2022–
2023

00

Bases
statistiques
et généralités

Neuchâtel 2025

Enquête PIAAC 2022–2023

Plan de sondage, pondération, estimation de variance et analyse du biais dû à la non-réponse.

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

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

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 PIAAC 2022–2023

Plan de sondage, pondération, estimation de variance et analyse du biais dû à la non-réponse.

Rédaction Blandine Bianchi, OFS
Michael Leuenberger, OFS
Anne Massiani, OFS
Christian Panchard, OFS

Éditeur Éditeur Office fédéral de la statistique (OFS)

Neuchâtel 2025

Éditeur: Office fédéral de la statistique (OFS)

Renseignements: Anne Massiani, tél. +41 58 463 62 65
Anne.Massiani@bfs.admin.ch

Rédaction: Blandine Bianchi, OFS
Michael Leuenberger, OFS
Anne Massiani, OFS
Christian Panchard, 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: Publishing et diffusion PUB, OFS
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. +41 58 463 60 60
Impression réalisée en Suisse

Copyright: OFS, Neuchâtel 2025
La reproduction est autorisée, sauf à des fins commerciales, si la source est mentionnée.

Numéro OFS: 338-0091

ISBN: 978-3-303-00768-6

Table des matières

Préambule et Résumé 5

1 Introduction 7

2 Présentation de l'enquête 8

3 Plan de sondage et tirage 9
3.1 Particularités/buts de l'enquête . . . 9
3.2 Population et stratégie d'enquête . . . 9
3.2.1 Cadre, population cible . . . 9
3.2.2 Construction du plan d'échantillonnage . . . 11
3.2.3 Tailles de l'échantillon brut . . . 13
3.2.4 Réserves . . . 13
3.2.5 Tirage coordonné et tailles d'échantillon effectives . . . 14
3.2.6 Table récapitulative . . . 14

4 Pondération 16
4.1 Principes généraux de la pondération . . . 16
4.1.1 Caractéristiques et déroulement de l'enquête . . . 16
4.1.2 Population cible pour l'extrapolation . . . 16
4.1.3 Poids initial et correction pour la non-éligibilité . . . 18
4.1.4 Correction de la non-réponse . . . 19
4.1.5 Calage . . . 20
4.2 Variables auxiliaires disponibles . . . 21
4.2.1 Variables auxiliaires pour la correction de la non-reponse . . . 21
4.2.2 Variables auxiliaires pour le calage . . . 23
4.3 Mise en oeuvre de la pondération . . . 27
4.3.1 Ajustement de la non-résponse . . . 27
4.3.2 Calage . . . 28

5 Estimation de la variance 30
5.1 Principe . . . 30
5.2 Analyse de l'ajustement des poids . . . 33
5.2.1 Comparaison des estimations avant et après les ajustements des poids . . . 33
5.2.2 Comparaison des estimations pondérées avec les totaux externes . . . 33
5.2.3 Analyse des variables recueillies lors de la collecte des données : codes liés à la littératie (literacy-related) . . . 34

6 Étude du biais lié à la non-réponse pour les variables mesurant les compétences 36
6.1 Comparaison des compétences en fonction du niveau d'effort . . . 36
6.2 Étendue du biais potentiel . . . 37

7 Conclusions 38

Références 39

Annexes 40

A Table de Notations 40

B Variables auxiliaires 42

C Niveau de formation 44

D Groupes Homogènes de Réponse (rhg) 45

Rapports de méthodes de la section méthodes statistiques de l'OFS 47

Préambule

Le programme PIAAC (Program for the International Assessment of Adult Competencies) est une étude internationale conduite par l'Organisation de Coopération et de Développement Économiques (OCDE) qui évalue des compétences de base des adultes. La Suisse y a participé pour la première fois lors de l'enquête qui s'est déroulée en 2022-2023, sous l'égide de la section BILD-S de l'Office fédéral de la statistique. Il s'agit d'une enquête complexe qui nécessite des traitements méthodologiques appropriés, aussi bien au niveau du plan de sondage que de la pondération. Les méthodes utilisées suivent les recommandations de l'OCDE et sont décrites dans ce rapport.

Les auteurs adressent leurs plus sincères remerciements à Daniel Kilchmann pour son accompagnement et son soutien au cours de ce stimulant projet. Également un immense merci à Jean-Pierre Renfer et Daniel Assoulin pour leur précieuse relecture. Leurs remarques et conseils ont permis d'améliorer la qualité de ce rapport. Les auteurs tiennent encore à remercier chaleureusement Emanuel von Erlach, Audrey Bovier-Michelet et Anouk Widmer de la section BILD-S pour avoir su créer un climat de travail à la fois constructif et agréable. Enfin, nous aimerions témoigner notre reconnaissance à Wendy Van de Kerckhove de la société Westat pour les riches échanges que nous avons eus durant ce projet.

Résumé

L'enquête PIAAC fournit une vue d'ensemble des niveaux de compétences en lecture, calcul et résolution adaptative de problèmes parmi la population permanente de Suisse âgée de 16 à 65 ans, tout en permettant des comparaisons internationales.

L'enquête PIAAC est une enquête complexe dont certaines particularités ont nécessité des traitements méthodologiques appropriés, aussi bien au niveau du plan de sondage que des méthodes d'estimation utilisées. Le plan de sondage tient notamment compte des régions linguistiques de la Suisse afin de permettre des exploitations à ce niveau d'agrégation. D'autre part, l'enquête étant réalisée en face-à-face, le plan de sondage a contribué à sa faisabilité sur le terrain en favorisant une certaine proximité géographique dans les strates. Elle est réalisée auprès d'un échantillon brut de 22 091 personnes, avec un taux de réponse d'environ 30%.

Une attention particulière a été portée sur la recherche de variables auxiliaires pertinentes pour la pondération, et différentes sources ont été considérées : informations sociodémographiques contenues dans le cadre de sondage, données AVS en lien avec les revenus, données du relevé structurel, etc. Les R^2 de la régression sur les variables de calage pour trois variables clés de l'enquête se situent aux alentours de 36%, ce qui entraîne en règle générale des gains appréciables sur la variance des estimations. De plus, la non-réponse liée à la thématique de l'enquête (literacy related non-response) est identifiée et fait l'objet d'un traitement spécifique.

L'estimation de la variance se base sur des réplications des poids des unités de l'échantillon. Les jeux de poids ainsi produits peuvent être livrés aux utilisateurs qui sont alors en mesure de calculer la variance de n'importe quel paramètre de façon très simple, sans qu'il soit nécessaire de leur livrer les variables auxiliaires, parfois sensibles du point de vue de la protection des données, ayant permis d'aboutir aux poids finaux.

Enfin, l'application de ces méthodes a été complétée par une analyse du biais potentiel lié à la non-réponse qui, si un risque de biais ne peut pas être écarté, livre néanmoins des éléments très rassurants quant à son étendue.

1 Introduction

Le programme PIAAC (Program for the International Assessment of Adult Competencies) est une étude internationale conduite par l'Organisation de Coopération et de Développement Économiques [1] (OCDE) qui évalue des compétences de base des adultes. La Suisse y a participé pour la première fois lors de l'enquête qui s'est déroulée en 2022-2023. L'objectif de ce programme est de fournir une image complète des compétences dans les domaines couverts par PIAAC pour les pays participants, tant au niveau national qu'au niveau international. Réalisée dans une première édition en 2011, puis reconduite en 2023 dans 31 pays, l'enquête PIAAC étudie les compétences de base des adultes à maîtriser des informations écrites et chiffrées du quotidien et à les mobiliser dans la vie de tous les jours, à la maison et au travail. Les compétences de base évaluées dans cette enquête couvrent les connaissances et des aptitudes fondamentales dans les domaines de la lecture (littératie), des mathématiques élémentaires (numératie) et de la résolution adaptative de problèmes. Ces compétences sont considérées comme essentielles dans la vie quotidienne, influençant l'intégration sociale, économique et politique des individus dans un monde de plus en plus technologique, numérisé et en manque de personnel qualifié. Elles constituent la base du développement de nombreuses autres aptitudes et compétences spécifiques.

L'enquête, qui est réalisée en face-à-face, est présentée plus en détails dans la section 2. Le plan de sondage est décrit dans la section 3. Il favorise une certaine proximité géographique dans les strates afin de faciliter le déroulement de l'enquête sur le terrain. Les principes généraux de la pondération, ainsi que les détails de son implémentation, sont présentées dans la section 4. La pondération se base sur une correction de la non-réponse suivie d'un calage, avec toutefois la spécificité que deux types de non-réponse sont considérés : une non-réponse qui n'est pas liée à la thématique de l'enquête (non literacy related non-response) et une non-réponse qui y est liée (literacy related non-response). Une attention particulière est portée à la prise en compte de variables de calage corrélées à la thématique de l'enquête. Le niveau de formation est ainsi intégré au calage. En ce qui concerne la précision des estimations, leur variance est estimée grâce à la méthode de Fay, qui est basée sur des réplications de l'échantillon et est une extention de la méthode « Balanced Repeated Replication » (BRR). Cette procédure, ainsi que son application dans l'enquête PIAAC en Suisse, est présentée dans la section 5. La section 6 est consacrée à l'analyse de l'étendue du biais potentiel, qui est notamment évalué en faisant différentes hypothèses sur les réponses qu'auraient fournies les non-répondants. Une hypothèse envisagée est que les non-répondants auraient fourni des réponses similaires aux répondants ayant été difficiles à contacter. Enfin, les conclusions de ce travail sont présentées dans la section 7.

1. En anglais, Organisation for Economic Cooperation and Development (OECD).

2 Présentation de l'enquête

L'enquête PIAAC fournit une vue d'ensemble des niveaux de compétences en lecture, calcul et résolution adaptative de problèmes parmi la population permanente de la Suisse âgée de 16 à 65 ans, tout en permettant des comparaisons internationales, (Widmer et al., 2024). Elle est réalisée en face-à-face auprès d'un échantillon brut de 22 091 personnes, avec un taux de réponse d'environ 30%. Les compétences sont mesurées par des exercices résolus de manière autonome sur une tablette. L'évaluation est complétée par un questionnaire contextuel dense renseignant sur de nombreuses caractéristiques des participants, notamment leurs caractéristiques socio-démographiques principales, leur parcours de formation, leur statut professionnel, leur contexte familial, leur contexte linguistique et migratoire, leur utilisation des différentes compétences de base au quotidien et au travail ou encore leurs compétences socio-émotionnelles.

3 Plan de sondage et tirage

3.1 Particularités/buts de l'enquête

L'enquête PIAAC a la particularité d'être une enquête avec une interview en face-à-face. Pour cette raison, un plan d'échantillon qui puisse tenir compte de la proximité des personnes de l'échantillon pour permettre aux interviewers de se déplacer facilement au sein d'une même strate a été choisi.
Une enquête test a été réalisée de mai à septembre 2021, puis l'enquête proprement dite de septembre 2022 à mai 2023. L'ensemble de ce qui suit dans cette section traite de l'échantillonnage PIAAC pour l'enquête principale. Certains aspects de l'enquête test seront évoqués dans les parties concernées.
A noter également que beaucoup d'aspects sont décrits par les « STANDARDS AND GUIDELINES PIAAC » OECD (2022a). Ces éléments seront évoqués dans les parties concernées, et les adaptations faites par l'OFS pour s'approcher le mieux possible de ces directives seront exposées.

3.2 Population et stratégie d'enquête

3.2.1 Cadre, population cible

Population cible définie par l'OCDE
La population cible de l'enquête PIAAC est définie au point 4.1 (Target Population) des « STANDARDS AND GUIDELINES PIAAC » OECD (2022a).
Globalement, selon la traduction de ce document en anglais :

> **Standard 4.1.1 La population cible de l'enquête PIAAC est constituée de tous les adultes non institutionnalisés âgés de 16 à 65 ans (inclus) qui résident dans le pays (dont le lieu de résidence habituel se trouve dans le pays) au moment de la collecte des données. Les adultes doivent être inclus sans distinction de citoyenneté, de nationalité ou de langue.**

Plus précisément :

> **Guideline 4.1.1A** La population cible exclut les adultes vivant dans des unités d'habitation collectives institutionnelles (ou logements collectifs) telles que les prisons, les hôpitaux et les maisons de retraite, ainsi que les adultes résidant dans des casernes et des bases militaires. Les militaires à temps plein et à temps partiel qui ne résident pas dans des casernes ou des bases militaires sont toutefois inclus dans la population cible. Les adultes vivant dans d'autres unités d'habitation collective non institutionnelles (ou logements collectifs), telles que les logements pour travailleurs ou les foyers de transition, sont également inclus dans la population cible.
>
> **Guideline 4.1.1B** La population cible comprend les adultes vivant à l'école dans des logements collectifs pour étudiants, tels que les dortoirs. Voir la recommandation 4.4.1 pour des suggestions sur l'échantillonnage de ces adultes.

Population cible mise en oeuvre
De manière générale, pour les enquêtes de l'OFS sur les personnes résidentes permanentes de la Suisse, l'élaboration du plan de sondage et de la population cible se base sur le dernier cadre de sondage SRPH (Stichprobenrahmen für Personen- und Haushaltsehebungen) disponible. Ce cadre est construit avec les données trimestrielles provenant de registres communaux, cantonaux et fédéraux.
Le SRPH, présenté avec la sélection suivante, est ce qui se rapproche au mieux de la population cible définie par l'OCDE :

- les personnes résidentes permanentes de la Suisse, à leur domicile principal, vivant dans un ménage privé, et âgées de 16 ans à 65 ans (inclus).

Cependant, les points 4.1.1A et 4.1.1B énoncés dans les guidelines font apparaître chacun un problème en lien avec la définition des ménages collectifs qui n'est pas identique entre l'OFS et l'OCDE. En particulier :

- les casernes et les bâtiments militaires ne sont pas considérés comme des ménages collectifs par l'ordonnance sur l'harmonisation de registres [2], ordonnance qui détermine les différents traitements permettant de définir la population résidente permanente du SRPH.
- par contre, l'ordonnance inclut les *internats et les foyers d'étudiants* dans les ménages collectifs, alors qu'ils devraient être compris dans la population cible.

Pour les personnes ne vivant pas en ménage collectif selon l'ordonnance, le numéro de commune que nous avons à disposition (et qui est la seule information disponible dans le SRPH sur le lieu de domicile) ne nous permet pas de savoir si la personne vit en caserne ou dans un logement extérieur à la caserne, dans un foyer d'étudiants ou dans celui d'un ou des parents. Il est donc aussi impossible de déterminer le nombre ou la proportion de personnes concernées.
D'autre part, la sélection usuelle pour les enquêtes tirées dans le SRPH est la suivante :

- les personnes résidentes permanentes de la Suisse, à leur domicile principal, vivant dans un ménage privé, et âgées de 16 ans à 65 ans (inclus).

sans prendre en compte :

- les demandeurs d'asile;
- les personnes sous tutelle / curatelle;
- les fonctionnaires internationaux;
- les personnes ayant des permis de séjour de courte durée;
- et les personnes n'ayant dans les registres qu'une adresse de référence hors de Suisse.

Cette sélection est différente de celle préconisée pour l'enquête, mais elle est en adéquation avec la plupart des tirages de l'OFS pour des enquêtes auprès des personnes et des ménages. Pour la suite du document c'est cette dernière sélection qui sera considérée, par manque d'informations permettant de s'approcher encore mieux de la définition de la population cible de l'OCDE.

Cadre d'échantillonnage défini par l'OCDE
Le cadre d'échantillonnage de l'enquête PIAAC est défini au point 4.2 (Sampling Frame) des « STANDARDS AND GUIDELINES PIAAC » OECD (2022a).

> **Standard 4.2.1 Le cadre d'échantillonnage doit comprendre 95% ou plus de la population cible PIAAC. En d'autres termes, le taux de sous-couverture, combiné à tous les stades de l'échantillonnage, ne doit pas dépasser 5% .**

Il s'agit ici principalement de déterminer la proportion des immigrants sans papiers, comme cela était fait par la majorité des pays participant au cycle précédent de PIAAC.
Un autre point important des « STANDARDS AND GUIDELINES PIAAC » OECD (2022a) est le suivant :

> **Guideline 4.2.2B** La ou les bases de sondage doivent contenir les variables nécessaires à la mise en oeuvre du plan d'échantillonnage prévu (par exemple, stratification, calcul des mesures de taille), à la pondération et à l'analyse du biais de non-réponse. Les variables pour la stratification et la pondération des ajustements de non-réponse doivent être disponibles pour tous les enregistrements de la base de sondage et être des variables corrélées avec la compétence, telles que l'âge et la formation.

Cadre d'échantillonnage adapté pour l'OFS
Pour l'élaboration du plan, la population cible prise en considération, de taille N, est généralement celle du dernier cadre de sondage SRPH disponible.
La taille de la population cible prise en considération pour le plan de sondage et qui correspond au SRPH avec la date de référence du 31 mars 2022, créé le 10 mai 2022, est $N = 5\,622\,772$.

2. https://www.fedlex.admin.ch/eli/cc/2007/868/fr

A préciser que dans le cas du cadre de sondage SRPH, la proportion des immigrants sans papiers a été estimée à bien moins que 2% (comme demandé par les recommandations de l'OCDE) lors de l'enquête de couverture de 2012 Massiani (2017). De plus, le SRPH correspond bien aux besoins du tirage de l'échantillon et sera complété pour la pondération au point 4.2.

3.2.2 Construction du plan d'échantillonnage

Selon les « STANDARDS AND GUIDELINES PIAAC » OECD (2022a), la taille nette m de l'échantillon doit se situer entre 4000 et 5000 personnes par langue considérée pour l'enquête.
Dans le cas de la Suisse et ses particularités linguistiques, il a été convenu en accord avec les groupes de travail de l'OCDE que la région linguistique principale serait la partie germanophone, et que les standards ne s'appliquerait qu'à cette région linguistique. Les parties francophone et italophone ont tout de même été couvertes par l'enquête mais en relâchant la contrainte sur la taille nette. Pour ces raisons, des tailles nettes de $m_{DE} = 4000$ pour la partie germanophone, $m_{FR} = 2000$ pour la partie francophone, et $m_{IT} = 1000$ pour la partie italophone ont été retenues. Nous avons ainsi une taille nette totale visée de $m =$ 7000 personnes.
De plus, du fait que les interviews ont dû être réalisées en face-à-face, il a été décidé de créer des strates d'échantillonnage pouvant permettre une certaine proximité entre les individus tirés. A cet effet, les bassins d'emploi (régions fonctionnelles dans lesquelles la majorité des personnes actives employées habitent et travaillent - von Siebenthal & Kohler (2019)) ont été retenus pour la définition de ces strates.
La table 1 présente les strates retenues par bassin d'emploi avec la population cible du cadre au 9 février 2021 (cadre utilisé pour le développement du plan de sondage). Ces strates sont allouées aux trois régions linguistiques selon les langues majoritaires des communes les composant.

TABLE 1 Strates selon les bassins d'emploi, et allocation des strates aux régions linguistiques.

Bassin d'emploi	Strate	Nombre de communes	Sous-population DE	Sous-population FR	Sous-population IT	Population Total
Biel/Bienne	101	54	100 782	3 896	0	104 678
Solothurn	102	66	107 669	0	0	107 669
Murten	103	30	34 774	3 924	0	38 698
Bern (*)	104	53	247 525	0	0	247 525
Burgdorf (*)	105	53	74 051	0	0	74 051
Brig-Glis (*)	106	62	52 599	0	0	52 599
Reinach (BL)-Allschwil (*)	107	53	146 008	0	0	146 008
Basel	108	3	126 541	0	0	126 541
Liestal	109	55	60 636	0	0	60 636
Rheinfelden	110	30	48 987	0	0	48 987
Thun (*)	111	88	136 659	2 879	0	139 538
Langenthal	112	41	47 974	0	0	47 974
Aarau-Olten	113	95	229 922	0	0	229 922
Sursee	114	36	79 542	0	0	79 542
Luzern	115	32	173 068	0	0	173 068
Stans (*)	116	45	86 364	0	0	86 364
Zug	117	38	130 234	0	0	130 234
Schwyz-Einsiedeln	118	17	48 966	0	0	48 966
Baden	119	87	190 080	0	0	190 080
Schaffhausen	120	37	68 024	0	0	68 024
Dietikon-Schlieren	121	30	115 500	0	0	115 500
Kloten	122	37	127 352	0	0	127 352
Zürich	123	1	289 385	0	0	289 385
Horgen-Wädenswil	124	18	105 231	0	0	105 231
Küsnacht (ZH)-Meilen	125	6	36 845	0	0	36 845
Uster-Dübendorf	126	18	134 976	0	0	134 976
Wetzikon (ZH)	127	16	94 543	0	0	94 543

(La table continue sur la page suivante...)

Table 1 Strates selon les bassins d'emploi, et allocation des strates aux régions linguistiques. *(suite)*

Bassin d'emploi	Strate	Nombre de communes	Sous-population DE	Sous-population FR	Sous-population IT	Population Total
Rapperswil-Jona	128	5	34 061	0	0	34 061
Winterthur	129	39	139 678	0	0	139 678
Freienbach-Glarus	130	19	80 706	0	0	80 706
Wil (SG)	131	23	73 842	0	0	73 842
Frauenfeld (*)	132	56	130 193	0	0	130 193
St. Gallen	133	44	170 993	0	0	170 993
Widnau-Au (*)	134	43	132 500	0	0	132 500
Chur (*)	135	34	69 972	0	0	69 972
Samedan-Pontresina (*)	136	56	47 474	0	0	47 474
Total DE		**1 420**	**3 973 656**	**10 699**	**0**	**3 984 355**
Vernier-Lancy	237	21	0	115 496	0	115 496
Genève	238	1	0	122 421	0	122 421
Nyon (*)	239	71	0	102 209	0	102 209
Thônex-Chêne-Bougeries	240	16	0	48 063	0	48 063
Renens-Ecublens	241	54	0	82 334	0	82 334
Prilly-Le Mont-sur-L.	242	38	0	50 554	0	50 554
Lausanne	243	33	0	149 297	0	149 297
Yverdon-les-Bains (*)	244	76	0	61 403	0	61 403
Monthey (*)	245	24	0	59 132	0	59 132
Montreux-Vevey	246	26	0	73 258	0	73 258
La Chaux-de-Fonds (*)	247	21	0	47 629	0	47 629
Neuchâtel	248	25	2 902	75 484	0	78 386
Payerne-Estavayer	249	39	0	44 717	0	44 717
Bulle	250	34	385	46 387	0	46 772
Fribourg	251	45	17 880	78 223	0	96 103
Delémont (*)	252	78	217	65 524	0	65 741
Martigny (*)	253	21	0	43 869	0	43 869
Sion (*)	254	29	978	85 544	0	86 522
Total FR		**652**	**22 362**	**1 351 544**	**0**	**1 373 906**
Locarno	355	27	0	0	42 668	42 668
Bellinzona (*)	356	32	0	0	56 773	56 773
Lugano	357	49	0	0	94 381	94 381
Mendrisio	358	15	0	0	33 597	33 597
Samedan-Pontresina (*)	359	3	0	0	3 594	3 594
Total IT		**126**	**0**	**0**	**231 013**	**231 013**
Total		**2 198**	**3 996 018**	**1 362 243**	**231 013**	**5 589 274**

Dans la table 1, les strates avec l'indication (*) mentionne une fusion de plusieurs bassins d'emploi. En effet, en accord avec l'institut de sondage mandaté pour l'enquête, il a été convenu de regrouper certain bassins d'emploi dont la taille était petite. La fusion de ces bassins d'emploi est présentée à la table 2.

Concernant le bassin d'emploi *Samedan-Pontresina*, il est constitué de tous les bassins d'emploi comme indiqué dans la table 2. Il est ensuite séparé en deux bassins d'emploi distincts (strate 136 et 359). La strate 359 contient uniquement les communes italophones (nom et numéro de communes : Brusio-3551, Poschiavo-3561 et Bregaglia-3792), alors que la strate 136 contient uniquement les communes germanophones.

Table 2 Fusions de bassins d'emploi.

Bassin d'emploi de référence	ID Bassin d'emploi de référence	ID Bassins d'emploi fusionnés
Bern	6012	6012, 6013, 6014
Burgdorf	6020	6015, 6020, 6030

(Continue sur la page suivante...)

Table 2 Fusions de bassins d'emploi. *(suite)*

Bassin d'emploi de référence	ID Bassin d'emploi de référence	ID Bassins d'emploi fusionnés
Brig-Glis	7090	7060, 7070, 7080, 7090
Reinach (BL)-Allschwil	8012	8011, 8012
Thun	9030	9010, 9020, 9030, 9040, 9050, 9060
Stans	11050	11010, 11030, 11050, 11070
Frauenfeld	15020	15020, 15030
Widnau-Au	15080	15040, 15060, 15070, 15080, 16020, 16040
Chur	16050	16010, 16050
Samedan-Pontresina	16100	16030, 16060, 16070, 16080, 16090, 16100, 16110
Nyon	1014	1013, 1014, 2021
Yverdon-les-Bains	2030	2010, 2030
Monthey	2040	2040, 2060
La Chaux-de-Fonds	3020	3010, 3020
Delémont	5040	3040, 5010, 5020, 5040, 5050
Martigny	7010	7010, 7020
Sion	7030	7030, 7040, 7050
Bellinzona	13040	13010, 13030, 13040

Pour chaque région linguistique, la taille nette visée m de l'échantillon est allouée aux strates proportionnellement à la taille de ces dernières. On a alors la taille nette par strate

$$m_h = m_L \cdot \frac{N_h}{N_L},$$

où $h = 1, 2, ..., H$ désigne les strates, N_h les tailles de la population cible de ces strates, $L \in \{DE, FR, IT\}$ la région linguistique à laquelle la strate h appartient, m_L les tailles nettes visées pour les trois régions linguistiques, et N_L les tailles de la population cible de ces régions linguistiques. Les tailles nettes m_h, par strate, sont données dans la table 3.

Pour la partie italophone du bassin d'emploi *Samedan-Pontresina* de population nettement plus petite que les autres bassins d'emploi, la taille nette visée a été fixée à 30.

3.2.3 Tailles de l'échantillon brut

Pour obtenir des tailles brutes $\widetilde{n}_h$ à partir des tailles nettes m_h, des taux de réponse u_h sont estimés au niveau des strates.

À partir de ces taux de réponse, on peut alors calculer :

$$\widetilde{n}_h = \frac{m_h}{u_h}.$$

Suite au taux de réponse de 25% obtenu pour l'enquête test, taux nettement plus bas que celui recommandé par les « Standards and Guidelines PIAAC » OECD (2022a), un taux de réponse de 38% a été retenu comme objectif pour l'enquête principal.
Les tailles brutes $\widetilde{n}_h$, par strate, sont données dans la table 3.

3.2.4 Réserves

Une réserve de 20% par strate a été créée et répartie dans quatre réserves de tailles équivalentes. Cependant, seule une réserve a été activée (voir table 3).

3.2.5 Tirage coordonné et tailles d'échantillon effectives

Le tirage d'un échantillon, de taille espérée égale à la somme des tailles brutes de l'échantillon principal et des réserves, est effectué avec le système de coordination des tirages du SRPH sous un plan de Poisson. Cet échantillon est ensuite réparti en un échantillon principal et en quatre réserves, en tenant compte de leur taille brute respective et toujours sous un plan de Poisson. Comme le tirage de l'échantillon et sa répartition sont réalisés sous des plans de Poisson, les tailles brutes obtenues ne sont pas exactement égales aux tailles brutes visées.

Le tirage est effectué dans le SRPH dont la date de référence est la plus récente possible par rapport à la période d'interview. Ce cadre est plus récent que celui qui est pris en considération pour l'élaboration du plan.
Dans le cas de cette enquête, le tirage a été effectué le 16 juin 2022 dans le SRPH avec la date de référence du 31 mars 2022, créé le 10 mai 2022, comme mentionné au point 3.2.1.

3.2.6 Table récapitulative

La table récapitulative 3 donne, par strate, les différentes tailles de l'enquête principale.

Table 3 Tailles, par strate, de l'enquête PIAAC.

Strate (h)	Strate ID	Taille échantillon						
		net	brut		réserve			total
		visée	visée	tirée	visée	tirée	activée	activée
		(m_h)	($\widetilde{n}_h$)	($\widetilde{\widetilde{n}}_h$)	(l_h)	($\widetilde{\widetilde{l}}_h$)	($\bar{l}_h$)	$n_h = (\widetilde{\widetilde{n}}_h + \bar{l}_h)$
Biel/Bienne	101	105	276	283	220	226	61	344
Solothurn	102	108	284	290	228	220	42	332
Murten	103	39	103	95	84	81	23	118
Bern	104	247	650	635	520	555	143	778
Burgdorf	105	73	192	190	152	155	41	231
Brig-Glis	106	53	139	144	112	116	29	173
Reinach (BL)-Allschwil	107	145	382	388	304	321	86	474
Basel	108	127	334	369	268	284	71	440
Liestal	109	61	161	176	128	130	38	214
Rheinfelden	110	49	129	117	104	104	37	154
Thun	111	138	363	350	292	290	67	417
Langenthal	112	48	126	126	100	91	31	157
Aarau-Olten	113	232	611	605	488	466	120	725
Sursee	114	80	211	227	168	168	41	268
Luzern	115	175	461	472	368	383	83	555
Stans	116	86	226	218	180	192	50	268
Zug	117	131	345	344	276	283	67	411
Schwyz-Einsiedeln	118	49	129	108	104	95	23	131
Baden	119	192	505	502	404	382	102	604
Schaffhausen	120	68	179	175	144	153	41	216
Dietikon-Schlieren	121	116	305	287	244	282	71	358
Kloten	122	128	337	331	268	260	62	393
Zürich	123	292	768	778	616	668	144	922
Horgen-Wädenswil	124	106	279	291	224	228	58	349
Küsnacht (ZH)-Meilen	125	37	97	116	76	78	18	134
Uster-Dübendorf	126	136	358	347	288	283	72	419
Wetzikon (ZH)	127	95	250	243	200	205	49	292
Rapperswil-Jona	128	35	92	76	72	79	20	96
Winterthur	129	141	371	346	296	317	68	414
Freienbach-Glarus	130	81	213	224	172	161	41	265
Wil (SG)	131	74	195	193	156	151	37	230
Frauenfeld	132	130	342	357	272	266	64	421
St. Gallen	133	171	450	428	360	360	94	522
Widnau-Au	134	134	353	354	284	277	64	418
Chur	135	70	184	188	148	146	36	224
Samedan-Pontresina	136	47	124	119	100	93	16	135
Total DE		**3 999**	**10 524**	**10 492**	**8 420**	**8 549**	**2 110**	**12 602**
Vernier-Lancy	237	169	445	484	356	351	85	569
Genève	238	174	458	450	368	376	80	530
Nyon	239	149	392	411	312	311	87	498

(Continue sur la page suivante...)

TABLE 3 Tailles, par strate, de l'enquête PIAAC. *(suite)*

Strate (h)	Strate ID	Taille échantillon						
		net	brut		réserve			total
		visée	visée	tirée	visée	tirée	activée	activée
		(m_h)	$(\widetilde{n}_h)$	$(\widetilde{\widetilde{n}}_h)$	(l_h)	$(\widetilde{\widetilde{l}}_h)$	$(\bar{l}_h)$	$n_h = (\widetilde{\widetilde{n}}_h + \bar{l}_h)$
Thônex-Chêne-Bougeries	240	70	184	178	148	136	42	220
Renens-Ecublens	241	122	321	323	256	237	64	387
Prilly-Le Mont-sur-L.	242	74	195	198	156	161	38	236
Lausanne	243	218	574	570	460	443	112	682
Yverdon-les-Bains	244	89	234	223	188	192	44	267
Monthey	245	86	226	214	180	179	46	260
Montreux-Vevey	246	107	282	278	224	242	57	335
La Chaux-de-Fonds	247	68	179	169	144	154	45	214
Neuchâtel	248	114	300	315	240	250	64	379
Payerne-Estavayer	249	66	174	168	140	149	34	202
Bulle	250	69	182	188	144	122	31	219
Fribourg	251	140	368	370	296	293	59	429
Delémont	252	94	247	234	196	190	48	282
Martigny	253	64	168	176	136	138	32	208
Sion	254	126	332	349	264	250	61	410
Total FR		**1 999**	**5 261**	**5 298**	**4 208**	**4 174**	**1 029**	**6 327**
Locarno	355	182	479	499	384	399	95	594
Bellinzona	356	243	639	631	512	532	146	777
Lugano	357	404	1063	1029	852	823	207	1236
Mendrisio	358	142	374	381	300	276	72	453
Samedan-Pontresina	359	30	79	80	64	73	22	102
Total IT		**1 001**	**2 634**	**2 620**	**2 112**	**2 103**	**542**	**3 162**
Total		**6 999**	**18 419**	**18 410**	**14 740**	**14 826**	**3 681**	**22 091**

4 Pondération

4.1 Principes généraux de la pondération

La pondération respecte les guidelines énoncées dans OECD (2022b). Cette section présente la façon dont est conçue la pondération, tandis que son implémentation est décrite dans la section 4.3. De façon générale, la pondération suit un schéma standard qui est principalement basé sur une correction de la non-réponse suivie d'un calage. L'enquête PIAAC présente cependant la particularité que l'on tente d'identifier et de traiter spécifiquement la non-réponse liée à la thématique de l'enquête (literacy related non-response).

4.1.1 Caractéristiques et déroulement de l'enquête

Les personnes sélectionnées pour participer à l'enquête sont tout d'abord sollicitées pour remplir un questionnaire de base (background questionnaire, appelé BQ dans la suite du document). Une fois ce questionnaire rempli, les participants sont invités à prendre part à une évaluation (assessment), qui correspond au coeur de l'enquête. Des difficultés spécifiques liées à la thématique de l'enquête peuvent néanmoins être détectées lors de l'interview. Il peut s'agir par exemple de difficultés liées à la barrière de la langue (language barrier) qui ne permettent pas à la personne enquêtée de remplir le questionnaire de base, auquel cas un autre type de questionnaire lui est proposé (doorstep interview). On parle alors de « background questionnaire literacy-related non-response ». Il peut également s'agir d'un handicap mental qui permet à la personne de remplir le questionnaire de base, mais pas de participer à l'évaluation. On parle alors d'« assessment literacy-related non-response ».
Le processus d'enquête vise également à détecter les personnes inéligibles, par exemple les personnes décédées ou parties à l'étranger entre le tirage et la réalisation de l'enquête. En Suisse, deux types de cas peuvent être distingués : les cas détectés sur le terrain, et les cas détectés de façon systématique entre le tirage dans le cadre de mars 2022 et l'actualisation des adresses en août 2022, en repérant les personnes de l'échantillon brut qui n'apparaissent plus à la deuxième date.
L'échantillon brut des personnes activées peut ainsi être réparti dans les différentes catégories listées dans la table 4, qui reprend la table 2.2 de OECD (2022b). Une cellule vide signifie que le processus d'enquête ne prévoit pas de soumettre le questionnaire correspondant dans le cas de figure considéré. Par exemple, dans un cas de « nonliteracy-related non-response » au questionnaire de base, il ne fait pas de sens de proposer à la personne interrogée de participer à l'évaluation. Les catégories qui reçoivent un poids final sont les catégories R et L1 de la table 4.

- La catégorie R (6409 unités) est constituée des personnes ayant rempli le questionnaire de base, et qui ne font pas partie de l'« assessment literacy-related non-response ». Autrement dit, soit ces personnes ont participé à l'évaluation qui suit le questionnaire de base, soit leur non-participation à l'évaluation n'est pas liée à la thématique de l'enquête.
- La catégorie L1 (239 unités) est constituée des personnes qui n'ont pas rempli le questionnaire de base pour des raisons liées à la thématique de l'enquête mais qui ont complété la « doorstep interview », ainsi que des personnes qui ont rempli le questionnaire de base mais qui n'ont pas participé à l'évaluation qui suit, pour des raisons liées à la thématique de l'enquête (« assessment literacy-related non-response »).

Les catégories R et L1 représentent 30.1% des 22 091 personnes de l'échantillon brut, ce qui est moins que le taux de réponse visé de 38%, cf. section 3.2.3.

4.1.2 Population cible pour l'extrapolation

Pour l'extrapolation, la population cible est la population résidante permanente de la Suisse âgée de 16 à 65 ans, au 31.12.2022, date qui correspond approximativement au milieu de

Table 4 Classification des unités de l'échantillon brut des personnes activées

Catégorie	Questionnaire de base (« background questionnaire »)	Évaluation (« assessment »)	« Doorstep interview »	Nombre de cas
R	Complet	Complet, ou « nonliteracy-related unit or item non-response »		6 409
L1	« Background questionnaire literacy-related non-response » (barrière de la langue)		Complet	239
	Complet	« Assessment literacy-related non-response » (barrière de la langue, difficulté à lire ou à écrire, troubles de l'apprentissage, handicap mental)		
L2	« Background questionnaire literacy-related non-response » (barrière de la langue, difficulté à lire ou à écrire, troubles de l'apprentissage handical mental)		Non-réponse ou non-complet	452
NR	« Background questionnaire nonliteracy-related non-response »			13 428
U (éligibilité inconnue, par exemple personne cible injoignable)				1 302
D (non participation due à un handicap, par exemple déficience visuelle ou auditive, etc.)				93
I (inéligibles) dont : « 25_1 » (détectés de façon systématique sur l'ensemble de l'échantillon brut) « 25_2 » (détectés sur le terrain				168 dont : 101 67
				22 091

la réalisation de l'enquête sur le terrain. Afin d'obtenir pour les besoins de la pondération un cadre correspondant au mieux à cette population cible, on utilise le cadre de tirage de PIAAC du 31.03.2022, que l'on complète avec des informations auxiliaires dont la date de référence est la plus proche possible du 31.12.2022. Ces informations auxiliaires sont ainsi extraites du « cadre de calage » de l'ESPA (Enquête Suisse sur la Population Active) du premier trimestre 2023, dont la date de référence est le 31.12.2022, cf. Massiani (2021) pour une description de ces

données. Elles contiennent notamment des informations sur les revenus des personnes actives (données AVS). Une très faible proportion d'unités du cadre de tirage de PIAAC du 31.03.2022 ne sont pas retrouvées dans le cadre de calage de l'ESPA mentionné ci-dessus. Pour ces unités, les informations auxiliaires sont recherchées dans le cadre de calage de l'ESPA du deuxième trimestre 2022, dont la date de référence est le 31.03.2022, comme le cadre de tirage de PIAAC (les informations auxiliaires obtenues sont alors plus anciennes que pour les unités retrouvées dans le cadre de calage du premier trimestre 2023). A l'issue de cette procédure, les informations auxiliaires sont trouvées pour toutes les unités du cadre de tirage de PIAAC.
Le résultat de ces appariements est appelé dans la suite du document « cadre de calage » de PIAAC. Son contenu est décrit plus en détails dans la section 4.2. Il est utilisé pour calculer les totaux de calage, ainsi que pour la correction de la non-réponse. A cet effet, les informations auxiliaires contenues dans le cadre de calage sont ajoutées aux données relatives à l'échantillon brut des personnes activées grâce à un appariement entre cet échantillon et le cadre de calage.

4.1.3 Poids initial et correction pour la non-éligibilité

Nous commençons par supprimer de l'échantillon brut des personnes activées les 101 unités inéligibles détectées de façon systématique sur cet échantillon (catégorie 25_1 de la table 4). Par souci de simplification, nous désignons dans la suite du document par échantillon brut l'échantillon de 21 990 personnes ainsi obtenu, et le dénotons par s. Pour toute unité k de l'échantillon brut s appartenant à une strate h, les poids initiaux sont alors donnés par :

$$w_k^{init} = \frac{n_h}{N_h}, \tag{1}$$

où on rappelle que n_h désigne la taille (effective) de l'échantillon brut s dans la strate h, tandis que N_h désigne la taille de la strate h.
Le travail sur le terrain permet d'identifier d'autres personnes inéligibles (catégorie 25_2 de la table 4), mais pas sur la totalité de l'échantillon brut s, de sorte que des personnes inéligibles peuvent encore se trouver dans la catégorie U de la table 4. Ceci nécessite un ajustement des poids initiaux basé sur le principe suivant.

- Une proportion

$$p_E = \frac{S_R^{init} + S_{NR}^{init} + S_{L1}^{init} + S_{L2}^{init} + S_D^{init}}{S_R^{init} + S_{NR}^{init} + S_{L1}^{init} + S_{L2}^{init} + S_D^{init} + S_{25_2}^{init}} \tag{2}$$

  de la catégorie U est considérée comme éligible, où

$$S_R^{init} = \sum_{k \in s_R} w_i^{init}, \tag{3}$$

  s_R désignant la partie de l'échantillon brut correspondant à la catégorie R de la table 4, avec des notations analogues pour les autres catégories. Afin de tenir compte du fait que seulement une partie de la catégorie U est considérée comme éligible, les poids initiaux w_k^{init} des unités de la catégorie U sont diminués en les multipliant par un facteur $F_{1,U} = p_E$.
- Une proportion $(1 - p_E)$ d'unités de la catégorie U est considérée comme inéligible et leurs poids initiaux sont redistribués sur la catégorie I, en appliquant un facteur d'ajustement $F_{1,I}$ aux unités de la catégorie I. Pour les unités inéligibles de la catégorie I, le facteur $F_{1,I}$ doit vérifier l'équation :

$$(1 - p_E) S_U^{init} + S_{25_2}^{init} = F_{1,I} \cdot S_{25_2}^{init}, \tag{4}$$

  Il vient alors, pour toute unité de la catégorie I :

$$F_{1,I} = \frac{S_R^{init} + S_{NR}^{init} + S_{L1}^{init} + S_{L2}^{init} + S_D^{init} + S_{25_2}^{init} + S_U^{init}}{S_R^{init} + S_{NR}^{init} + S_{L1}^{init} + S_{L2}^{init} + S_D^{init} + S_{25_2}^{init}}. \tag{5}$$

- Le poids des autres catégories que U et I reste inchangé à cette étape.

On aboutit ainsi pour toute unité k de s au facteur d'ajustement $F_{1,k}$ suivant, pour tenir compte de l'éligibilité inconnue :

$$F_{1,k} = \begin{cases} \dfrac{S_R^{init} + S_{NR}^{init} + S_{L1}^{init} + S_{L2}^{init} + S_D^{init} + S_{25_2}^{init} + S_U^{init}}{S_R^{init} + S_{NR}^{init} + S_{L1}^{init} + S_{L2}^{init} + S_D^{init} + S_{25_2}^{init}} & \text{si } k \text{ appartient à la catégorie I,} \\ \dfrac{S_R^{init} + S_{NR}^{init} + S_{L1}^{init} + S_{L2}^{init} + S_D^{init}}{S_R^{init} + S_{NR}^{init} + S_{L1}^{init} + S_{L2}^{init} + S_D^{init} + S_{25_2}^{init}} & \text{si } k \text{ appartient à la catégorie U,} \\ 1 & \text{si } k \text{ appartient aux catégories R, NR, L1, L2, et D.} \end{cases} \tag{6}$$

Il correspond au facteur F_1 de la table 2.1 de OECD (2022b). On désigne par w_k^{UE} les poids ajustés selon ce facteur, qui sont appelés poids de base :

$$w_k^{UE} = w_k^{init} \cdot F_{1,k}. \tag{7}$$

A noter que cet ajustement a peu d'impact compte tenu du faible nombre d'unités inéligibles, cf. table 4. Conformément aux guidelines de OECD (2022b), l'ajustement est réalisé dans les cellules utilisées pour la correction de la non-réponse, cf. section 4.1.4.

4.1.4 Correction de la non-réponse

Les poids w_k^{UE} des répondants de la catégore R de la table 4 sont ajustés pour la non-réponse en les multipliant par l'inverse d'une estimation de la probabilité de réponse de cette catégorie, c'est-à-dire par un facteur :

$$F_{2,k} = \frac{S_R^{UE} + S_{NR}^{UE} + S_D^{UE} + S_U^{UE}}{S_R^{UE}}, \tag{8}$$

où

$$S_R^{UE} = \sum_{k \in s_R} w_k^{UE}, \tag{9}$$

avec des notations analogues pour les autres catégories. Ce facteur correspond au facteur F_3 de la table 2.1 de OECD (2022b) [3]. Les catégories D et U sont assimilées aux non-répondants de la catégore NR (les poids w_k^{UE} des unités de la catégorie U ayant été préalablement corrigés pour tenir compte du fait que des unités inéligibles se trouvent potentiellement dans la catégorie U, cf. section 4.1.3). En pratique, afin que la correction de la non-réponse soit plus performante et réduise le risque de biais, l'ajustement est réalisé au sein de cellules construites pour qu'au sein de chacune d'entre elles, les probabilités de réponse soient le plus homogènes possible. Elles sont créées sur la base de variables auxiliaires corrélées avec le mécanisme de réponse, cf. section 4.3.

Les poids w_k^{UE} des unités k de la catégorie L1 de la table 4 sont ajustés pour la « literacy related non-response » en les multipliant par le facteur :

$$F_{3,k} = \frac{S_{L1}^{UE} + S_{L2}^{UE}}{S_{L1}^{UE}}. \tag{10}$$

3. A noter qu'il ne figure pas de facteur F_2 dans la table 2.1 de OECD (2022b).

Ce facteur correspond au facteur F_4 de la table 2.1 de OECD (2022b). Compte tenu du relativement faible nombre de cas dans les catégories L1 et L2, cf. table 4, il est appliqué directement sur l'ensemble des catégories L1 et L2, sans recourir à une subdivision en cellules d'ajustement. Finalement, le facteur d'ajustement pour la non-réponse donné par les formules (8) et (10) se résume par :

$$F_k^{NR} = \begin{cases} \dfrac{S_R^{UE} + S_{NR}^{UE} + S_D^{UE} + S_U^{UE}}{S_R^{UE}} & \text{si } k \text{ appartient à la catégorie R} \\ \dfrac{S_{L1}^{UE} + S_{L2}^{UE}}{S_{L1}^{UE}} & \text{si } k \text{ appartient à la catégorie L1} \\ 1 & \text{si } k \text{ appartient à la catégorie I} \\ 0 & \text{si } k \text{ appartient aux catégories NR, D, U, L2.} \end{cases} \tag{11}$$

le facteur étant appliqué au sein de cellules d'ajustement pour les unités de la catégorie R. On dénote par w_k^{NR} les poids corrigés pour la non-réponse selon ce facteur :

$$w_k^{NR} = F_k^{NR} \cdot w_k^{UE}. \tag{12}$$

Ces poids sont nuls pour les unités des catégories NR, D, U et L2. Leur somme fournit ainsi une approximation de la taille du cadre de tirage.

4.1.5 Calage

Un calage est ensuite appliqué aux poids des unités de l'union des catégories R et L1, que l'on dénote par $s_{R,L1}$. Selon les guidelines de OECD (2022b), les totaux de calage doivent inclure au minimum le croisement de groupes d'âge avec le genre. Nous avons également recherché d'autres variables de calage potentiellement en lien avec la thématique de l'enquête afin d'améliorer la précision des estimations. Ces variables auxiliaires peuvent être issues du cadre de calage introduit à la section 4.1.2, mais d'autres sources ont également été considérées, cf. sections 4.2 et 4.3. Les g-weights de ce calage correspondent au facteur F_5 de la table 2.1 de OECD (2022b). On désigne par w_k les poids finaux obtenus à l'issue de ce calage, qui sont définis pour toute unité k des catégories R et L1. Le trimming suivi d'un deuxième calage suggéré par OECD (2022b) (cf. facteurs F_6 et F_7 du tableau 2.1 de ces guidelines) n'a pas été implémenté pour l'enquête PIAAC en Suisse, afin de ne pas diminuer « artificiellement » la variance au détriment du biais. La dispersion des poids a néanmoins été limitée en fixant à 100 le nombre minimum d'observations dans les cellules d'ajustement utilisées pour la non-réponse, cf. section 4.3.1. Ce choix est un compromis entre le risque de biais qui pourrait résulter de cellules trop grandes, et le risque d'explosion des poids qui pourrait découler de cellules trop petites, rendant alors nécessaire un trimming des poids.

4.2 Variables auxiliaires disponibles

4.2.1 Variables auxiliaires pour la correction de la non-reponse

Une grande partie des questions couvertes par l'enquête PIAAC est corrélée au niveau de formation. Des variables auxiliaires qui sont généralement considérées comme corrélées avec le niveau de formation d'après la littérature ou d'autres enquêtes (par exemple, le revenu) ont été sélectionnées.

En outre, le taux de réponse étant assez faible, d'autre variables auxiliaires utilisées pour expliquer le mécanisme de non-réponse dans d'autres enquêtes (par exemple la région et la typologie de ménage) et pas nécessairement corrélées avec le thème de l'enquête PIAAC ont été retenues pour la modélisation du mécanisme de non-réponse de PIAAC.

Il convient de noter qu'il aurait été utile d'utiliser le niveau de formation dans la modélisation du mécanisme de non-réponse, mais cette information n'était disponible que pour l'échantillon net. Cependant, les totaux de contrôle étaient disponibles au niveau de la population. Le niveau de formation a donc été utilisé pour le calage.

Pour la liste complète des variables auxiliaires utilisées pour la correction de la non-réponse et leurs modalités voir annexe B.

Les tables 5 - 16 montrent les distributions des variables auxiliaires utilisées pour la correction de la non-réponse : pour l'échantillon brut des personnes activées, sans les inéligibles 25_1 (s), pour les statut de réponse R, NR, D, U ensemble, ainsi que pour les répondants (R).

Table 5 Non-réponse : distribution de "Classe d'âge" selon le statut de réponse.

Classe d'âge	s	$s[\%]$	$s_{R,NR,D,U}$	$s_{R,NR,D,U}[\%]$	s_R	$s_R[\%]$
16-18	961	4.4	949	4.5	317	5.0
19-24	2 077	9.5	2 034	9.6	642	10.0
25-29	2 040	9.3	1 974	9.3	596	9.3
30-39	4 570	20.8	4 426	20.9	1280	20.0
40-54	7 149	32.5	6 860	32.3	2083	32.5
55-64	4 810	21.9	4 617	21.8	1365	21.3
65	383	1.7	372	1.8	126	2.0
Total	21 990	100	21 232	100	6 409	100

Table 6 Non-réponse : distribution de "Genre" selon le statut de réponse.

Genre	s	$s[\%]$	$s_{R,NR,D,U}$	$s_{R,NR,D,U}[\%]$	s_R	$s_R[\%]$
Homme	11 073	50.4	10 729	50.5	3 188	49.7
Femme	10 917	49.7	10 503	49.5	3 221	50.3
Total	21 990	100	21 232	100	6 409	100

Table 7 Non-réponse : distribution de "Nationalité" selon le statut de réponse.

Nationalité	s	$s[\%]$	$s_{R,NR,D,U}$	$s_{R,NR,D,U}[\%]$	s_R	$s_R[\%]$
Suisse	15 697	71.4	15 408	72.6	4 966	77.5
Europe du Nord/Ouest	1 563	7.1	1 508	7.1	509	8.0
Europe du Sud	2 353	10.7	2 241	10.6	535	8.4
Autres Pays	2 377	10.8	2 075	9.8	399	6.2
Total	21 990	100	21 232	100	6 409	100

TABLE 8 Non-réponse : distribution de "Percentile de la classe de revenu" selon le statut de réponse.

Classe de revenu	s	$s[\%]$	$s_{R,NR,D,U}$	$s_{R,NR,D,U}[\%]$	s_R	$s_R[\%]$
P10	1 614	7.3	1 549	7.3	496	7.7
P20	1 452	6.6	1 402	6.6	380	5.9
P30	1 666	7.6	1 608	7.6	452	7.1
P40	1 650	7.5	1 584	7.5	432	6.7
P50	1 660	7.6	1 613	7.6	459	7.2
P60	1 569	7.1	1 524	7.2	431	6.7
P70	1 611	7.3	1 565	7.4	480	7.5
P80	1 563	7.1	1 537	7.2	503	7.9
P90	1 516	6.9	1 495	7.0	545	8.5
P100	1 571	7.1	1 522	7.17	609	9.5
Pas de revenus	5 978	27.2	5 706	26.9	1 592	24.8
Pas connu	140	0.6	127	0.6	30	0.5
Total	21 990	100	21 232	100	6 409	100

TABLE 9 Non-réponse : distribution de "Permis de séjour" selon le statut de réponse.

Permis	s	$s[\%]$	$s_{R,NR,D,U}$	$s_{R,NR,D,U}[\%]$	s_R	$s_R[\%]$
Suisse	15 697	71.4	15 408	72.6	4 966	77.5
Permis C	4 100	18.6	3 874	18.3	981	15.3
Autres	2 193	10.0	1 950	9.2	462	7.2
Total	21 990	100	21 232	100	6 409	100

TABLE 10 Non-réponse : distribution de "Grande région" selon le statut de réponse.

Region	s	$s[\%]$	$s_{R,NR,D,U}$	$s_{R,NR,D,U}[\%]$	s_R	$s_R[\%]$
Région lémanique	4 775	21.7	4 625	21.8	1 596	24.9
Espace Mittelland	4 300	19.6	4 177	19.7	1 347	21.0
Suisse du Nord-Ouest	2 472	11.2	2 356	11.1	625	9.6
Zürich	3 240	14.7	3 093	14.6	813	12.7
Suisse orientale	2 560	11.6	2 488	11.7	671	10.5
Suisse centrale	1 705	7.8	1 613	7.6	548	8.6
Tessin	2 938	13.4	2 880	13.6	809	12.6
Total	21 990	100	21 232	100	6 409	100

TABLE 11 Non-réponse : distribution de "Langue" selon le statut de réponse.

Langue	s	$s[\%]$	$s_{R,NR,D,U}$	$s_{R,NR,D,U}[\%]$	s_R	$s_R[\%]$
Allemand	12 547	57.1	12 031	56.7	3 394	53.0
Français	6 296	28.6	6 115	28.8	2 145	33.5
Italien	3 147	14.3	3 086	14.5	870	13.6
Total	21 990	100	21 232	100	6 409	100

TABLE 12 Non-réponse : distribution de "Durée d'arrivé en Suisse" selon le statut de réponse.

Durée	s	$s[\%]$	$s_{R,NR,D,U}$	$s_{R,NR,D,U}[\%]$	s_R	$s_R[\%]$
Suisse	15 697	71.4	15 408	72.6	4 966	77.5
Depuis 0 à 1 ans	454	2.1	409	1.9	95	1.5
Depuis 2 à 4	899	4.1	805	3.8	205	3.2
Depuis plus de 5 ans	4 940	22.5	4 610	21.7	1 143	17.8
Total	21 990	100	21 232	100	6 409	100

TABLE 13 Non-réponse : distribution de "Rform" (forme juridique de l'entreprise) selon le statut de réponse.

Rform	s	$s[\%]$	$s_{R,NR,D,U}$	$s_{R,NR,D,U}[\%]$	s_R	$s_R[\%]$
Chef indépendant (1)	1 497	6.8	1 449	6.8	353	5.5
Sociétées (chef salarié) (2)	10 419	47.4	10 080	47.5	3 010	47.0
Entreprise publiques (3)	750	3.4	733	3.5	265	4.1
Administrations publiques (4)	2 195	10.0	2 164	10.2	850	13.3
Non concerné (910)	6 989	31.8	6 679	31.5	1 901	29.7
Pas encore présent dans la pop. (911)	140	0.6	127	0.6	30	0.5
Total	21 990	100	21 232	100	6 409	100

TABLE 14 Non-réponse : distribution de "Typologie de Ménage (réduite)" selon le statut de réponse.

Typologie de Ménage (réduite)	s	$s[\%]$	$s_{R,NR,D,U}$	$s_{R,NR,D,U}[\%]$	s_R	$s_R[\%]$
Homme seul majeur (101)	2 086	9.5	2 035	9.6	572	8.9
Femme seule majeure (102)	1 535	7.0	1 488	7.0	435	6.8
2 adultes de genre différent (110)	4 729	21.5	4 566	21.5	1 337	20.9
2 adultes de même genre (111)	581	2.6	554	2.6	157	2.5
3 adultes ou plus (120)	1 547	7.0	1 468	6.9	419	6.5
Personnes mineures (201)	315	1.4	311	1.5	76	1.2
Homme seul avec mineur(s) (210)	404	1.8	396	1.9	107	1.7
Femme seule avec mineur(s) (211)	1 147	5.2	1 113	5.2	336	5.2
2 adultes de genre diff. + mineur(s) (220)	8 226	37.4	7 952	37.5	2 605	40.7
Autres (230)	1 377	6.3	1 306	6.2	358	5.6
Inconnu	43	0.2	43	0.2	7	0.1
Total	21 990	100	21 232	100	6 409	100

TABLE 15 Non-réponse : distribution de "Typologie de Bâtiment" selon le statut de réponse.

Bâtiment	s	$s[\%]$	$s_{R,NR,D,U}$	$s_{R,NR,D,U}[\%]$	s_R	$s_R[\%]$
Maisons individuelles (1021)	5 759	26.2	5 613	26.4	1 909	29.8
Maisons à plusieurs logements (1025)	11 862	53.9	11 402	53.7	3 285	51.3
Habitation avec usage annexe (1030)	3 650	16.6	3 521	16.6	1 025	16.0
Partiellement à usage d'habitation (1040)	532	2.4	511	2.4	138	2.2
Inconnu (910)	187	0.9	185	0.9	52	0.8
Total	21 990	100	21 232	100	6 409	100

TABLE 16 Non-réponse : distribution de "Taille SILC" selon le statut de réponse.

Taille SILC	s	$s[\%]$	$s_{R,NR,D,U}$	$s_{R,NR,D,U}[\%]$	s_R	$s_R[\%]$
Moins de 20m2 par personne (1)	968	4.4	915	4.3	237	3.7
20m2 et plus par personne (2)	20 748	94.4	20 051	94.4	6 111	95.4
Pas disponible (999)	274	1.3	266	1.3	61	1.0
Total	21 990	100	21 232	100	6 409	100

4.2.2 Variables auxiliaires pour le calage

Pour le calage, 7 variables du SRPH et de l'ESPA ont été utilisées pour construire les marges de calage. Les tables 17 - 22, et 24 montrent les distributions respectives des variables : (1) genre croisé avec âge, (2) grande région, (3) permis de séjour, (4) nationalité, (5) percentile de revenu, (6) plus haut niveau de formation achevé croisé avec langue, et (7) durée d'arrivée en Suisse, c.-à-d. depuis combien temps la personne est en Suisse. Les sommes des poids finaux vont être égales à toutes les marges de calage.

TABLE 17 Variable de calage : genre croisé avec âge.

Modalité	Marges de calage
Homme 16-25	444 315
Homme 26-35	583 621
Homme 36-50	901 408
Homme 51-65	905 125
Femmme 16-25	420 491
Femmme 26-35	569 199
Femmme 36-50	893 580
Femmme 51-65	905 033
Total	5 622 772

TABLE 18 Variable de calage : grande région de Suisse.

Modalité	Marges de calage
Régon lémanique	1 066 614
Espace Mittelland	1 209 971
Suisse du Nord-Ouest	770 095
Zürich	1 035 705
Suisse orientale	774 347
Suisse centrale	544 423
Tessin	221 617
Total	5 622 772

TABLE 19 Variable de calage : permis de séjour.

Modalité	Marges de calage
Suisses	4 043 415
Permis C	1 003 601
Autres	575 756
Total	5 622 772

TABLE 20 Variable de calage : nationalité.

Modalité	Marges de calage
Suisse	4 043 415
Europe du Nord/Ouest	444 136
Europe du Sud	494 316
Autres Pays	640 905
Total	5 622 772

Le plus haut niveau de formation achevé n'est pas disponible au niveau du cadre de sondage. Il l'est en revanche, ainsi que la langue avec laquelle il a été croisé, dans le relevé structurel, qui est une enquête annuelle basée sur un échantillon d'au moins 200 000 personnes. Moyennant quelques adaptations détaillées ci-dessous, un calage du niveau de formation relevé dans l'enquête PIAAC sur des totaux estimés sur le relevé structurel peut être envisagé.

Il convient tout d'abord d'appliquer des filtres sur l'échantillon du relevé structurel de façon à ce que la population couverte corresponde à la population cible de PIAAC. Il s'agit notamment de filtres sur l'âge et certains permis de séjour. On désigne par relevé structurel adapté à PIAAC l'échantillon ainsi construit. Nous nous basons sur le relevé structurel disponible au moment de la pondération et le plus proche de la date de référence pour PIAAC (31.12.2022, cf. section 4.1.2), qui est le relevé structurel de 2021, avec date de référence 31.12.2021.

TABLE 21 Variable de calage : percentiles du revenu.

Modalité	Marges de calage
P10	423 094
P20	384 089
P30	412 739
P40	404 598
P50	428 950
P60	416 969
P70	420 728
P80	407 482
P90	409 735
P100	411 777
Pas de revenu d'employé	1 502 611
Total	5 622 772

TABLE 22 Variable de calage : durée d'arrivée en Suisse.

Modalité	Marges de calage
Suisse	4 043 415
Depuis 0 à 1 an	123 493
Depuis 2 à 4 ans	228 818
Depuis plus de 5 ans	1 227 046
Total	5 622 772

Quatre niveaux de plus haute formation achevée sont considérés pour la pondération de PIAAC :

- école obligatoire inachevée,
- école obligatoire,
- secondaire II,
- tertiaire.

Pour les répondants ayant complété le questionnaire de base, cf. table 4, les quatre niveaux de formation découlent directement des questions posées au sein de ce questionnaire. Pour les répondants ayant participé à la "Doorstep interview", le niveau de formation est déduit du nombre total d'années d'étude par la section BILD-S, illustré dans la table 36 en annexe C. Les modalités du niveau de formation du relevé structurel doivent être agrégées de façon à correspondre aux quatre catégories considérées pour la pondération de PIAAC. Les regroupements effectués sont décrits dans l'annexe C.

Comme le montrent les CVs de la table 23, les estimations (basées sur le relevé structurel adapté à PIAAC) du nombre de personnes dans le croisement du niveau de formation avec la langue sont relativement précises, à l'exception du niveau de formation "Ecole obligatoire inachevée" en Suisse italienne. Compte tenu du lien important entre le niveau de formation et la thématique de l'enquête PIAAC, il a été décidé de tenir compte du croisement du niveau de formation et de la langue dans le calage. A cet effet, les catégories de la Suisse romande et de la Suisse italienne ont été fusionnées, donnant lieu aux estimations du nombre de personnes présentées dans la troisième colonne de la table 24. Le total de 5 632 600 apparaissant à la dernière ligne de la table 24 étant une estimation de la taille de la population PIAAC sur la base du relevé structurel adapté à PIAAC, qui diffère légèrement de la taille de cette population qui s'élève à 5 622 772. Les estimations de la troisième colonne (Nb. de personnes) doivent donc être ajustées en appliquant le ratio 5 622 772 / 5 632 600 aux totaux des modalités, afin que la somme des marges (Marges de calage) égale la taille de la population (cf. quatrième colonne de la table 24).
Nous donnons dans la table 25 le R^2 de la régression sur les variables de calage pour trois variables clés. Les variables clés considérées sont calculées à partir de variables relevées lors de l'enquête et synthétisent les compétences des personnes interrogées en termes de scores. Il

TABLE 23 Répartition du plus haut niveau de formation achevé en fonction de la langue selon le relevé structurel de 2021 adapté à PIAAC, avec le coefficient de variation (CV) correspondant.

Langue	Niveau de formation	Nb de personnes	CV (en %)
Suisse alémanique et romanche	École obligatoire inachevée	93 581	1.9
	École obligatoire	557 013	0.7
	Secondaire II	1 714 386	0.4
	Tertiaire	1 655 279	0.4
Suisse romande	École obligatoire inachevée	39 976	2.4
	École obligatoire	267 863	0.9
	Secondaire II	559 547	0.6
	Tertiaire	511 893	0.6
Suisse italienne	École obligatoire inachevée	4 123	6.8
	École obligatoire	37 423	2.1
	Secondaire II	110 981	1.1
	Tertiaire	80 535	1.3
Total		5 632 600	

TABLE 24 Variable de calage : plus haut niveau de formation achevé croisé avec langue.

Modalité	Niveau de formation	Nb de personnes	Marges de calage
Suisse alémanique et romanche	École obligatoire inachevée	93 581	93 418
	École obligatoire	557 013	556 041
	Secondaire II	1 714 386	1 711 395
	Tertiaire	1 655 279	1 652 391
Suisse romande et italienne	École obligatoire inachevée	44 099	44 022
	École obligatoire	305 286	304 753
	Secondaire II	670 528	669 358
	Tertiaire	592 428	591 394
Total		5 632 600	5 622 772

y a une variable clé pour chacun des trois champs de compétences considérés : la littératie, la numératie, et la résolution adaptative de problèmes (« Adaptive Problem Solving », APS). Les trois variables clés sont respectivement désignées par PVLIT1, PVNUM1, et PVAPS1. Les résultats présentés dans cette table illustrent la pertinence du choix des variables de calage et laisse espérer des gains appréciables sur la variance des estimations.

TABLE 25 R^2 de la régression de trois variables clés sur les variables de calage.

Variable d'intérêt	R^2
PVLIT1	0.38
PVNUM1	0.36
PVAPS1	0.35

4.3 Mise en oeuvre de la pondération

Cette section décrit la mise en oeuvre de la pondération. La procédure de pondération garantit que les estimations se réfèrent à la population cible et réduit ainsi le risque de biais dû à la non-réponse.

Pour le calcul des poids d'échantillonnage et d'extrapolation, les guidelines du document "Part II of the PIAAC Sampling Plan on weighting and variance estimation" OECD (2022b) ont été suivies comme décrit sous 4 à l'aide des variables énoncées sous 4.2.
L'utilisation de la même procédure de pondération pour tous les pays participant à PIAAC permet d'obtenir des estimations et des estimations de précision comparables d'un pays à l'autre.

L'OCDE a mandaté Westat pour calculer les poids d'extrapolation pour tous les pays participants. Toutefois, afin de garantir la protection des données, l'Office Fédéral de la Statistique a renoncé à ce service et a calculé ses propres poids d'extrapolation pour la Suisse.

Avant de procéder à la pondération les variables envisagées pour les ajustements de la pondération ont été évaluées. Les variables utilisées pour les ajustements liés à la non-réponse doivent être disponibles pour toutes les unités éligibles, et les variables utilisées pour les ajustements liés au calage doivent être disponibles pour tous les répondants au BQ (R) et pour les non-répondants (non-réponse liée à la literatie) qui ont répondu à l'enquête de porte-à-porte (L1). Pour l'ajustement par calage, toutes les variables sélectionnées. Ces variables ne doivent pas avoir plus de 5% de valeurs manquantes.

4.3.1 Ajustement de la non-résponse

Les potentielles variables d'ajustement pour la non-réponse proviennent du cadre de calage introduit à la section 4.2.
D'abord des groupes homogènes pour la probabilité de réponse (rhg - 'response homogeneity groups' en anglais) ont étés créés à l'aide de la macro SAS CHAIDBIN, basé sur un algorithme CHAID, 'chi-square automatic interaction detection', que pour les répondants (R). L'algorithme de segmentation divise ainsi les observations en sous-groupes homogènes par rapport à la propension à répondre.
Les paramètres utilisés étaient :

1. Nombre pondéré minimum de personnes dans un rhg : lim_pond = 100,
2. Nombre non-pondéré minimum de personnes dans un rhg : lim_np = 100,
3. Taux de réponse minimum dans un rhg : lim_taux = 0.1,
4. Seuil utilisé pour les tests du Chi-carré : seuil = 0.05.
5. Poids de base (w^{UE}) standardisé par la moyenne des poids.

Le pourcentage de concordance était de 59% et la C-Statistique, c.f. Hosmer & Lemeshow (2000), de 0.61. Un modèle avec une C-Statistique de l'ordre de 0.6 indique un modèle non pas extraordinaire mais intrinsèquement suffisant par manque de variables auxiliaires supplémentaires. L'analyse du biais dû à la non-réponse est décrite dans la section 6.

L'algorithme de segmentation a fourni 47 cellules ou groupes homogènes de réponse (rhg), cf. figure 2 dans l'annexe D.

Le taux de réponse dans les cellules varie entre 12.7% et 50.8%, cf. figure 3 dans l'annexe D. Parmi les variables testées, la nationalité joue un rôle important. Les personnes ayant immigré de l'Europe de l'Est ou de pays non-européens (cl_nation4=1) ont généralement un plus faible taux de

réponse. Parmi les 4 cellules couvrant cette sous-population, 2 ont un taux de réponse inférieur à 20% et le minimum de 12.7% est atteint dans cette sous-population. A l'inverse, les personnes de nationalité Suisse ou immigrées du Nord, de l'Ouest ou du Sud de l'Europe (cl_nation4=0) et travaillant pour une administration publique (rform4=1) ont des taux de réponse généralement élevés (par rapport au taux de réponse moyen), et le maximum de 50.8% est atteint dans une cellule de cette sous-population. Il est à noter que le genre joue un rôle mineur dans cet arbre. La taille du ménage, le permis de séjour et la durée depuis l'arrivée en Suisse n'ont pas été sélectionnés dans cet arbre (voir l'arbre de segmentation en annexe D).

La corrélation entre le mécanisme de réponse et la nationalité pourrait être problématique, puisque ce type de non-réponse est également susceptible d'être corrélé avec le thème de l'enquête (par exemple, il pourrait être lié à la barrière de la langue, au niveau de formation,...). Il semble donc particulièrement pertinent d'utiliser la nationalité dans l'ajustement pour la non-réponse. Vu que le niveau de formation ne peut être utilisé que dans le calage (et non dans l'ajustement pour la non-réponse), la nationalité peut jouer un rôle important afin de diminuer le risque de biais dû à la non-réponse.

Les poids pour la correction de la non-réponse ont étés calculés à l'aide du taux de réponse au sein des rhg en utilisant les variables auxiliaires des tables 5 - 16.

TABLE 26 Statistiques des poids après correction de la non-réponse : w^{UE} : Unknown eligibility adjustment (sans les I), $w^{NR}(R)$: Nonresponse adjustment pour les répondants, $w^{NR}(L1)$: Nonresponse adjustment pour les L1, $w^{NR}(R, L1)$: Nonresponse adjustment.

Statistiques	w^{UE}	$w^{NR}(R)$	$w^{NR}(L1)$	$w^{NR}(R, L1)$
N. unitées	21 923	6 409	239	6 648
Somme des poids	5 580 859	5 388 924	191 935	5 580 859
Moyenne des poids	254.57	840.84	803.08	839.48
Poids minimum	34.88	75.15	201.70	75.15
Poids maximum	374.20	2 850.35	957.98	2 850.35
CV des poids	33.95	52.16	19.91	51.43

On constate tout d'abord que la correction de la non-réponse génère une augmentation de la moyenne des poids, en adéquation avec le taux de réponse de 30.1%, voir table 27. Si la moyenne des poids corrigés pour la non-réponse est similaire pour les catégories R et L1, la dispersion de ces poids est en revanche plus élevée pour la catégorie R que pour la catégorie L1. Il en va de même pour le poids maximum. Cela vient du fait que pour la catégorie R, la correction de la non-réponse est réalisée au sein de cellules d'ajustement dont les taux de réponse varient de 12.7% à 50.8%, tandis que pour la catégorie L1, une seule cellule est utilisée du fait du nombre restreint d'observations.

4.3.2 Calage

L'étape suivante de la pondération consiste à ajuster les poids pour qu'ils correspondent aux totaux des marges du cadre de calage.
Les variables de calage décrites dans la section 4.2, tables 17 - 22, et 24 : (1) genre croisé avec âge, (2) grande région, (3) permis de séjour, (4) nationalité, (5) percentile de revenu, (6) niveau de formation croisé avec langue, et (7) durée d'arrivée en Suisse, ont été utilisées comme marges en vérifiant que le nombre d'observations était suffisamment élevé au sein des cellules qui en découlaient. Le calage a été réalisé à l'aide de la macro **CALMAR** (CALage sur MARges) dans SAS, avec la méthode du RAKING RATIO.

La table 27 montre les statistiques des poids pour la correction de la non-réponse et pour le calage.

TABLE 27 Statistiques des poids pour la correction de la non-réponse : w^{UE} : Unknown eligibility adjustment (sans les I du type 25_2), $w^{NR}(R)$: Nonresponse adjustment pour les répondants, $w^{NR}(L1)$: Nonresponse adjustment pour les L1, $w^{NR}(R, L1)$: Nonresponse adjustment, w : Poids finaux après calage.

Statistiques	w^{UE}	$w^{NR}(R)$	$w^{NR}(L1)$	$w^{NR}(R, L1)$	w
N. unitées	21 923	6 409	239	6 648	6 648
Somme des poids	5 580 859	5 388 924	191 935	5 580 859	5 622 772
Moyenne des poids	254.57	840.84	803.08	839.48	845.78
Poids minimum	34.88	75.15	201.70	75.15	65.43
Poids maximum	374.20	2 850.35	957.98	2 850.35	5 236.07
CV des poids	33.95	52.16	19.91	51.43	58.08

On constate que si le calage a un impact non négligeable sur les poids maximum, il a en revanche un effet plus modéré sur la dispersion des poids en termes de coefficient de variation.

5 Estimation de la variance

5.1 Principe

Comme recommandé dans OECD (2022b), la variance est estimée selon la méthode de Fay, qui est basée sur des réplications de l'échantillon. Il s'agit d'une extention de la méthode « Balanced Repeated Replication » (BRR). Ces deux méthodes sont décrites dans Wolter (2007), et des explications sont également fournies dans OECD (2022b).

Balanced Repeated Replication (BRR)

Le principe de la méthode BRR est le suivant :

1. Sélection d'un nombre R^* de demi-échantillons bruts qui imitent le design de l'échantillon brut complet. Le poids initial des unités de chaque demi-échantillon est multiplié par 2.
2. Application du processus de pondération à chaque demi-échantillon (correction pour l'éligibilité inconnue, correction de la non-réponse et calage). On obtient alors R^* jeux de « replicate weights ».
3. Estimation du paramètre d'intérêt $\hat{\Theta}_r$ pour chaque jeu r de « replicate weights », avec $r = 1, \ldots, R^*$.
4. Estimation de la variance sur la base de la différence entre l'estimation du paramètre sur l'échantillon complet $\hat{\Theta}$ et de celles sur les demi-échantillons $\hat{\Theta}_r$:

$$\mathrm{var}\left(\hat{\Theta}\right) = \frac{1}{R^*} \sum_{r=1}^{R^*} \left(\hat{\Theta} - \hat{\Theta}_r\right)^2. \tag{13}$$

Afin d'éviter un nombre trop élévé R^* de jeux de « replicate weights », les demi-échantillons sont sélectionnés à l'étape 1 selon la procédure suivante (ici $R^* = 80$, cf. section 5.2).

1.a Création de « variance strata » et de « variance units ».
Il s'agit de diviser l'échantillon brut complet en H' groupes appelés « variance strata », eux même sub-divisés en deux « variance units ». Un sous-groupe de type « variance unit » comporte donc plusieurs unités. Par exemple, si un échantillon brut de taille $n = 100$ est divisé en 10 variance strata, chaque variance unit comporte 5 unités. A noter que le nombre H' de variance strata doit être inférieur ou égal au nombre R^* de jeux de replicate weights. Les sous-groupes de type « variance strata » et « variance unit » doivent être créés alétoirement de façon à imiter le design de l'échantillon brut complet. On procède ici de la façon suivante (variante 2 de OECD (2022b)), séparément dans chaque strate h, cf. exemple de la table 28.
   - Faire en sorte que l'ordre dans lequel les unités sont listées dans la strate soit aléatoire.
   - Grouper deux unités consécutives par paires.
   - Numéroter les H' premières paires de 1 à H', et faire de même avec les suivantes. Cela permet d'identifier les variance strata.
   - Pour chaque paire, attribuer aléatoirement avec une probabilité de 1/2 le chiffre 1 à l'une des unités de la paire. L'autre unité reçoit le chiffre 2. Les unités ayant reçu le même chiffre au sein d'une même variance strata consituent une variance unit.

   Ainsi dans l'exemple de la table 28, la variance strata $h' = 1$ est constituée des unités 1, 2, 21, 22, 41, 42, 61, 62, 81, 82, 91, 92 de la strate 1, et des unités 101 et 102 de la strate 2. Au sein de cette variance strata, l'unité 1 appartient à la variance unit 1, tandis que l'unité 2 appartient à la variance unit 2, et ainsi de suite [4].

4. Si une strate compte un nombre impair d'unités, alors l'une des variance units d'une des variance strata contient une unité de plus que l'autre variance unit de la même variance strata.

TABLE 28 Exemple de création de $H' = 10$ variance strata pour un échantillon divisé en deux strates

Strate	Unit	Variance stratum	variance unit
	1	1	1
	2	1	2
	3	2	1
	4	2	2
	5	3	2
	6	3	1
	⋮	⋮	⋮
	19	10	2
1	20	10	1
	21	1	1
	22	1	2
	23	2	2
	24	2	1
	⋮	⋮	⋮
	97	9	2
	98	9	1
	99	10	1
	100	10	2
	101	1	2
	102	1	1
	103	2	1
2	104	2	2
	⋮	⋮	⋮
	115	8	2
	116	8	1

1.b Les R^* demi-échantillons sont sélectionnés de la façon suivante.

- Pour chaque demi-échantillon, une variance unit sur deux est sélectionnée dans chaque variance strata.
- Les R^* demi-échantillons sont sélectionnés en fonction d'une matrice de Hadamard. Il s'agit d'une matrice carrée de taille R^*, dont les coefficients ne sont que des 1 et des -1, et dont les lignes sont orthogonales entre elles, cf. exemple ci-dessous :

$$H_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

Les lignes correspondent aux demi-échantillons, tandis que les colonnes correspondent aux variance strata. Le r-ième demi-échantillon est sélectionné en fonction de la r-ième ligne :

— si l'élément (r, h') de la matrice de Hadamard est égal à 1, la première variance unit de la variance strata h' est sélectionnée dans le demi-échantillon r, tandis que la deuxième variance unit de h' n'est pas conservée dans le demi-échantillon.

— Si l'élément (r, h') de la matrice de Hadamard est égal à -1, c'est la deuxième variance unit de la variance strata h' qui est sélectionnée dans le demi-échantillon r.

Du fait de l'orthogonalité des lignes de la matrice de Hadamard, on parle de « balanced replication ». On rappelle que le nombre R^* de demi-échantillons doit être supérieur ou égal au nombre H' de variance strata. Si $R^* > H'$, alors seulement une partie de

la matrice de Hadamard est utilisée et la propriété de « balanced replication » n'est vérifiée qu'approximativement. En dehors des matrices de Hadamard de taille 1 et 2, la taille d'une matrice de Hadamard est nécessairement un multiple de 4. C'est la raison pour laquelle le nombre de demi-échantillons R^* est le plus petit multiple de 4 qui est supérieur ou égal au nombre H' de variance strata. Le logiciel SAS que nous avons utilisé pour sélectionner les demi-échantillons impose que $R^* > H'$. Nous avons créé ici 79 variance strata, donnant lieu à 80 jeux de « replicate weights ».

Méthode de Fay

La méthode de Fay consiste à adapter la méthode BRR de la façon suivante. Lorsqu'un demi-échantillon est sélectionné à l'étape 1, ses poids initiaux sont multipliés par un facteur $(1-K)$, où $0 \leq K < 1$, tandis que le sous-échantillon restant est également conservé, avec des poids initiaux multipliés par un facteur K. Les étapes 2 et 3 sont identiques. A l'étape 4, la variance de $\hat{\Theta}$ est obtenue par la formule suivante :

$$\mathrm{var}\left(\hat{\Theta}\right) = \frac{1}{R^* \cdot (1-K)^2} \sum_{r=1}^{R^*} \left(\hat{\Theta} - \hat{\Theta}_r\right)^2. \tag{14}$$

La méthode de BRR correspond donc à la méthode Fay avec $K = 0$. Ici le facteur K est choisi égal à 0.3, conformément aux guidelines de OECD (2022b).
Une fois que les 80 jeux de replicates weights sont créés, ils peuvent être livrés aux utilisateurs qui sont alors en mesure de calculer la variance de n'importe quel paramètre de façon très simple en appliquant la formule (14). Les méthodes de Fay et BRR présentent de plus un avantage important du point de vue de la protection des données. Elles permettent en effet de tenir compte du plan de sondage et de différents aspects de la pondération tels que la correction de la non-réponse ou le calage, sans qu'il soit nécessaire de livrer aux utilisateurs les variables auxiliaires, parfois sensibles du point de vue de la protection des données, ayant permis d'aboutir aux poids finaux.
A titre d'illustration, une partie des résultats de la table T 3.3.1 de Widmer et al. (2024) est reproduite dans la table 29 ci-après. Ces résultats montrent notamment qu'au niveau national, le pourcentage de personnes ayant de faibles compétences en littératie peut être estimé avec un intervalle de confiance à 95% de $\pm$ 1.3%. Les intervalles de confiance s'élargissent bien sûr lorsque le pourcentage est ventilé en fonction de différentes caractéristiques, par exemple par classes d'âge. Ils restent néanmoins inférieurs à $\pm$ 5% pour la plupart des ventilations considérées dans la table T 3.3.1 de Widmer et al. (2024).

TABLE 29 Proportion de personnes ayant de faibles compétences en littératie

		Pourcentage	Bornes de l'intervalle de confiance à 95%
Total		22.3	$\pm$ 1.3
Classes d'âge	16-25 ans	10.4	$\pm$ 3.1
	26-35 ans	14.7	$\pm$ 2.6
	36-45 ans	24.3	$\pm$ 3.0
	46-55 ans	26.8	$\pm$ 3.1
	56-65 ans	31.7	$\pm$ 3.4

5.2 Analyse de l'ajustement des poids

Cette section décrit l'analyse de l'effet de l'ajustement des poids sur d'éventuels biais dus à la non-réponse. Cette analyse est requise si le taux de réponse global est inférieur à 70% et fait partie de la Non-Response Bias Analysis (NRBA). L'objectif est d'évaluer l'effet des ajustements de pondération et d'évaluer les biais potentiels sur les scores mesurant les compétences des personnes.

5.2.1 Comparaison des estimations avant et après les ajustements des poids

Pour cette analyse une variable registre qui n'a pas été utilisée dans la pondération et disponible pour tous les répondants doit être utilisée.

Pour cette analyse on a utilisé la variable "Permis de séjour". Cette variable a été utilisée dans l'arbre de segmentation pour l'ajustement de la non-réponse, mais n'a pas été sélectionnée par l'algorithme. Cela peut s'expliquer par le fait que la nationalité, corrélée au "Permis de séjour", a été sélectionnée par l'algorithme.

Le choix de cette variable était favorisé par les réflexions suivantes :
- L'objectif de cette tâche est d'étudier l'effet de l'ajustement pour la non-réponse.
- Le permis de séjour est susceptible d'être corrélé avec les variables de scores.
- En raison du taux de réponse relativement faible, on a décidé d'inclure dans la procédure de la pondération une grande partie des informations auxiliaires disponibles, afin de réduire le risque de biais.

Il s'ensuit que les variables disponibles restantes sont peu nombreuses et ne sont pas nécessairement les plus informatives.

TABLE 30 Comparaison des estimations avant et après les ajustements de pondération et p-valeur d'un T-test.

Permis de Séjour	w^{UE} (1)	$w^{UE}(R, L1)$ (2)	$w^{UE}(NR)$ (3)	p-valeur (1) vs (2)	p-valeur (1) vs (3)
Suisse	71.85	75.83	71.07	0	0.0639
Permis C	18.24	15.49	18.61	0	0.4467
Autre	9.9	8.68	10.32	0.0001	0.3115

La comparaison entre les poids de base (1) et les poids de base des répondants (2) de la table 30 montre que la non-réponse conduit à une surreprésentation des Suisses et la distribution est significativement différente selon un T-test, $p <= 0.05$. La comparaison entre les poids de base (1) et les poids ajustés pour la non-réponse (3) montre que l'ajustement pour la non-réponse corrige cet effet et que les poids ajustés pour la non-réponse reproduisent très bien la distribution du permis de séjour calculée avec les poids de base sur toutes les unités éligibles (le T-test ne montre plus de différences significatives). Ceci peut être interprété comme un signe positif pour la réduction du biais dû à la non-réponse.

5.2.2 Comparaison des estimations pondérées avec les totaux externes

Pour cette analyse on a utilisé les micro-données de la Statistique de la population et des ménages (STATPOP) basée sur les registres (date de référence : 31.03.2022), qui peuvent être appariées avec le cadre de sondage de l'enquête PIAAC. Seules les unités qui correspondent au cadre de sondage de PIAAC sont retenues pour calculer les totaux externes (toutes les unités du cadre PIAAC ont été retrouvées dans le STATPOP). Cette procédure garantit que la population cible correspond exactement à la population de l'enquête PIAAC.

Pour calculer le total externe ainsi que les estimations basées sur l'échantillon PIAAC, on a utilisé la variable "état civil". Ainsi, la définition de la variable d'analyse est exactement la même dans les deux sources. L'état civil n'a pas été utilisé pour la pondération, ce qui est un prérequis de cette analyse.

D'autres totaux externes provenant d'autres sources ont été envisagés dans cette analyse, mais leurs définitions n'étaient pas cohérentes avec celles utilisées dans PIAAC, ou ces totaux étaient déjà pris en compte dans le calage, comme le niveau de formation.

Selon les résultats de la table 31 on n'observe pas de différences significatives selon le T-test et $p <= 0.05$ entre les distributions de l'état civile dans PIAAC et celui de STATPOP.

TABLE 31 Comparaison des estimations pondérées avec les totaux externes.

Etat Civil	PIAAC Proportion [%]	PIAAC Ecart-type	STATPOP Proportion [%]	STATPOP Ecart-type	T-Test	p-valeur
Célibataire	42.42	0.55	42.40	0	0.041	0.967
Marié	47.00	0.61	46.92	0	0.124	0.901
Veuf	0.85	0.12	1.05	0	1.717	0.086
Divorcé	9.73	0.35	9.63	0	0.296	0.767

5.2.3 Analyse des variables recueillies lors de la collecte des données : codes liés à la littératie (literacy-related)

La non-réponse liée à la littératie pourrait être liée à un faible niveau de formation ou à une barrière linguistique. On s'attend à des taux de non-réponse liés à la littératie plus élevés pour les étrangers ou pour les personnes à faible revenu.

Pour cette analyse, nous avons utilisé une variable binaire liée au revenu "Income above P20", qui indique si le revenu de l'unité considérée est supérieur ou inférieur au 20ème percentile P20 (1 si supérieur, 0 si égal ou inférieur), ainsi que la variable binaire "Swiss" indiquant si la personne est de nationalité Suisse ou non.

TABLE 32 Distribution, non-réponse liée à la littératie (L2) et les autres non-répondants (NR).

Modalité	Proportion (L2)	Ecart-type (L2)	Estimate (NR)	Ecart-type (NR)	T-Test	p-valeur
Suisse	52.7749	2.4536	28.5632	0.3991	9.6397	0
Non Suisse	47.2251	2.4536	71.4368	0.3991	-9.6397	0
Revenu $\leq$ P20	44.7006	2.6934	33.4973	0.4616	4.1277	0
Revenu $>$ P20	55.2994	2.6934	66.5026	0.4616	-4.1277	0

La table 32 montre que le pourcentage d'étrangers est plus important pour les non-répondants liés à la littératie (L2) que pour les autres non-répondants (NR). De même, le pourcentage de personnes dont le revenu est inférieur au 20ème percentile est plus élevé chez les non-répondants que dans la population et il est même beaucoup plus élevé chez les non-répondants liés à la littératie. Les différences entre les deux types de non-répondants sont clairement significatives comme le montre le T-test avec des valeurs p > 0.05.

Ces résultats confirment qu'il était judicieux d'utiliser un questionnaire raccourci (interview "doorstep") pour les personnes ayant des difficultés liées à la littératie.

On peut ainsi s'attendre à ce que les répondants à ce questionnaire spécial (L1) soient comparables aux non-répondants liés à la littératie (L2), puisque les deux ont eu des problèmes liés à la littératie, tandis que les autres non-répondants ont une autre composition par rapport aux deux variables analysées.
La procédure de pondération proposée par l'OCDE (ajustements séparés pour la non-réponse liée à la littératie et pour les autres non-réponses, suivis d'un calage des unités répondantes) semble particulièrement pertinente.

L'analyse de l'ajustement des poids pour d'éventuels biais dus à la non-réponse n'a montré aucun biais potentiel dans les estimations finales pour les variables considérées dans cette section.

6 Étude du biais lié à la non-réponse pour les variables mesurant les compétences

Comme recommandé par OECD (2023), nous avons procédé à une analyse du biais lié à la non-réponse pour les variables mesurant les compétences. Nous en présentons ici les principaux éléments. Nous nous sommes concentrés pour cette analyse sur trois variables clés qui synthétisent les compétences des personnes interrogées en termes de scores pour la littératie, la numératie, et la résolution adaptative de problèmes (« Adaptive Problem Solving », APS). Les trois variables clés sont respectivement désignées par PVLIT1, PVNUM1, et PVAPS1.

6.1 Comparaison des compétences en fonction du niveau d'effort

Nous nous intéressons ici à la moyenne des variables PVLIT1, PVNUM1, et PVAPS1, en fonction du niveau d'effort qu'il a fallu fournir pour joindre la personne interrogée. La variable retenue pour représenter le niveau d'effort est le nombre de tentatives de contact qui ont été nécessaires pour joindre la personne. Le graphique 1 représente les résultats obtenus. La courbe rouge correspond aux moyennes cumulées de la variable PVLIT1 relative à la littéracie en fonction du nombre de tentatives de contact. Par exemple, pour une valeur de 7 sur l'axe horizontal, la valeur sur l'axe vertical correspond à la moyenne de la variable PVLIT1 pour les personnes pour lesquelles au moins 7 tentatives de contact ont été nécessaires. Les courbes verte et bleue sont l'équivalent pour les variables PVNUM1, et PVAPS1. Les moyennes sont calculées avec les poids finaux. Sous l'hypothèse que les répondants les plus difficiles à joindre ont des scores similiaires aux non-répondants, des différences entre les répondants faciles ou difficiles à contacter pourraient indiquer un risque de biais. Le graphique 1 montre que les scores en

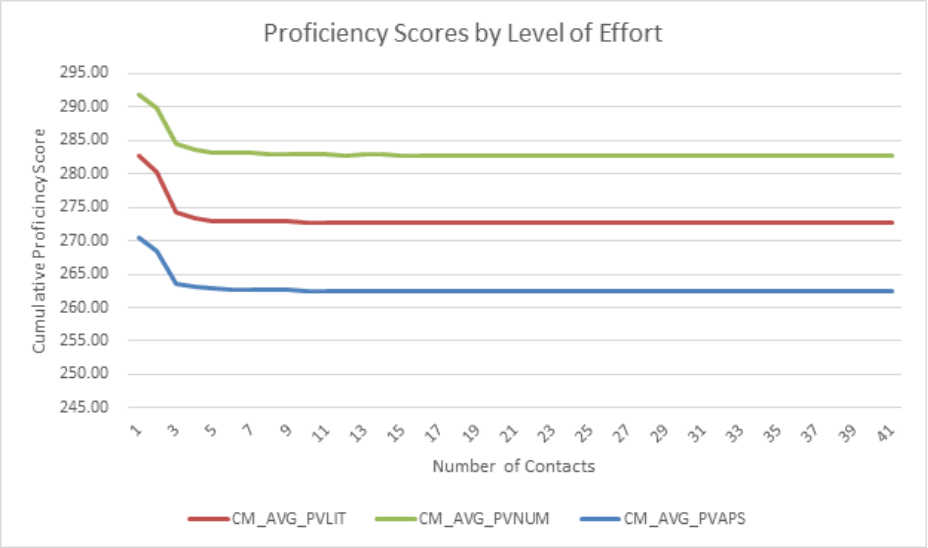

FIGURE 1 Moyennes cumulées des variables PVLIT1, PVNUM1, et PVAPS1 en fonction du nombre de tentatives de contact

littératie, numératie, et résolution adaptative de problèmes decroissent lorsque le nombre de tentatives de contacts augmente. Cela suggère qu'un risque de biais ne peut pas être exclu sous l'hypothèse formulée ci-dessus. Cependant, les différences entre les répondants faciles ou difficiles à contacter demeurent modérées, ce qui est rassurant quant au risque de biais. De plus, le recours à de multiples tentatives de contact aide à limiter le risque de biais, dans la mesure où cela permet d'intégrer à l'enquête des répondants difficiles à contacter, supposés avoir des scores similiaires aux non-répondants. Enfin, on observe que les courbes se stabilisent après 3 tentatives.

6.2 Étendue du biais potentiel

Le principe de cette analyse est de formuler des hypothèses sur les scores qu'auraient obtenus les non-répondants, puis de calculer le biais qui en résulte. Deux hypothèses sont envisagées.

- Au sein des cellules d'ajustement utilisées pour la correction de la non-réponse (rhg), les non-répondants ont tous un score égal au 40ème percentile du score des répondants (H1).
- Au sein de ces cellules, les non-répondants ont tous un score égal au 60ème percentile du score des répondants (H2).

Les résultats obtenus sont présentés dans la table 33.

Table 33 Biais potentiel

Variable	Moyenne basée sur les poids corrigés pour l'éligibilité inconnue w^{UE}	Moyenne basée sur les poids finaux w	Moyenne sous H1	Moyenne sous H2	Biais sous H1	Biais sous H2
PVLIT1	273.26	265.87	261.22	279.49	4.65	-13.62
PVNUM1	283.86	276.48	272.56	290.23	3.92	-13.75
PVAVS1	263.05	257.12	252.90	269.09	4.22	-11.97

On constate tout d'abord qu'en comparaison avec les poids corrigés pour l'éligibilité inconnue, l'utilisation des poids finaux réduit les scores moyens. Cela montre que la pondération accorde plus de poids aux répondants ayant des scores peu élevés et corrige, au moins partiellement, l'effet de la non-réponse. Les résultats de la section 6.1 suggèrent en effet que, sous l'hypothèse que les scores des non-répondants sont similaires à ceux des répondants difficiles à contacter, les non-répondants ont en moyenne des scores inférieurs aux répondants. De plus, sous l'hypothèse H1, le biais s'élèverait pour la variable PVLIT1 à 4.65, soit 1.7%, et à -13.62 sous l'hypothèse H2, soit -5.1%, avec des résultats similaires pour les variables PVNUM1 et PVAPS1. On constate donc que le risque de biais n'est pas symétrique, et qu'il est, en valeur absolue, plus élevé sous H2 que sous H1. Or les résultats de la section 6.1 montrent qu'il est plus vraisemblable que les non-répondants aient en moyenne des scores inférieurs aux répondants, ce qui est un élément rassurant quant au risque de biais.

7 Conclusions

L'enquête PIAAC est une enquête complexe dont certaines particularités ont nécessité des traitements méthodologiques appropriés, aussi bien au niveau du plan de sondage que de la pondération. Tout d'abord, en favorisant une certaine proximité géographique dans les strates, le plan de sondage a contribué à la faisabilité de l'enquête en face-à-face.
D'autre part, le taux de réponse de 30.1% a rendu cruciale la recherche de variables auxiliaires pertinentes pour la correction de la non-réponse et du calage. Des variables sociodémographiques et des variables de revenu disponibles au niveau individuel dans les cadres ont ainsi été prises en compte dans la pondération. Des variables en lien avec la thématique de l'enquête non disponibles au niveau individuel mais dont des totaux agrégés sont connus pour la population ont également été recherchées. Ceci a conduit à prendre en compte le niveau de formation dans le calage. Les R^2 de la régression sur les variables de calage pour trois variables clés de l'enquête, qui synthétisent les compétences des personnes interrogées en termes de scores pour la littératie, la numératie et la résolution adaptative de problèmes, se situent aux alentours de 36%. Cela entraîne en règle générale des gains appréciables sur la variance des estimations, qui est estimée ici sur la base de réplications des poids des unités de l'échantillon.
Enfin, l'application de ces méthodes a été complétée par une analyse du biais potentiel lié à la non-réponse. Cette analyse montre que les scores en littératie, numératie, et résolution adaptative de problèmes decroissent lorsque le nombre de tentatives de contact augmente. Cela suggère que, si les répondants les plus difficiles à joindre ont des scores similiaires aux non-répondants, un risque de biais ne peut pas être écarté. Sous l'hypothèse qu'au sein des cellules d'ajustement utilisées pour la correction de la non-réponse, les non-répondants ont un score égal au 40ème percentile du score des répondants, l'étendue du biais potentiel serait néanmoins très limitée (de l'ordre de 1.5%). Pour conclure, cette analyse a donc livré des éléments rassurants quant au risque de biais lié à la non-réponse.

Références

Hosmer, P. & Lemeshow, S. (2000). *Applied Logistic Regression*. Wiley.

Massiani, A. (2017). Estimation de la couverture du nouveau recensement en 2012. Rapport de méthodes, Office Fédéral de la Statistique, Neuchâtel. Numéro OFS : 338-0078.

Massiani, A. (2021). Enquête suisse sur la population active 2015-2020. Révision de la pondération de l'estimation de la variance. Rapport de méthodes, Office Fédéral de la Statistique, Neuchâtel. Numéro OFS : 338-0083.

OECD (2022a). CYCLE 2 PIAAC TECHNICAL STANDARDS AND GUIDELINES. Tech. rep.

OECD (2022b). PIAAC CYCLE 2 SAMPLING PLAN, PART II : PIAAC WEIGHTING AND VARIANCE ESTIMATION PLANS. Tech. rep.

OECD (2023). PIAAC CYCLE 2 SAMPLING PLAN, PART III : RESPONSE RATES AND NONRESPONSE BIAS ANALYSIS PLANS. Tech. rep.

von Siebenthal, D. & Kohler, F. (2019). Bassin d'emploi 2018 : Rapport explicatif. Office Fédéral de la Statistique, Neuchâtel. Numéro OFS : 1925-1800.

Widmer, A., Stolz, E., Bovier-Michelet, A. & von Erlach, E. (2024). Compétences en littératie, numératie et résolution de problèmes des adultes en Suisse. Premiers résultats de l'enquête PIAAC 2022/2023. Office Fédéral de la Statistique, Neuchâtel. Numéro OFS : 2334-2300.

Wolter, K. (2007). *Introduction to Variance Estimation*. New York, NY : Springer, 2nd ed.

Annexes

A Table de Notations

TABLE 34 Notations

Symbole	Description
N	Taille de la population de tirage.
H	Nombre de strates.
L	Indice des régions linguistiques.
h	Indice des strates.
k	Indice des unités.
n	Taille brute (i.e. taille de l'échantillon activé).
l	Taille des réserves.
u	Taux de réponse.
m	Taille nette (i.e. taille de $s_{R,L1}$, cf. ci-dessous).
s	Echantillon brut activé, sans les inéligibles de type « 25_1 », cf. table 4.
s_R	Ensemble des individus ayant rempli le questionnaire de base, et qui ne font pas partie de l'« assessment literacy-related nonresponse », cf. table 4.
s_{L1}	Ensemble des individus L1 (personnes qui n'ont pas rempli le questionnaire de base pour des raisons liées à la thématique de l'enquête mais qui ont complété la « doorstep interview », ou bien qui ont rempli le questionnaire de base mais qui n'ont pas participé à l'évaluation qui suit, pour des raisons liées à la thématique de l'enquête, cf. table 4.
$s_{R,L1}$	$s_R \cup s_{L1}$
s_{L2}	Ensemble des non-répondants dont la non-participation est liée à la thématique de l'enquête, cf. table 4.
s_{NR}	Ensemble des non-répondants dont la non-participation n'est pas liée à la thématique de l'enquête, cf. table 4.
s_U	Ensemble des individus dont l'éligibilité est inconnue, cf. table 4.

Symbole	Description
s_D	Ensemble des individus n'ayant pas participé à l'enquête du fait d'un handicap, cf. table 4.
s_I	Ensemble des individus détectés comme inéligibles, cf. table 4.
w_k^{init}	Poids initiaux.
w_k^{UE}	Poids ajustés pour l'éligibilité inconnue.
w_k^{NR}	Poids corrigés pour la non-réponse.
w_k	Poids finaux calés.

B Variables auxiliaires

Table 35 Liste des variables auxiliaires pour la correction de la non-réponse et leurs modalités.

Variable auxiliaire	Nb de modalité	Modalité	Source
Classe d'age	7	16-18 19-24 25-29 30-39 40-54 55-64 65	SRPH
Genre	2	Homme Femme	SRPH
Nationalité	4	Suisse Pays de l'Europe du Nord/Ouest Pays de l'Europe du Sud Autres Pays	SRPH
Revenu	11	P10 P20 P30 P40 P50 P60 P70 P80 P90 P100 Pas de revenus Pas connus	Cadre de calage ESPA
Permis	3	Suisse Permis C Autres	SRPH
Grande région	7	Région lémanique Espace Mittelland Suisse du Nord-Ouest Zürich Suisse orientale Suisse centrale Tessin	SRPH
Forme juridique	6	Chef indépendant (1) Sociétées (chef salrié) (2) Entreprise publiques (3) Administrations publiques (4) Non concerné (910) Pas encore présent dans la pop. (911)	Cadre de calage ESPA
Taille SILC	3	Moins de 20m2 par personne (1) 20ms et plus par personne (2) Pas disponible (999)	Cadre de calage ESPA
Langue	3	Allemand Français Italien	SRPH

Variable auxiliaire	Nb de catégorie	Modalité	Source
Durée	4	Suisse Depuis 0 à 1 an en Suisse Depuis 2 à 4 ans en Suisse Depuis plus de 5 ans en Suisse	Cadre de calage ESPA
Typologie de Ménage (réduite)	11	Homme seul majeur (101) Femme seule majeure (102) 2 adultes de genre différent (110) 2 adultes du même genre (111) 3 adultes ou plus (120) Personnes mineures (201) Homme seul avec mineur(s) (210) Femme seule avec mineur(s) (211) 2 adultes de genres diff. + mineur(s) (220) Autre (230) Inconnu (999)	SRPH
Bâtiment	5	Maisons individuelles (1021) Maisons à plusieurs logements (1025) Habitation avec usage annexe (1030) Partiellement à usage d'habitation (1040) Inconnu (910)	Cadre de calage ESPA

C Niveau de formation

Table 36 Niveau de formation en fonction du nombre d'années d'études pour les personnes ayant complété la "Doorstep interview"

Nombre d'années d'études (question Q04 de la "Doorstep interview")		Niveau de formation utlisées pour la pondération de PIAAC
1	Aucune année d'études	Ecole obligatoire inachevée
2	Moins de 4 années	
3	4 à 7 années	Ecole obligatoire
4	8 à 10 années	Secondaire II
5	11 à 13 années	Tertiaire
6	Plus de 13 années	

Table 37 Correspondance entre le niveau de formation selon le relevé structurel et les catégories utilisées pour la pondération de PIAAC

Q8 du relevé structurel		Niveau de formation utilisés pour la pondération de PIAAC
1	Aucune	École obligatoire inachevée
2	Ecole obligatoire inachevée	
3	Ecole obligatoire	École obligatoire
4	Formation d'un an / offre de transition	
5	Ecole de culture générale / école de degré diplôme	Secondaire II
6	Apprentissage en entreprise (CFC, AFP) / formation professionnelle élémentaire / école professionnelle / école de commerce	
7	Maturité gymnasiale / brevet d'enseignement	
8	Maturité professionnelle / spécialisée	
9	Examen professionnel avec brevet fédéral / examen professionnel supérieur avec diplôme fédéral/maîtrise	Tertiaire
10	Ecole professionnelle supérieure (ES)	
11	Bachelor : université, EPF, haute école spécialisée, haute école pédagogique (y compris diplôme HES/HEP)	
12	Master : université, EPF, haute école spécialisée, haute école pédagogique (y compris licence/diplôme université / EPF)	
13	Doctorat / habilitation	

D Groupes Homogènes de Réponse (rhg)

FIGURE 2 Arbre de segmentation et groupes homogènes de réponse (rhg).

split1	val1	split2	val2	split3	val3	split4	val4	split5	val5	split6	val6	split7	val7	split8	val8	split9	val9	split10	val10	split11	val11	split12	val12	split13	val13	rhg
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	0	cl_salary9	0	rform910	0	Typo_Men220	0	cl_salary7	0	cl_age4	0			1
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	0	cl_salary9	0	rform910	0	Typo_Men220	0	cl_salary7	0	cl_age4	1			2
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	0	cl_salary9	0	rform910	0	Typo_Men220	0	cl_salary7	1	sex1	0			3
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	0	cl_salary9	0	rform910	0	Typo_Men220	0	cl_salary7	1	sex1	1			4
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	0	cl_salary9	0	rform910	0	Typo_Men220	1	cl_salary8	0	cl_salary3	0	sex1	0	5
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	0	cl_salary9	0	rform910	0	Typo_Men220	1	cl_salary8	0	cl_salary3	0	sex1	1	6
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	0	cl_salary9	0	rform910	0	Typo_Men220	1	cl_salary8	0	cl_salary3	1			7
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	0	cl_salary9	0	rform910	0	Typo_Men220	1	cl_salary8	1					8
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	0	cl_salary9	0	rform910	0									9
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	0	cl_salary9	0	rform910	1	cl_nation1	1	cl_age6	0	cl_age5	0	Typo_Men220	0	10
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	0	cl_salary9	0	rform910	1	cl_nation1	1	cl_age6	0	cl_age5	0	Typo_Men220	1	11
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	0	cl_salary9	0	rform910	1	cl_nation1	1	cl_age6	0	cl_age5	1	Typo_Men220	0	12
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	0	cl_salary9	0	rform910	1	cl_nation1	1	cl_age6	0	cl_age5	1	Typo_Men220	1	13
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	0	cl_salary9	0	rform910	1	cl_nation1	1	cl_age6	1	Typo_Men110	0			14
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	0	cl_salary9	0	rform910	1	cl_nation1	1	cl_age6	1	Typo_Men110	1			15
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	0	cl_salary9	1											16
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	0	region6	1													17
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	1	cl_age1	0	cl_age4	0	region2	0	rform1	0							18
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	1	cl_age1	0	cl_age4	0	region2	0	rform1	1							19
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	1	cl_age1	0	cl_age4	0	region2	1									20
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	1	cl_age1	0	cl_age4	1	rform2	0									21
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	1	cl_age1	0	cl_age4	1	rform2	1									22
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	0	GKATS1021	1	cl_age1	1													23
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	1	lang3	0	cl_age5	0													24
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	1	lang3	0	cl_age5	1													25
cl_nation4	0	rform4	0	lang2	0	cl_salary10	0	cl_nation3	1	lang3	1															26
cl_nation4	0	rform4	0	lang2	0	cl_salary10	1	region4	0	rform2	0															27
cl_nation4	0	rform4	0	lang2	0	cl_salary10	1	region4	0	rform2	1	Typo_Men110	0													28
cl_nation4	0	rform4	0	lang2	0	cl_salary10	1	region4	0	rform2	1	Typo_Men110	1													29
cl_nation4	0	rform4	0	lang2	0	cl_salary10	1	region4	1																	30
cl_nation4	0	rform4	0	lang2	1	cl_nation3	0	Typo_Men220	0	cl_salary8	0	cl_salary10	0	cl_age5	0											31
cl_nation4	0	rform4	0	lang2	1	cl_nation3	0	Typo_Men220	0	cl_salary8	0	cl_salary10	0	cl_age5	1											32
cl_nation4	0	rform4	0	lang2	1	cl_nation3	0	Typo_Men220	0	cl_salary8	0	cl_salary10	1													33
cl_nation4	0	rform4	0	lang2	1	cl_nation3	0	Typo_Men220	0	cl_salary8	1															34
cl_nation4	0	rform4	0	lang2	1	cl_nation3	0	Typo_Men220	1																	35
cl_nation4	0	rform4	0	lang2	1	cl_nation3	1	cl_salary910	0																	36
cl_nation4	0	rform4	0	lang2	1	cl_nation3	1	cl_salary910	1																	37
cl_nation4	0	rform4	1	region6	0	GKATS1030	0	lang2	0	region2	0															38
cl_nation4	0	rform4	1	region6	0	GKATS1030	0	lang2	0	region2	1	sex1	0													39
cl_nation4	0	rform4	1	region6	0	GKATS1030	0	lang2	0	region2	1	sex1	1													40
cl_nation4	0	rform4	1	region6	0	GKATS1030	0	lang2	1																	41
cl_nation4	0	rform4	1	region6	0	GKATS1030	1																			42
cl_nation4	0	rform4	1	region6	1																					43
cl_nation4	1	lang1	0																							44
cl_nation4	1	lang1	1	cl_salary1	0	Typo_Men220	0																			45
cl_nation4	1	lang1	1	cl_salary1	0	Typo_Men220	1																			46
cl_nation4	1	lang1	1	cl_salary1	1																					47

FIGURE 3 Taux de réponse des groupes homogènes de réponse (rhg).

rhg	n_pond1	n_np1	tauxrep1	n_pond2	n_np2	tauxrep2	n_pond3	n_np3	tauxrep3	n_pond4	n_np4	tauxrep4	n_pond5	n_np5	tauxrep5	n_pond6	n_np6	tauxrep6	n_pond7	n_np7	tauxrep7	n_pond8	n_np8	tauxrep8	n_pond9	n_np9	tauxrep9	n_pond10	n_np10	tauxrep10	n_pond11	n_np11	tauxrep11	n_pond12	n_np12	tauxrep12	n_pond13	n_np13	tauxrep13
1	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	6448.00	6070.00	0.26	5979.76	5664.00	0.25	3911.83	3646.00	0.26	2866.66	2661.00	0.25	2462.61	2302.00	0.24	1900.02	1790.00	0.26			
2	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	6448.00	6070.00	0.26	5979.76	5664.00	0.25	3911.83	3646.00	0.26	2866.66	2661.00	0.25	2462.61	2302.00	0.24	562.58	512.00	0.20			
3	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	6448.00	6070.00	0.26	5979.76	5664.00	0.25	3911.83	3646.00	0.26	2866.66	2661.00	0.25	404.05	359.00	0.31	155.48	143.00	0.39			
4	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	6448.00	6070.00	0.26	5979.76	5664.00	0.25	3911.83	3646.00	0.26	2866.66	2661.00	0.25	404.05	359.00	0.31	248.57	216.00	0.26			
5	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	6448.00	6070.00	0.26	5979.76	5664.00	0.25	3911.83	3646.00	0.26	1045.17	985.00	0.29	900.94	857.00	0.28	759.42	721.00	0.26	406.40	372.00	0.31
6	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	6448.00	6070.00	0.26	5979.76	5664.00	0.25	3911.83	3646.00	0.26	1045.17	985.00	0.29	900.94	857.00	0.28	759.42	721.00	0.26	353.02	349.00	0.21
7	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	6448.00	6070.00	0.26	5979.76	5664.00	0.25	3911.83	3646.00	0.26	1045.17	985.00	0.29	900.94	857.00	0.28	141.52	136.00	0.36			
8	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	6448.00	6070.00	0.26	5979.76	5664.00	0.25	3911.83	3646.00	0.26	1045.17	985.00	0.29	144.22	128.00	0.38						
9	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	6448.00	6070.00	0.26	5979.76	5664.00	0.25	2067.94	2018.00	0.23	266.98	223.00	0.31									
10	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	6448.00	6070.00	0.26	5979.76	5664.00	0.25	2067.94	2018.00	0.23	1800.96	1795.00	0.22	1420.69	1415.00	0.23	1050.84	1051.00	0.25	618.12	611.00	0.22
11	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	6448.00	6070.00	0.26	5979.76	5664.00	0.25	2067.94	2018.00	0.23	1800.96	1795.00	0.22	1420.69	1415.00	0.23	1050.84	1051.00	0.25	432.72	440.00	0.29
12	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	6448.00	6070.00	0.26	5979.76	5664.00	0.25	2067.94	2018.00	0.23	1800.96	1795.00	0.22	1420.69	1415.00	0.23	369.85	364.00	0.19	213.16	215.00	0.23
13	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	6448.00	6070.00	0.26	5979.76	5664.00	0.25	2067.94	2018.00	0.23	1800.96	1795.00	0.22	1420.69	1415.00	0.23	369.85	364.00	0.19	156.69	149.00	0.14
14	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	6448.00	6070.00	0.26	5979.76	5664.00	0.25	2067.94	2018.00	0.23	1800.96	1795.00	0.22	380.26	380.00	0.17	227.28	232.00	0.14			
15	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	6448.00	6070.00	0.26	5979.76	5664.00	0.25	2067.94	2018.00	0.23	1800.96	1795.00	0.22	380.26	380.00	0.17	152.99	148.00	0.22			
16	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	6448.00	6070.00	0.26	468.24	406.00	0.34															
17	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	7583.84	6979.00	0.27	1135.84	909.00	0.34																		
18	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	3036.70	2960.00	0.32	2786.63	2709.00	0.31	2340.28	2300.00	0.32	1856.28	1912.00	0.31	1678.16	1716.00	0.32									
19	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	3036.70	2960.00	0.32	2786.63	2709.00	0.31	2340.28	2300.00	0.32	1856.28	1912.00	0.31	178.12	196.00	0.24									
20	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	3036.70	2960.00	0.32	2786.63	2709.00	0.31	2340.28	2300.00	0.32	483.99	388.00	0.37												
21	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	3036.70	2960.00	0.32	2786.63	2709.00	0.31	446.35	409.00	0.26	156.85	147.00	0.20												
22	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	3036.70	2960.00	0.32	2786.63	2709.00	0.31	446.35	409.00	0.26	289.50	262.00	0.29												
23	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	10620.54	9939.00	0.28	3036.70	2960.00	0.32	250.07	251.00	0.42																		
24	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	1033.61	1283.00	0.19	862.57	693.00	0.18	515.67	414.00	0.15																		
25	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	1033.61	1283.00	0.19	862.57	693.00	0.18	346.90	279.00	0.21																		
26	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	11654.15	11222.00	0.28	1033.61	1283.00	0.19	171.04	590.00	0.27																					
27	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	988.83	888.00	0.39	648.98	615.00	0.42	130.38	120.00	0.34																					
28	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	988.83	888.00	0.39	648.98	615.00	0.42	518.61	495.00	0.44	374.94	363.00	0.47																		
29	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	988.83	888.00	0.39	648.98	615.00	0.42	518.61	495.00	0.44	143.67	132.00	0.37																		
30	18990.41	19157.00	0.31	16898.24	17044.00	0.30	12642.98	12110.00	0.29	988.83	888.00	0.39	339.84	273.00	0.33																								
31	18990.41	19157.00	0.31	16898.24	17044.00	0.30	4255.26	4934.00	0.35	3544.66	4109.00	0.37	2237.07	2591.00	0.36	2078.06	2406.00	0.35	1932.15	2237.00	0.34	1505.78	1743.00	0.35															
32	18990.41	19157.00	0.31	16898.24	17044.00	0.30	4255.26	4934.00	0.35	3544.66	4109.00	0.37	2237.07	2591.00	0.36	2078.06	2406.00	0.35	1932.15	2237.00	0.34	426.38	494.00	0.30															
33	18990.41	19157.00	0.31	16898.24	17044.00	0.30	4255.26	4934.00	0.35	3544.66	4109.00	0.37	2237.07	2591.00	0.36	2078.06	2406.00	0.35	145.90	169.00	0.44																		
34	18990.41	19157.00	0.31	16898.24	17044.00	0.30	4255.26	4934.00	0.35	3544.66	4109.00	0.37	2237.07	2591.00	0.36	159.01	185.00	0.43																					
35	18990.41	19157.00	0.31	16898.24	17044.00	0.30	4255.26	4934.00	0.35	3544.66	4109.00	0.37	1307.60	1518.00	0.39																								
36	18990.41	19157.00	0.31	16898.24	17044.00	0.30	4255.26	4934.00	0.35	710.59	825.00	0.25	539.15	627.00	0.23																								
37	18990.41	19157.00	0.31	16898.24	17044.00	0.30	4255.26	4934.00	0.35	710.59	825.00	0.25	171.44	198.00	0.32																								
38	18990.41	19157.00	0.31	2092.17	2113.00	0.40	1900.86	1960.00	0.39	1622.23	1678.00	0.40	1185.79	1172.00	0.38	879.18	927.00	0.37																					
39	18990.41	19157.00	0.31	2092.17	2113.00	0.40	1900.86	1960.00	0.39	1622.23	1678.00	0.40	1185.79	1172.00	0.38	306.61	245.00	0.43	178.79	143.00	0.49																		
40	18990.41	19157.00	0.31	2092.17	2113.00	0.40	1900.86	1960.00	0.39	1622.23	1678.00	0.40	1185.79	1172.00	0.38	306.61	245.00	0.43	127.82	102.00	0.35																		
41	18990.41	19157.00	0.31	2092.17	2113.00	0.40	1900.86	1960.00	0.39	1622.23	1678.00	0.40	436.44	506.00	0.45																								
42	18990.41	19157.00	0.31	2092.17	2113.00	0.40	1900.86	1960.00	0.39	278.63	282.00	0.33																											
43	18990.41	19157.00	0.31	2092.17	2113.00	0.40	191.31	153.00	0.51																														
44	2183.86	2075.00	0.18	543.50	758.00	0.26																																	
45	2183.86	2075.00	0.18	1640.36	1317.00	0.15	1513.75	1215.00	0.15	888.60	713.00	0.13																											
46	2183.86	2075.00	0.18	1640.36	1317.00	0.15	1513.75	1215.00	0.15	625.15	502.00	0.17																											
47	2183.86	2075.00	0.18	1640.36	1317.00	0.15	126.61	102.00	0.26																														

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

Bianchi, B., Leuenberger, M., Massiani, A., Panchard, C. (2025). Enquête PIAAC 2022-2023. Plan de sondage, pondération, estimation de variance et analyse du biais dû à la non-réponse. Numéro OFS : 338-0091

Meskaldji, D.-E., Panchard, C., Qualité, L., Saliba, J. (2024). Enquêtes thématiques 2018-2023. Plans de sondage, tirages, pondération et estimation de variance. Numéro OFS : 338-0090

Chevalier, C. (2024). Statistique de la valeur ajoutée 2021. Révision du plan de sondage. Numéro OFS : 338-0089

Nedyalkova, D., Assoulin, D. (2024). Swiss structural business statistics (STATENT) 2015-2021. Estimation and precision of the full-time equivalents. FSO number : 338-0088

Qualité, L. (2023). Relevé structurel de la population suisse 2018. Révision de la pondération et sélections répétées. Numéro OFS : 338-0087

Potterat, J., Assoulin, D. (2023). Forschung und Entwicklung 2021. Stichprobenplan, Gewichtung und Schätzverfahren. BFS-Nummer : 338-0086

Chevalier, C. (2023). Méthodes de correction des effets calendrier en usage à l'OFS en 2023. Construction, choix et application des régresseurs mis à disposition. Numéro OFS : 338-0085

Mariétan, R. (2022). Comportement en matière de voyages 2020. Révision de la pondération. Numéro OFS : 338-0084

Massiani, A. (2021). Enquête suisse sur la population active 2015-2020 : Révision de la pondération et de l'estimation de la variance. Numéro OFS : 338-0083

Chevalier, C., Qualité, L. (2021). Enquête suisse sur la structure des salaires 2018. Révision de la pondération. Numéro OFS : 338-0082

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

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

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

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

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

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

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

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

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

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

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

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

Potterat, J. (2012). Use of conversion keys for NOGA2002-2008. FSO 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 OFS : 338-0066

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

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

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

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

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

Eichenberger P., Hulliger B., Potterat J. (2010). Describing the Anticipated Accuracy of the Swiss Population Survey. FSO 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 OFS : 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. FSO 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 OFS : 338-0057-05

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

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

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

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

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

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

Kilchmann, D. (2008). Statistik der sozialmedizinischen Institutionen 1999-2004 und Krankenhausstatistik 1999-2002. Einsetzungen für fehlende Daten. BFS-Nummer : 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 OFS : 338-0049

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

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

Gabler, S., Häder, S. (2007). Haushalts- und Personenerhebungen. Machbarkeit von Random Digit Dialing in der Schweiz. BFS-Nummer : 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 OFS : 338-0045

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

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

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

Renfer, J.-P. (2006). Enquête sur les chiffres d'affaires du commerce de détail. Plan d'échantillonnage et méthodes d'estimation. Numéro OFS : 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 OFS : 338-0040

Renaud, A. (2006). Statistique suisse des bénéficiaires de l'aide sociale. Pondération des communes 2004. Numéro OFS : 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. FSO number : 338-0038

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

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

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

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

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

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

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

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

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

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

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

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

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

Renaud, A. (2004). Analyse de données d'enquêtes. Quelques méthodes et illustration avec des données de l'OFS. Numéro OFS : 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 OFS : 338-0023

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

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

Potterat, J. (2003). Landwirtschaftliche Betriebszählung 2003. Stichprobenplan der Zusatzerhebung. BFS-Nummer : 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 OFS : 338-0019

Hulliger, B. (2003). Bereinigung der Stichprobe, Ersatz fehlender Werte und Schätzverfahren. Erhebung über F+E in der schweizerischen Privatwirtschaft 2000. BFS-Nummer : 338-0018

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

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

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

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

Renaud, A. (2002). Programme international pour le suivi des acquis des é lèves (PISA). Plans d'échantillonnage pour PISA 2000 en Suisse. Numéro OFS : 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 OFS : 338-0012

Potterat, J., Salamin, P.A. (2002). Betriebszählung 2001. Stichprobenplan und Schätzverfahren für die provisorischen Ergebnisse. BFS-Nummer : 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 OFS : 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 OFS : 338-0009

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

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

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

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

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

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

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

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

Hulliger, B., Eichenberger, P. (2000). Stichprobenregister für Haushalterhebungen : Umstellung auf Telefonnummern ohne Namen und Adressen, Abläufe für Erstellung und Stichprobenziehung. BFS-Nummer : 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 (+41 58 463 60 60) ou par e-mail (order@bfs.admin.ch).
www.statistique.ch → Statistiques → Catalogue

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
+41 58 463 60 11, info@bfs.admin.ch

L'enquête PIAAC fournit une vue d'ensemble des niveaux de compétences en lecture, calcul et résolution adaptative de problèmes parmi la population permanente de Suisse âgée de 16 à 65 ans, tout en permettant des comparaisons internationales.

L'enquête PIAAC est une enquête complexe dont certaines particularités ont nécessité des traitements méthodologiques appropriés, aussi bien au niveau du plan de sondage que des méthodes d'estimation utilisées. Le plan de sondage tient notamment compte des régions linguistiques de la Suisse afin de permettre des exploitations à ce niveau d'agrégation. D'autre part, l'enquête étant réalisée en face-à-face, le plan de sondage a contribué à sa faisabilité sur le terrain en favorisant une certaine proximité géographique dans les strates. Elle est réalisée auprès d'un échantillon brut de 22 091 personnes, avec un taux de réponse d'environ 30%.

Une attention particulière a été portée sur la recherche de variables auxiliaires pertinentes pour la pondération, et différentes sources ont été considérées informations sociodémographiques contenues dans le cadre de sondage, données AVS en lien avec les revenus, données du relevé structurel, etc. Les R^2 de la régression sur les variables de calage pour trois variables clés de l'enquête se situent aux alentours de 36%, ce qui entraîne en règle générale des gains appréciables sur la variance des estimations.

De plus, la non-réponse liée à la thématique de l'enquête (literacy related non-response) est identifiée et fait l'objet d'un traitement spécifique.

L'estimation de la variance se base sur des réplications des poids des unités de l'échantillon. Les jeux de poids ainsi produits peuvent être livrés aux utilisateurs qui sont alors en mesure de calculer la variance de n'importe quel paramètre de façon très simple, sans qu'il soit nécessaire de leur livrer les variables auxiliaires, parfois sensibles du point de vue de la protection des données, ayant permis d'aboutir aux poids finaux.

Enfin, l'application de ces méthodes a été complétée par une analyse du biais potentiel lié à la non-réponse qui, si un risque de biais ne peut pas être écarté, livre néanmoins des éléments très rassurants quant à son étendue.

En ligne
www.statistique.ch

Imprimés
www.statistique.ch
Office fédéral de la statistique
CH-2010 Neuchâtel
order@bfs.admin.ch
tél. +41 58 463 60 60

Numéro OFS
338-0091

ISBN
978-3-303-00768-6

Les informations publiées ici contribuent à mesurer la réalisation des objectifs de développement durable (ODD).

Système d'indicateurs MONET 2030
www.statistique.ch → Statistiques → Développement durable → Système d'indicateurs MONET 2030

La statistique compte pour vous.
www.la-statistique-compte.ch