

MESURE DE LA CHARGE D'ENQUÊTE À L'OFFICE FÉDÉRAL DE LA STATISTIQUE (SUISSE)

Lionel Qualité ¹

¹ *Office Fédéral de la Statistique, Espace de l'Europe 10, 2010 Neuchâtel,
lionel.qualite@bfs.admin.ch*

Résumé. L'Office Fédéral de la Statistique veut et doit mettre en place une procédure systématique de mesure du fardeau qui découle de son activité. La définition de ce fardeau et la manière de le mesurer ne sont pas encore totalement figées. Un rapide examen des pratiques des autres instituts nationaux de statistique montre qu'il n'y a pas d'unanimité sur le sujet. On tente ici de donner un aperçu des difficultés et choix qui se présentent à nous. On montre aussi comment l'utilisation d'un seul outil par lequel les échantillons de la plupart des enquêtes sont sélectionnés permet de faciliter la mesure de la charge sans toutefois apporter toutes les solutions.

Mots-clés. Fardeau d'enquête, Coordination d'échantillons.

1 Contexte

La mesure du fardeau d'enquête est un des objectifs de l'Office Fédéral de la Statistique (OFS), selon le programme pluriannuel de la statistique fédérale (voir Schütz 2016, p.15). La charge pesant sur le domaine des entreprises a été évaluée ponctuellement en 2012 dans un mandat confié à l'université de Saint-Gall (voir Müller et Bergmann 2013). D'après cette étude, le coût engendré par les statistiques d'entreprise est annuellement d'environ 9.3 millions de francs Suisse, dont 7.3 millions sont liés à des enquêtes pour lesquelles la participation est obligatoire. Récemment, deux motions parlementaires visant à exclure des relevés les entreprises de moins de 50 employés ont été acceptées par la première chambre parlementaire avant d'être rejetées par la seconde¹. La charge de travail imposée par le système statistique était mise en cause par les initiants. Si ces motions avaient été acceptées, les conséquences pour la production de statistiques d'entreprise auraient été sévères puisqu'il s'agissait d'en exclure 98% des entreprises, employant 44% des actifs occupés. Des domaines entiers d'activité auraient été absents des enquêtes. La possibilité de comparer la charge liée au système statistique avec la charge totale liée aux réglementations, évaluée à 50 milliards de francs suisses annuels (voir Savary, 2016), était essentielle.

¹voir: <https://www.parlament.ch/fr/ratsbetrieb/amtliches-bulletin/amtliches-bulletin-die-verhandlungen?SubjectId=36634>

Parallèlement, et sans que l'on ait aucune information chiffrée sur ce sujet, on ne peut que s'inquiéter de l'exaspération sensible de la population des personnes privées sollicitées téléphoniquement à leur domicile. Bien que ces sollicitations proviennent aussi et peut-être surtout d'acteurs sans aucun rapport avec la statistique publique, l'OFS ne peut se dispenser de mesurer sa contribution et de la communiquer de manière transparente. Une bonne communication peut permettre d'accroître la propension à répondre des personnes et ménages sélectionnés ou d'en limiter l'érosion. Ainsi, un investissement dans la mesure du fardeau d'enquête devrait permettre d'améliorer ou de soutenir la qualité des statistiques produites, tant dans le domaine des personnes et ménages que dans celui des entreprises et établissements.

L'OFS a engagé et poursuit des projets visant à réduire et répartir le fardeau lié à ses enquêtes : l'utilisation des données des registres quand celles-ci sont adaptées à une production statistique est systématisée (voir encore Schütz 2016). Dans certains cas, l'OFS participe à l'harmonisation des registres pour les rendre utilisables par la statistique. Cela a été le cas, par exemple, des registres cantonaux et communaux de population. Dans le domaine des statistiques sur la population et les ménages, la réforme la plus importante ayant permis de réduire le fardeau d'enquête est certainement la rénovation du système de recensement : le recensement traditionnel (visant à l'exhaustivité) décennal a été remplacé depuis janvier 2011 par l'utilisation des registres, accompagnée d'une enquête annuelle de grande ampleur (voir par exemple Graf et Qualité 2014). Chaque année, entre 3 et 4% de la population participe à cette enquête. Les sélections des personnes et des ménages sont en outre réparties au mieux en utilisant une méthode d'échantillonnage avec coordination (voir encore Graf et Qualité 2014, Qualité 2009).

Dans le domaine des enquêtes auprès des entreprises, le système de tirage coordonné est également utilisé depuis 2009 pour répartir au mieux les échantillons, sélectionner et mettre à jour les échantillons de panels et panels rotatifs. Le principal effort visant à réduire le fardeau d'enquête est probablement le projet "profiling" grâce auquel les grandes entreprises, qui sont sollicitées pour quasiment toutes les enquêtes, bénéficient d'un contact personnalisé et d'une procédure unique pour livrer toutes les informations nécessaires pour la statistique. La plus grande utilisation des données issues des registres (notamment ceux de la Centrale de compensation qui concernent les cotisations sociales des employés) est également poursuivie.

2 Problématique

L'OFS a la chance de pouvoir sélectionner les échantillons pour ses enquêtes les plus volumineuses dans deux familles de bases de sondage très riches : pour les personnes et les ménages sont utilisés des fichiers trimestriels obtenus par consolidation des registres cantonaux et communaux de population ainsi que certains registres fédéraux sur les migrants, travailleurs étrangers et demandeurs d'asile. Pour les entreprises, la base de sondage est

un extrait semestriel du registre des entreprises qui est lui mis à jour en continu. Toutes les informations des enquêtes coordonnées (probabilités de tirages, probabilités jointes de sélection à différentes enquêtes, sélections effectives) sont centralisées dans ce que l'on appelle un "système de coordination". On dispose donc de quantité d'information pour décrire les fréquences de sélection dans ces enquêtes.

Pour autant, toutes les enquêtes ne sont pas sélectionnées dans ces fichiers. On peut citer par exemple l'enquête aux frontières sur les personnes et celle sur les marchandises, qui sont réellement des enquêtes réalisées aux frontières. Elles touchent les personnes privées et les entreprises. Leurs échantillons ne sont pas ou pas nécessairement appariés avec les cadres de sondage. D'ailleurs ce ne sont pas non plus nécessairement les mêmes populations : des touristes étrangers, des transporteurs étrangers sont enquêtés. On peut également citer l'enquête sur la parahotellerie dont la base de sondage est obtenue en collectant les offres de location, l'enquête sur le transport de marchandise pour laquelle un échantillon de véhicules est sélectionné dans un registre spécifique. Ce ne sont pas, loin de là, les seuls exemples.

À contrario, un certain nombre d'échantillons sont sélectionnés dans les bases de sondage de l'OFS pour des partenaires externes et ne sont pas directement liés aux statistiques que l'OFS a reçu le mandat de produire. Par exemple, le cadre de sondage pour les enquêtes auprès de la population et des ménages est mobilisable pour tous les projets de recherche soutenus par le Fonds National Suisse de la recherche scientifique.

Il nous faut donc premièrement définir le champ des enquêtes qui entrent dans le calcul d'une mesure du fardeau d'enquête : seulement celles de l'OFS, toutes celles commandées par l'administration fédérale, les extensions d'échantillon cantonales, les enquêtes des chercheurs d'universités et hautes-écoles publiques ?

La définition du fardeau aussi doit être examinée : dans Müller et Bergmann (2013), le fardeau d'enquête des entreprises est exprimé en durée et en coût financier. Pour les enquêtes auprès des personnes et des ménages, il paraît raisonnable de chercher à exprimer le fardeau en durée d'enquête.

La mesure de ces charges n'est pas évidente et est tributaire du mode de collecte. Les enquêtes auprès des personnes et des ménages de l'OFS sont essentiellement réalisées par questionnaire papier ou téléphone avec toujours la possibilité de répondre sur internet. Pour les enquêtes par téléphone, on peut envisager de récupérer les durées d'appel enregistrées par les instituts de sondage mandatés. Ces données servent à la facturation, et des estimations sont réalisées sur des échantillons tests avant l'enquête principale. Pour les enquêtes sur questionnaire papier ou formulaire en ligne, il faudra certainement faire des estimations sur des échantillons tests, voire des prédictions à l'aide de modèles ajustés sur des données relevées. Enfin, pour les enquêtes qui ne sont pas commandées par l'OFS, il faudra organiser un retour d'information qui n'est pour l'instant pas prévu : l'OFS doit fournir les échantillons mais ne peut, à l'heure actuelle, rien exiger en retour.

La plupart des enquêtes de l'OFS auprès des entreprises sont effectuées par questionnaire papier ou par transmission électronique d'informations. Les résultats de Müller et

Bergmann (2013) sont obtenus en prédisant des fardeaux d'enquête par avis d'experts. Si l'on veut remplacer ces prédictions par un relevé statistique, on devra trouver une solution admissible par toutes les entreprises. Certaines pourraient être en mesure de nous donner une évaluation de la durée mais pas du coût et inversement. Pour certaines, ni l'une ni l'autre n'ont un sens immédiat : comment comptabiliser l'achat d'un logiciel qui permet de répondre un nombre indéterminé de fois aux demandes de l'OFS ? Il semble difficile de parvenir à une situation où la statistique soit entièrement basée sur des relevés.

Connaître le nombre de questionnaires envoyés, retournés et idéalement le fardeau associé à la réponse permet de quantifier la charge totale d'enquête. Mais ce total cache des disparités : toutes les unités ne sont pas sollicitées avec la même fréquence ni les mêmes fardeaux. Pour arriver à une description pertinente du fardeau d'enquête, il faut pouvoir le ventiler dans une population et l'étudier en regard de caractéristiques de cette population, telles la taille ou le domaine d'activité d'une entreprise, le lieu de résidence l'âge ou la nationalité d'une personne, la taille et composition d'un ménage.

La définition de la population est elle-même l'objet d'un choix. En effet, les ménages changent avec le temps, les entreprises aussi, les personnes entrent et sortent de la population, les entreprises sont créées et sont détruites. Un certain nombre de règles ont été définies pour traiter ces évolutions dans les systèmes de tirages coordonnés de l'OFS. Elles ont été choisies en fonction de l'objectif premier qui est d'étaler le fardeau d'enquête le mieux possible. S'agissant des ménages qui changent de composition, le principe retenu est de faire transiter l'historique et l'identifiant de et vers les plus grands groupes de personnes sélectionnables impliqués dans ces mutations. Les nouveaux ménages qui ne reçoivent pas d'historiques sont traités comme des nouvelles unités indépendantes. S'agissant des entreprises là aussi les historiques sont transférés autant que possible vers les nouvelles unités. Annuellement, dans quelques centaines de cas, une analyse détaillée de la structure de l'entreprise, de ses établissements et de leur taille est nécessaire pour choisir les transitions à opérer. Faut de mieux, de disposer d'un concept abouti de ménage et d'entreprise suivi dans le temps, ces transferts d'identifiants et d'historiques seront déterminants dans la mesure de la charge cumulée d'enquête.

Les choix effectués pour suivre les transformations des ménages et des entreprises dans le temps peuvent être discutés. Toutefois, la méthode de tirage des échantillons assure que le plan transversal de chaque enquête est exactement respecté. En particulier, un groupe d'unités surreprésenté dans une enquête ne risque pas d'être manquant dans l'échantillon d'une enquête ultérieure, et une unité qui hériterait *à tort* de l'historique d'une autre ne voit pas sa probabilité de sélection aux enquêtes futures faussée par cet historique.

3 Début de résolution

Notre objectif est, au cours de l'année 2016, de mettre en place un système de mesure du fardeau d'enquête. Il concernera dans un premier temps les enquêtes sélectionnées dans

les systèmes de coordination, soit environ les deux tiers des sélections dans les enquêtes considérées dans Müller et Bergmann (2013), et la quasi totalité des enquêtes auprès des personnes et des ménages. Il utilisera les données progressivement mobilisées : les indicatrices de sélection, disponibles immédiatement, d'activation (tentative de contact) et de réponse, disponibles à la fin du processus de collecte, puis les fardeaux individuels, relevés ou estimés. Les enquêtes sélectionnées hors des systèmes de coordination, et celles qui touchent des unités d'une autre nature seront intégrées dans le calcul d'un fardeau total.

Des premiers résultats sont présentés dans les tableaux 1 et 2. Il s'agit de nombres de sélections uniques et répétées dans la population des personnes et des ménages de novembre 2010 à mars 2016. Toutes les enquêtes, y compris celles externes à l'OFS, toutes les sélections, y compris pour les échantillons dits de réserve habituellement inutilisés sont comptabilisées. Les périodes 2011-2013 et 2014-2016 sont distinguées : entre 2011 et 2013, la coordination était réalisée au niveau des personnes. Depuis fin 2013 les tirages sont réalisés en deux temps : une sélection de ménages avec coordination puis une sélection d'une personne cible, sans coordination, au sein de chaque ménage. Les effectifs dans ces tableaux peuvent être rapprochés des quelques 6.5 millions de personnes sélectionnables, qui vivent dans 3.5 millions de ménages environ. N.B. : Il faut comprendre dans le

Table 1: Sélections de personnes et sélections multiples

Année	1 sél.	2 sél.	3 sél.	4 sél.
2011	524'451	13	0	0
2012	491'518	9	0	0
2013	514'095	1'078	0	0
<i>2011-2013</i>	<i>1'524'539</i>	<i>3'861</i>	<i>1</i>	<i>0</i>
2014	686'984	562	1	0
2015	815'235	810	0	0
2016	531'941	67	0	0
<i>2014-2016</i>	<i>1'975'392</i>	<i>30'543</i>	<i>205</i>	<i>0</i>
2011-2016	3'109'292	223'513	4'327	32

tableau 1 que 13 personnes ont été sélectionnées deux fois dans la seule année 2011, mais que 3'861 personnes ont été sélectionnées deux fois sur toute la période 2011-2013.

Ces statistiques peuvent être détaillées par canton de résidence, classe d'âge, taille de ménage. On constate sans surprise dans les tableaux 1 et 2 que le passage d'une coordination (négative) au niveau des personnes à une coordination (négative) au niveau des ménages a entraîné une augmentation du nombre de sélections répétées de personnes et une diminution du nombre de sélections répétées de ménages. Ce dernier est tributaire des mutations des ménages et ne peut pas être ramené à zéro : on ne saurait prévenir que deux personnes précédemment sélectionnées dans des échantillons forment un ménage commun. On constate aussi les accidents de parcours : les 13 et 9 sélections répétées de personnes

Table 2: Sélections de ménages et sélections multiples

Année	1 sél.	2 sél.	3 sél.	4 sél.	5 sél.	6 sél.	7 sél.	8 sél.
2011	470'552	9'507	95	0	0	0	0	0
2012	451'090	9'920	83	1	0	0	0	0
2013	461'208	12'409	145	1	0	0	0	0
<i>2011-2013</i>	<i>1'132'933</i>	<i>123'870</i>	<i>5'816</i>	<i>184</i>	<i>14</i>	<i>0</i>	<i>0</i>	<i>0</i>
2014	640'463	1'599	9	0	0	0	0	0
2015	769'774	2'896	16	0	0	0	0	0
2016	512'114	285	1	0	0	0	0	0
<i>2014-2016</i>	<i>1'712'568</i>	<i>92'273</i>	<i>4'218</i>	<i>152</i>	<i>5</i>	<i>1</i>	<i>0</i>	<i>0</i>
2011-2016	2'108'358	437'219	84'914	14'940	2'077	232	41	4

en 2011 et 2012 qui sont des personnes ayant quitté la Suisse au moins un trimestre puis qui sont revenues et dont l'historique n'était pas correctement suivi, les 1'078 sélections répétées de personnes en 2013 qui ont en réalité été sélectionnées indirectement par leur ménage dans un échantillon pour le panel suisse des ménages. Notre objectif dans les prochains mois est de remplacer la variable indicatrice de sélection qui est à la base du calcul par des variables de plus en plus pertinentes : indicatrice d'activation, de réponse, mesure de charge.

Bibliographie

- [1] Graf, E. et Qualité, L. (2014), Sondage dans des registres de population et de ménages en Suisse : coordination d'échantillons, pondération et imputation, *Journal de la Société Française de Statistique*, 155(4), 95–133.
- [2] Müller, C. et Bergmann, H. (2013), *Schätzung der Kosten von Regulierungen und Identifizierung von Potenzialen für die Vereinfachung und Kostenreduktion im Bereich Statistik. Forschungsbericht im Auftrag des Bundesamtes für Statistik.*, KMU-HSG, Universität St. Gallen.
- [3] Qualité, L. (2009), *Unequal probability sampling and repeated surveys*, Thèse de Doctorat, Université de Neuchâtel, Suisse.
- [4] Savary, G. (2016), Rapport de la Commission de la science, de l'éducation et de la culture du 16 novembre 2015, *Conseil des États*, Suisse.
https://www.parlament.ch/centers/kb/Documents/2015/Rapport_de_la_commission_CSEC-E_15.3433_2015-11-16.pdf.
- [5] Schütz (2016), *Programme pluriannuel de la statistique fédérale 2016-2019*, OFS, Neuchâtel.