightarrow \psi(\beta) \preceq \psi(\gamma).$$

Ces propriétés sont des conséquences de la monotonie des $n_h(\alpha)$ dans la section 3.4.1.

Une procédure itérative peut être utilisée pour chercher des paramètres β tels que $\psi(\beta) \preceq \alpha$ et $\mathcal{A}(\beta) \preceq \mathcal{A}(\alpha)$. Soit ρ et θ deux nombres positifs et ϕ la fonction :

$$\phi = \beta \mapsto \beta + \rho [\alpha - \psi(\beta)]_+ + \theta [\alpha - \psi(\beta)]_- ,$$

où $[\alpha - \psi(\beta)]_+$ est le vecteur obtenu à partir de $[\alpha - \psi(\beta)]$ en remplaçant les termes négatifs par 0 et $[\alpha - \psi(\beta)]_-$ est le vecteur obtenu à partir de $[\alpha - \psi(\beta)]$ en remplaçant les termes positifs par 0. La fonction ϕ appliquée à un vecteur β de nombre positifs a pour effet d'augmenter les termes β_g qui correspondent à des variances finales v_g inférieures aux variances cibles α_g et de diminuer les termes β_g qui correspondent à des variances finales v_g supérieures aux variances cibles α_g .

Considérons les suites de vecteurs définies par :

$$\beta_0 = \alpha, \quad \beta_{t+1} = \phi(\beta_t), \quad \mathcal{A}_t = \mathcal{A}(\beta_t) \text{ et } \mathbf{v}_t = \psi(\beta_t), \quad t \geq 0.$$

Si ρ est suffisamment petit alors $\mathbf{v}_t = \psi(\beta_t) \preceq \alpha$ pour tout t . Les suites $(\beta_t)_{t \geq 0}$ et $(\mathbf{v}_t)_{t \geq 0}$ sont alors des suites croissantes au sens de $\preceq$, et $(\mathcal{A}_t)_{t \geq 0}$ est une suite décroissante. Il en résulte que $(\mathbf{v}_t)_{t \geq 0}$ et $(\mathcal{A}_t)_{t \geq 0}$ convergent. Indépendamment de la convergence, ce programme permet de trouver un vecteur de paramètres β_t tel que $\mathbf{v}_t$ satisfait les contraintes de variance et $\mathcal{A}_t$ est inférieur à $\mathcal{A}(\alpha)$.

En pratique, il est difficile de déterminer avec certitude un paramètre ρ convenable car on ne connaît pas la différentielle de la fonction ψ . Le choix d'un paramètre θ suffisamment grand est

alors utile pour trouver, parmi les β_t , des solutions qui respectent les contraintes de variance tout en conduisant à une taille d'échantillon réduite par rapport à celle obtenue à partir de $\beta_0 = \alpha$.

Pour l'échantillon d'entreprises publiques et privées de la LSE2018, cette idée a été appliquée en travaillant sur les écarts-types plutôt que sur les variances. En effet, les variances des estimateurs de totaux prennent des valeurs relativement élevées qui peuvent entraîner des problèmes de calcul numérique. Il a été constaté qu'après quinze itérations de la fonction ϕ avec les paramètres $\rho = 0.9$ et $\theta = 1.1$, les valeurs de v_t et $\mathcal{A}_t$ ne changent quasiment plus. Le vecteur β_t obtenu après ces quinze itérations a été utilisé pour déterminer l'allocation effective de l'échantillon OFS. Lors de l'allocation des échantillons densifiés, cinq itérations ont été jugées suffisantes. Pour l'échantillon d'administrations publiques, cette procédure d'optimisation n'a pas montré d'effet notable et n'a pas été utilisée.

C Allocation optimale

TABLE 17 Taille d'échantillon de l'allocation optimale par grande région ou zone de densification.

Grande région	GRS	Nb. Ents.	Alloc. OFS	densif.	Total
1- Région lémanique	GE	11'489	2'169	563	2'733
	VD	15'134	2'685	653	3'339
	Reste (VS)	7'994	1'127	0	1'127
2- Espace Mittelland	FR	5'755	1'043	1'148	2'191
	NE	3'617	646	849	1'495
	JU	1'647	290	638	928
	Reste (BE, SO)	26'579	5'249	0	5'249
3- Suisse du Nord-ouest	AG, BL, BS	22'616	5'155	0	5'155
4- Zurich	Ville de Zurich	13'118	2'955	310	3'264
	Reste ZH	20'163	3'686	0	3'686
5- Suisse orientale	TG	5'678	1'155	972	2'127
	Reste (AR, AI, GL, GR, SG, SH)	21'365	4'414	0	4'414
6- Suisse centrale	LU	8'822	2'352	480	2'831
	ZG	5'012	1'390	869	2'259
	Reste (LU, NW, OW, SZ, UR)	6'938	1'685	0	1'685
7- Tessin	TI	10'817	2'896	0	2'896
Total		186'744	38'897	6'483	45'380

Rapports de méthodes de la section méthodes statistiques de l'OFS
Methodenberichte der Sektion Statistische Methoden des BFS
Methodology reports published by the FSO's Statistical Methods Section

Qualité, L., Potterat, J. (2020). Enquête suisse sur la structure des salaires 2018. Révision du plan de sondage. Numéro de commande : 338-0081

Potterat, J., Qualité, L., Assoulin, D. (2019). Strukturhebung der eidgenössischen Volkszählung. Stichprobenplan und -ziehung, Gewichtung, Schätzverfahren und Pooling 2010-2018. Bestellnummer: 338-0080

Potterat, J., Assoulin, D., Nicoletti, J.-M. (2019). Beschäftigungsstatistik BESTA. Revision 2015: Stichprobenrahmen, Stichprobenplan und Hochrechnung. Bestellnummer: 338-0079

Massiani, A. (2017). Estimation de la couverture du nouveau recensement en Suisse en 2012. Numéro de commande : 338-0078

Nedyalkova, D., Assoulin, D. (2017). Construction of full-time equivalents for the Swiss structural business statistics. Order number: 338-0077

Ferster, M. (2016). Energieverbrauchsstatistik EVS2014 - Stichprobe, Hochrechnung und Vergleichbarkeit mit der EVS2013. Bestellnummer: 338-0076

Panchard, C. (2014). Enquête sur la participation aux activités culturelles 2008. Plan de sondage et estimations. Bestellnummer: 338-0073

Assoulin, D. (2014). Wertschöpfungsstatistik. Revision 2009: Statistische Datenaufbereitung und Hochrechnung. Bestellnummer: 338-0072

Wilhelm, M. (2014). Echantillonnage boule de neige. La méthode de sondage déterminé par les répondants. Numéro de commande: 338-0071

Assoulin, D. (2013). Wertschöpfungsstatistik. Revision 2009: Stichprobenrahmen und Stichprobenplan. Bestellnummer: 338-0070

Ferster, M. (2013). EVS I - Energieverbrauchsstatistik 2002 bis 2007: Stichprobenplan und Hochrechnung. Bestellnummer: 338-0069

Assoulin, D. (2013). Zusatzerhebung für die landwirtschaftliche Betriebszählung 2010: Stichprobenplan und Hochrechnung. Bestellnummer: 338-0068

Potterat, J. (2012). Use of conversion keys for NOGA2002-2008. Order number: 338-0067-05

Andrade, B., Salamin P.-A. (2012). Enquête sur la situation sociale et économique des étudiant-e-s des hautes écoles suisses 2009. Cadre de sondage, plan d'échantillonnage et méthodes d'estimation. Numéro de commande: 338-0066

Potterat, J., Panchard, C., Kilchmann, D. (2012). Umweltschutzausgaben der Unternehmen 2009 (UWSA2009). Stichprobenplan, Einsetzungen, Gewichtung und Schätzverfahren. Bestellnummer: 338-0065

Potterat, J. (2012). Benutzung der Umsteigeschlüssel NOGA 2002-2008. Bestellnummer: 338-0064

Potterat, J. (2011). Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe im Jahr 2009 (KNBB09). Stichprobenplan, Gewichtung und Schätzverfahren. Bestellnummer: 338-0063

Kilchmann, D., Potterat, J., Genoud, S. (2011). Gütertransporterhebung 2008. Stichprobenplan, Datenaufbereitung, Gewichtung und Schätzverfahren. Bestellnummer: 338-0062

Graf, E. (2010). Enquête suisse sur la santé 2007. Plan d'échantillonnage, pondérations et analyses pondérées des données. Numéro de commande: 338-0061

Eichenberger P., Hulliger B., Potterat J. (2010). Describing the Anticipated Accuracy of the Swiss Population Survey. Order number: 338-0060

- Graf, E. (2010). Étude empirique de l'attrition du Panel Suisse de Ménages : vers une caractérisation du profil du non-répondant. Numéro de commande: 338-0059
- Graf, E. (2009). Weightings of the Swiss Household Panel: SHP_I wave 9, SHP_II wave 4, SHP_I et SHP_II combined. Order number: 338-0058
- Graf, E. (2009). Pondérations du Panel Suisse de Ménages: PSM_I vague 9, PSM_II vague 4, PSM_I et PSM_II combinés. Numéro de commande: 338-0057-05
- Qualité, L., Tillé, Y. (2009). Estimation de la précision d'évolutions dans l'enquête sur la valeur ajoutée. Numéro de commande: 338-0056
- Renaud, A., Panchard, C. et Potterat, J. (2008). Statistique de l'emploi. Révision 2007 : méthodes d'estimation. Numéro de commande: 338-0055
- Graf, E. (2008). Pondérations du PSM. PSM_I vague 8, PSM_II vague 3, PSM_I et PSM_II combinés. Numéro de commande: 338-0054
- Andrade, B., Graf, M. (2008). Enquête suisse sur la structure des salaires 2006. Aspects méthodologiques du modèle des salaires "Salarium". Numéro de commande: 338-0053
- Renaud, A. (2008). Statistique de l'emploi. Révision 2007 : cadre de sondage et échantillonnage. Numéro de commande: 338-0052
- Graf, E. (2008). Pondérations du SILC pilote. SILC_I vague 2, SILC_II vague 1, SILC_I et SILC_II combinés. Numéro de commande: 338-0051
- Kilchmann, D. (2008). Statistik der sozialmedizinischen Institutionen 1999-2004 und Krankenhausstatistik 1999-2002. Einsetzungen für fehlende Daten. Bestellnummer: 338-0050
- Renaud, A. (2008). Technologies de l'information et de la communication. Estimations sur la base de la statistique de la valeur ajoutée. Numéro de commande: 338-0049
- Assoulin, D. (2007). Wertschöpfungsstatistik. Einsetzungsversuche für fehlende Antworten grosser Unternehmen. Bestellnummer: 338-0048
- Kilchmann, D. (2007). Beherbergungsstatistik Campingplätze. Stichprobenrahmen und Schätzverfahren 2005/06. Bestellnummer: 338-0047
- Gabler, S., Häder, S. (2007). Haushalts- und Personenerhebungen. Machbarkeit von Random Digit Dialing in der Schweiz. Bestellnummer: 338-0046
- Ferrez, J., Graf, M. (2007). Enquête suisse sur la structure des salaires. Programmes R pour l'intervalle de confiance de la médiane. Numéro de commande: 338-0045
- Renaud, A. (2007). Harmonisation de la scolarité obligatoire en Suisse (HarmoS). Design général de l'enquête et échantillon des écoles. Numéro de commande: 338-0044
- Potterat, J. (2007). Betriebszählung 2005. Statistische Methoden zur Schätzung der provisorischen Ergebnisse. Bestellnummer: 338-0043
- Hulliger, B. (2006). Umweltschutzausgaben der Unternehmen 2003, Stichprobenplan, Datenaufbereitung und Schätzverfahren. Bestellnummer: 338-0042
- Renfer, J.-P. (2006). Enquête sur les chiffres d'affaires du commerce de détail. Plan d'échantillonnage et méthodes d'estimation. Numéro de commande: 338-0041
- Salamin, P.-A. (2006). Statistique de l'aide sociale dans le domaine de l'asile. Plan de sondage et extrapolations pour l'enquête pilote 2005. Numéro de commande: 338-0040
- Renaud, A. (2006). Statistique suisse des bénéficiaires de l'aide sociale. Pondération des communes 2004. Numéro de commande: 338-0039
- Graf, M. (2006). Swiss Earnings Structure Survey 2002-2004. Compositional data in a stratified two-stage sample: Analysis and precision assessment of wage components. Order number: 338-0038
- Potterat, J. (2006). Pensionskassenstatistik 2004. Statistische Methoden zur Schätzung der provisorischen Ergebnisse. Bestellnummer: 338-0037

- Potterat, J. (2006). Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe im Jahr 2004. Stichprobenplan, Gewichtung und Schätzverfahren. Bestellnummer: 338-0036
- Kilchmann, D. (2006). Vierteljährliche Wohnbaustatistik. Stichprobenplan, statistische Datenaufarbeitung und Schätzverfahren 2005. Bestellnummer: 338-0035
- Kilchmann, D. (2006). Erhebung über Forschung und Entwicklung in der schweizerischen Privatwirtschaft 2004. Bereinigung der Stichprobe, Ersatz fehlender Werte und Schätzverfahren. Bestellnummer: 338-0034
- Kilchmann, D., Eichenberger, P., Potterat, J. (2005). Volkszählung 2000. Statistische Einsetzungsverfahren Band 2. Bestellnummer: 338-0033
- Kilchmann, D., Eichenberger, P., Potterat, J. (2005). Volkszählung 2000. Statistische Einsetzungsverfahren Band 1. Bestellnummer: 338-0032
- Graf, M., Matei, A. (2005). Enquête suisse sur la structure des salaires 2002. La précision du salaire brut standardisé médian. Numéro de commande: 338-0031
- Graf, E., Renfer, J.-P. (2005). Enquête suisse sur la santé 2002. Plan d'échantillonnage, pondération et estimation de la précision. Numéro de commande: 338-0030
- Potterat, J. (2005). Mietpreis-Strukturerhebung 2003. Gewichtung und Schätzverfahren. Bestellnummer: 338-0029
- Potterat, J. (2005). Landwirtschaftliche Betriebszählung 2003. Schätzverfahren für die Zusatzerhebung. Bestellnummer: 338-0028
- Renaud, A. (2004). Coverage estimation for the Swiss population census 2000. Estimation methodology and results. Order number: 338-0027
- Kilchmann, D. (2004). Revision des Schweizerischen Lohnindex. Schätzmethoden der Lohnindices und deren Varianzschätzer. Bestellnummer: 338-0026
- Graf, M. (2004). Enquête suisse sur la structure des salaires 2002. Plan d'échantillonnage et extrapolation pour le secteur privé. Numéro de commande: 338-0025
- Renaud, A. (2004). Analyse de données d'enquêtes. Quelques méthodes et illustration avec des données de l'OFS. Numéro de commande 338-0024
- Renaud, A., Potterat, J. (2004). Estimation de la couverture du recensement de la population de l'an 2000. Echantillon pour l'estimation de la sous-couverture (P-sample) et qualité du cadre de sondage des bâtiments. Numéro de commande: 338-0023
- Graf, M. (2004). Fusion de données. Etude de faisabilité. Numéro de commande: 338-0022
- Potterat, J. (2003). Mietpreis-Strukturerhebung 2003. Entwicklung des Stichprobenplans und Ziehung der Stichprobe. Bestellnummer: 338-0021
- Potterat, J. (2003). Landwirtschaftliche Betriebszählung 2003. Stichprobenplan der Zusatzerhebung. Bestellnummer: 338-0020.
- Renaud, A. (2003). Estimation de la couverture du recensement de la population de l'an 2000. Echantillon pour l'estimation de la sur-couverture (E-sample). Numéro de commande: 338-0019
- Hulliger, B. (2003). 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 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

Programme des publications de l'OFS

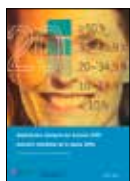
En tant que service statistique central de la Confédération, l'Office fédéral de la statistique (OFS) a pour tâche de rendre les informations statistiques accessibles à un large public. Il utilise plusieurs moyens et canaux pour diffuser ses informations statistiques par thème.

Les domaines statistiques

- 00 Bases statistiques et généralités
- 01 Population
- 02 Espace et environnement
- 03 Travail et rémunération
- 04 Économie nationale
- 05 Prix
- 06 Industrie et services
- 07 Agriculture et sylviculture
- 08 Énergie
- 09 Construction et logement
- 10 Tourisme
- 11 Mobilité et transports
- 12 Monnaie, banques, assurances
- 13 Sécurité sociale
- 14 Santé
- 15 Éducation et science
- 16 Culture, médias, société de l'information, sport
- 17 Politique
- 18 Administration et finances publiques
- 19 Criminalité et droit pénal
- 20 Situation économique et sociale de la population
- 21 Développement durable, disparités régionales et internationales

Les principales publications générales

L'Annuaire statistique de la Suisse



L'Annuaire statistique de la Suisse de l'OFS constitue depuis 1891 l'ouvrage de référence de la statistique suisse. Il englobe les principaux résultats statistiques concernant la population, la société, l'État, l'économie et l'environnement de la Suisse.

Le Mémento statistique de la Suisse



Le mémento statistique résume de manière concise et attrayante les principaux chiffres de l'année. Cette publication gratuite de 52 pages au format A6/5 est disponible en cinq langues (français, allemand, italien, romanche et anglais).

Le site Internet de l'OFS: www.statistique.ch

Le portail «Statistique suisse» est un outil moderne et attrayant vous permettant d'accéder aux informations statistiques actuelles. Nous attirons ci-après votre attention sur les offres les plus prisées.

La banque de données des publications pour des informations détaillées

Presque tous les documents publiés par l'OFS sont disponibles gratuitement sous forme électronique sur le portail Statistique suisse (www.statistique.ch). Pour obtenir des publications imprimées, vous pouvez passer commande par téléphone (058 463 60 60) ou par e-mail (order@bfs.admin.ch). www.statistique.ch → Trouver des statistiques → Catalogues et banques de données → Publications

Vous souhaitez être parmi les premiers informés?



Abonnez-vous à un Newsmail et vous recevrez par e-mail des informations sur les résultats les plus récents et les activités actuelles concernant le thème de votre choix. www.news-stat.admin.ch

STAT-TAB: la banque de données statistiques interactive



La banque de données statistiques interactive vous permet d'accéder simplement aux résultats statistiques dont vous avez besoin et de les télécharger dans différents formats. www.stattab.bfs.admin.ch

Statatlas Suisse: la banque de données régionale avec ses cartes interactives



L'atlas statistique de la Suisse, qui compte plus de 4500 cartes, est un outil moderne donnant une vue d'ensemble des thématiques régionales traitées en Suisse dans les différents domaines de la statistique publique. www.statatlas-suisse.admin.ch

Pour plus d'informations

Centre d'information statistique

058 463 60 11, info@bfs.admin.ch

Le plan de sondage de l'enquête suisse sur la structure des salaires (ESS) a fait l'objet d'une révision approfondie pour l'enquête de l'année 2018. Le principal objectif de cette révision est d'utiliser les données administratives sur les revenus des personnes salariées pour améliorer l'efficacité de l'enquête. Le plan de sondage, comme pour les enquêtes précédentes, assigne à chaque entreprise ou administration du cadre de sondage une probabilité de sélection déterminée en fonction de sa grande région, de son canton ou de sa commune de densification, de son activité économique et de son nombre d'employés. La méthode utilisée pour calculer les taux de sondage de l'ESS2018 repose sur le choix de paramètres d'intérêt principaux (salaires médians notamment) et de la précision avec laquelle ces paramètres doivent être estimés.

Le résultat de la révision a conduit à la sélection de 45 877 unités pour l'ESS2018. Cela représente une diminution de 4691 unités, soit 9,3% par rapport à l'ESS2016 qui comptait 50 568 unités dans son échantillon.

En ligne

www.statistique.ch

Imprimés

www.statistique.ch

Office fédéral de la statistique

CH-2010 Neuchâtel

order@bfs.admin.ch

tél. 058 463 60 60

Numéro OFS

338-0081

ISBN

978-3-303-00657-3

**La statistique
compte pour vous.**

www.la-statistique-compte.ch