

DÉFIS MÉTHODOLOGIQUES DANS LES MESURES DE DÉ(CHARGE) DES ENQUÊTES DE L'OFS

Jean-Pierre Renfer¹

¹ *Office fédéral de la statistique, section des méthodes statistiques, Espace de l'Europe 10, CH-2010 Neuchâtel, SUISSE. E-mail : jean-pierre.renfer@bfs.admin.ch*

Résumé. On aborde ici dans un contexte de forte demande de réduction de la charge des unités enquêtées les défis méthodologiques pour y faire face. Une première partie est consacrée à un système de coordination des échantillons qui a été mis en place aussi bien pour les enquêtes auprès des entreprises que des enquêtes auprès des personnes et des ménages. Ce système de répartition de la charge a permis de développer un dispositif de mesure de la charge des enquêtes. Les difficultés inhérentes à ce dispositif sont également décrites. Les mesures de réduction de la charge sont évidemment au centre de toutes les attentions avec les défis méthodologiques y relatifs allant de l'optimisation des plans d'échantillonnage à l'utilisation directe ou indirecte des données administratives. Pour terminer, quelques pistes sur les défis futurs dans ce domaine sont proposées et des projets pilotes en lien avec la stratégie d'innovation des données ont été lancés.

Mots-clés. Cadre de sondage, calage, coordination, données administratives, échantillon aléatoire, fardeau de réponse, imputations, mesure, plan de sondage, pondérations.

Abstract. In a context of strong demand for reducing the burden of surveyed units, the methodological challenges to face these questions are addressed. The first part is dedicated to a system of sample coordination which was implemented for the business surveys as well as for surveys of individuals and households. This burden balancing system allowed to develop a measuring device of the surveys burden. The inherent difficulties of this device are also described. The burden reduction measures are central with the corresponding methodological challenges ranging from sampling plans optimization to direct or indirect use of administrative data. To conclude, some avenues for future challenges in this field are provided and pilot projects linked to the data innovation strategy were launched.

Keywords. Administrative data, calibration, coordination, imputations, measure, random sample, response burden, sampling frame, sampling plan, weights.

1. Introduction

Il s'agit ici de décrire les mesures prises par l'Office fédéral de la statistique (OFS) pour maîtriser le fardeau de la charge de réponse imposé aux personnes et aux entreprises.

Sur ce chemin, les problèmes posés relèvent dans certains cas de véritables défis méthodologiques pour lesquels les solutions méritent d'être présentées.

Cette démarche s'inscrit dans un contexte qu'il convient de rappeler et qui fera l'objet du paragraphe suivant. Nous traiterons ensuite des mesures prises pour coordonner les échantillons, des mesures qui ont conduit l'OFS à implémenter un système de répartition de la charge aussi bien au niveau des ménages que des entreprises. Nous exposerons ensuite les travaux entrepris permettant de mesurer la charge des enquêtes de l'OFS (et des défis qu'il s'agit là aussi encore de

relever). Enfin, toute une batterie de mesures ont été prises pour réduire de manière ciblée la charge des unités enquêtées. Finalement, et en guise de conclusions, nous nous proposons de considérer quelques défis futurs qui pourraient bien apporter leur contribution pour soulager le fardeau de réponse des unités enquêtées.

2. Contexte

Le fardeau de réponse imposé aux personnes et aux entreprises est sans nul doute un sujet de préoccupation constant pour un office national de statistique. Il n'est pas rare aujourd'hui qu'une enquête statistique soit perçue comme une charge (trop) lourde en comparaison de son utilité à titre individuel. Cette perception s'est faite plus vive encore dans l'économie que dans la population, pour preuve plusieurs interpellations parlementaires ces dernières années, notamment : les postulats 10.3429 et 10.3592 « Mesure des coûts réglementaires » (déposés en 2010 et classés en 2014 et 2015 respectivement), les motions 15.3433 et 15.3439 « Libérer les entreprises de la charge statistique » (déposées en 2015 et rejetées en 2016), la motion 15.3843 « Libérer les entreprises et les citoyens de la charge statistique. Réduire les effectifs de l'office fédéral de la statistique » (déposée en 2015 et rejetée en 2017) et enfin la motion 16.4011 « Numérisation. Eviter les récolte de données en parallèle » (déposée en 2016 et adoptée en 2017).

Cette question de la (dé)charge des unités enquêtées est ainsi au centre des débats et il convenait dès lors de mettre en place un système global de mesure de la charge de toutes les enquêtes de l'OFS. Ce système exploite largement l'infrastructure mise en place pour coordonner les échantillons (aussi bien dans les enquêtes entreprises que dans les enquêtes ménages) décrite au paragraphe suivant.

3. Système de répartition de la charge

Pour ses enquêtes par échantillonnage aléatoire, l'OFS utilise principalement deux cadres de sondage. Le premier, développé pour les enquêtes auprès des personnes et des ménages, appelé SRPH (StichprobenRahmen für Personen und Haushalte) est mis régulièrement à jour (chaque trimestre actuellement) grâce aux données des registres des habitants des communes et des cantons, des registres fédéraux de personnes ainsi que du registre fédéral des bâtiments et des logements. Le second cadre permet de tirer les échantillons d'entreprises et est mis à jour deux fois par année grâce aux données du registre des entreprises et des établissements (REE) qui comprend toutes les entreprises et tous les établissements de droit privé et public établis et générant une activité économique sur le territoire suisse.

Pour ces deux cadres de sondage, le système de répartition de la charge d'enquête implémenté est le même et repose sur les travaux cités en référence (Qualité, 2009) et (Nedyalkova, Qualité et Tillé, 2011).

Pour l'OFS, le cahier des charges pour organiser ses enquêtes est le suivant :

- Respecter pour chaque enquête un plan de sondage ;
- Tenir compte des échantillons qui ont déjà été tirés ;
- Limiter autant que possible les sollicitations pour des enquêtes ;
- Contrôler l'étalement dans le temps des sélections ;
- Pouvoir sélectionner aussi bien des enquêtes ponctuelles que des panels (rotatifs) ;

Ce dernier point signifie que les échantillons doivent pouvoir être coordonnés aussi bien négativement que positivement (ce dernier aspect étant particulièrement important si l'on veut

estimer avec précision une évolution).

L'algorithme de coordination repose sur des échantillons poissonniens tirés à partir de numéros aléatoires permanents selon la démarche suivante :

1. Chaque unité du cadre de sondage reçoit un numéro aléatoire permanent u dans $[0;1]$;
2. Pour chaque enquête, on associe à chaque unité un intervalle de sélection (chaque unité étant traitée indépendamment des autres unités) de longueur égale aux probabilités d'inclusion ;
3. Une unité est tirée si son numéro aléatoire appartient à l'intervalle de sélection ;
4. La coordination (négative ou positive) s'obtient par un placement judicieux des intervalles de sélection dans $[0;1]$.

A noter que lors d'une nouvelle enquête, le sens de coordination et l'ordre de priorité avec les enquêtes précédentes doit être déterminé.

4. Mesures de la charge

Dans le contexte évoqué au paragraphe 2, l'OFS a décidé de développer un dispositif de mesure de la charge des enquêtes. De manière à obtenir un prototype relativement rapidement, il a été décidé de limiter dans un premier temps cette mesure aux enquêtes directes de l'OFS. Cet outil est avant tout conçu dans le but de mesurer, suivre et maîtriser l'évolution de la charge des enquêtes en objectivant le nombre d'unités contactées ainsi que le nombre d'unités répondantes (tout comme leur quotient !) et pour ces dernières le temps consacré à répondre à l'enquête. Selon le type d'enquêtes, ces exploitations peuvent être faites en fonction du caractère obligatoire ou facultatif des enquêtes ou encore dans des domaines particuliers (taille/activité économique pour les entreprises ou sexe/âge pour les personnes par exemple).

Le fait d'avoir mis en place un système centralisé de coordination des échantillons est un avantage indéniable pour la mise en œuvre de ce système de mesure de la charge. Cette mesure est cependant loin d'être triviale et dépend largement de l'enquête et de son mode de collecte !

S'il est relativement facile de mesurer la durée des enquêtes téléphoniques (encore nombreuses pour les enquêtes personnes/ménages), cela l'est nettement moins dans le cas des enquêtes papier ou encore par internet. Dans ces derniers cas, il est fait recours dans certains cas à des tests et souvent à des avis d'expert.

5. Mesures de réduction de la charge

Les mesures de réduction de la charge des unités concernées se fondent sur une rationalisation des relevés qui comprend les éléments suivants :

- a) L'optimisation du questionnaire ;
- b) La facilitation de la réponse par le mode de collecte ;
- c) L'utilisation systématique (directe ou indirecte) des données administratives (registres) ;
- d) La répartition la plus judicieuse de la charge.

Bien entendu, l'élaboration et les tests de questionnaires tout comme le mode de collecte y relatifs jouent un rôle important dans les mesures de réduction de la charge. Cependant, sur le plan des méthodes statistiques, le recours aux données administratives (registres) y jouent un rôle fondamental. En effet, qu'il s'agisse de l'optimisation des plans d'échantillonnages, du traitement de la non-réponse ou des méthodes d'estimation basées sur le calage, on ne compte plus les exemples d'application de réduction de la charge et/ou (à qualité égale !) d'augmentation de la qualité (à charge égale !).

Finalement, le fait de disposer d'un système de répartition de la charge, s'il ne permet pas à proprement parler de réduire la charge, il a cependant l'indéniable avantage de réduire les plaintes grâce à une communication adaptée.

6. Défis futurs

Les défis méthodologiques dans les mesures de (dé)charge des enquêtes de l'OFS sont nombreux et peuvent être synthétisés de la manière suivante :

- i. Mise en œuvre de la mesure et du monitoring de la charge ;
- ii. Méthodes statistiques : plans de sondage, pondérations, imputations, calage, ... ;
- iii. Nouvelles sources de données, y compris à l'aide d'appariements (mesure de la qualité !) ;
- iv. Applications dans la production statistique (statistique des entreprises, de la valeur ajoutée) ;
- v. Transfert de connaissances et formation continue : nouvelles méthodes (estimations sur petits domaines, apprentissage profond,...) ;
- vi. Projets pilotes à l'OFS en lien avec la stratégie d'innovation des données.

Bibliographie

Qualité, L. (2009), *Unequal probability sampling and repeated surveys*, Thèse de doctorat, Université de Neuchâtel, Suisse.

Nedyalkova, D., Qualité, L. & Tillé, Y. (2011). *Tirages coordonnées d'échantillons poissoniens*. In M.E. Trambly, P. Lavallée & M. El Haj Tirari (Eds). *Pratiques et méthodes de sondage* (pp. 323-328). Paris, Dunod.