

Introduction du CAWI dans l'enquête LU-EBM/HBS :

Défis, enjeux et solutions

Adrien CANNET-DELBOSQ

¹ STATEC, 13, rue Erasme, Luxembourg, Luxembourg

adrien.cannet-delbosq@statec.etat.lu

Résumé.

L'Enquête Budget des Ménages (EBM) est une enquête menée depuis 1977 au Luxembourg qui participe au calcul des facteurs de pondération de l'Indice des Prix à la consommation Nationale (ICPN) et permet d'établir des modèles de dépenses selon la composition et la situation financière des ménages enquêtés. Menée auprès de 5000 ménages, cette enquête est traditionnellement conduite en mode PAPI (Pen-and-Paper Personal Interview) par des enquêteurs, avant transmission et encodage manuel au STATEC. Un projet de modernisation est aujourd'hui mené, avec pour objectif d'introduire un second mode de collecte CAWI (Computer Assisted Web Interviewing) auto-administré par le biais d'un projet pilote, déployé en 2024 en production auprès de 1500 ménages afin de tester les performances du mode de collecte CAWI dans l'enquête EBM.

La présentation est un retour d'expérience, d'abord axé sur la présentation des enjeux liés à l'introduction d'un nouveau mode de collecte dans l'enquête LU-EBM, puis sur les méthodes, techniques et non-techniques, qui ont été utilisées pour répondre à certains défis opérationnels.

La stratégie sur plusieurs fronts mise en place pour répondre à ces défis sera détaillée de l'amont à l'aval du projet ;

- D'abord au niveau non-technique, notamment : analyse des besoins, cadre méthodologique (*AGILE*), recrutement, communication entre les différentes parties prenantes
- Au niveau technique, nous présenterons les outils mis en place au niveau de la collecte des données, de la gestion de l'enquête, de la gestion des bases des données en équipe pour la codification des données ainsi que pour leur validation.

Mots-clés. Enquêtes multimodales: collecte et traitement, EBM/HBS, statistiques sociales

1. Contexte :

L'Enquête Budget des Ménages (EBM) est une enquête menée depuis 1977 au Luxembourg qui participe au calcul des facteurs de pondération de l'Indice des Prix à la consommation Nationale (ICPN) et permet d'établir des modèles de dépenses selon la composition et la situation financière des ménages enquêtés.

L'enquête EBM fournit également de précieuses données permettant de mieux suivre les évolutions des habitudes de consommation des ménages pour l'ensemble des postes de dépenses (alimentation, logement, services, ...)

L'enquête est menée annuellement par le STATEC auprès de 5000 ménages résidants au Luxembourg. Cette enquête est traditionnellement conduite en mode PAPI (Pen-and-Paper Personal Interview) par des enquêteurs, avant transmission et encodage au STATEC des questionnaires.

Les ménages sont invités à renseigner l'ensemble de leurs dépenses sur des périodes allant de 15 jours à 1 an pour les dépenses les moins courantes. Cela rend l'enquête EBM particulièrement lourde pour les ménages et compliquée à gérer sur le terrain. Un projet de modernisation est aujourd'hui mené, avec pour objectif d'introduire un second mode de collecte CAWI (Computer Assisted Web Interviewing) auto-administré. L'objectif est d'utiliser les avantages du CAWI pour :

1. Proposer une alternative au PAPI
2. Augmenter le taux de participation des ménages
3. Meilleure couverture géographique du territoire National
4. Réduire la charge de travail d'encodage pour les agents du STATEC
5. Améliorer la qualité des données
6. Meilleure couverture de la population pour augmenter la représentativité

Un projet pilote a été déployé en production au début de l'année 2024 afin de tester les performances du mode de collecte CAWI dans l'enquête EBM. La fin de la collecte CAWI est prévue pour août 2024. Une fois l'analyse des performances du pilote CAWI réalisée, et, le cas échéant, des correctifs appliqués, l'objectif est de proposer à horizon 2025 l'enquête LU-EBM en multimode c'est-à-dire en modes PAPI/CAWI.

L'objectif de la présentation est de rendre compte des enjeux liés à l'introduction d'un nouveau mode de collecte dans l'enquête LU-EBM et des méthodes qui ont été utilisées pour répondre à certains défis.

2. Principaux enjeux

2.1 Enjeux liés à la représentativité et à l'introduction de la multimodalité

L'introduction du CAWI a aussi comme objectif d'augmenter la représentativité de l'enquête. En effet, proposer le PAPI sans aucune alternative a comme première conséquence l'introduction d'un biais de sélection. En effet, le PAPI requérant le déplacement à domicile d'un enquêteur, et chaque enquêteur étant assigné à une zone géographique, les résidents de certaines régions du pays moins accessibles étaient parfois sous-représentés. Le CAWI doit permettre de mieux toucher les populations résidentes de ces zones ainsi que les ménages les plus jeunes.

Par ailleurs, l'introduction d'une alternative au PAPI doit permettre à terme de diminuer la non-réponse (E. Sala, P. Lynn, 2007, E. De leeuw, 2018). Dans ce pilote, il n'a pas été proposé aux ménages de choisir leur méthode de participation. L'échantillon a été scindé en 2 sous-échantillons : un échantillon PAPI, un échantillon CAWI pilote. Il était possible pour l'enquêté de passer du CAWI vers le PAPI, sans que cette possibilité ne lui soit directement proposée. Il n'était en revanche pas possible de passer du PAPI vers le CAWI pour éviter une baisse du taux de participation PAPI, le CAWI étant en 2024 uniquement une enquête pilote.

2.2 Enjeux liés au caractère auto-administré du CAWI

L'introduction d'un nouveau mode de collecte n'est pas totalement nouveau dans le contexte de l'enquête EBM au Luxembourg. En effet, la pandémie de la COVID-19 et la nécessité d'éviter le contact physique avait conduit le STATEC à déployer un mode de convocation CATI (Computer-Assisted Telephonic Interview). L'enquêté recevait par courrier un questionnaire papier l'invitant à répondre à l'enquête, mais toujours avec une première prise de contact par un enquêteur chargé de présenter l'enquête, son but et son caractère obligatoire. L'enquêteur constituait toujours, comme dans le PAPI, un référent pro actif pour le ménage.

2.3 Améliorer l'efficacité des procédés de traitement

L'introduction du CAWI est aussi l'occasion de réduire la charge de la collecte pour les équipes de gestion par rapport au PAPI:

- Par la réduction de la charge d'encodage
- Par l'automatisation de la codification afférente à certaine nomenclature (NACE, ISCO, COICOP)

2.4 Adaptation au cadre opérationnel

La modernisation de l'enquête LU-EBM s'inscrit dans une stratégie de digitalisation plus large du STATEC. A ce titre, le Centre des technologies de l'information de l'Etat (CTIE), le bras technologique du ministère de la digitalisation Luxembourgeois, administre un guichet unique, MyGuichet, pour l'ensemble des démarches administratives nationales, incluant les enquêtes statistiques. C'est sur cette plate-forme pré-existante que l'enquête LU-EBM a été digitalisée en intégrant les exigences techniques propres à la plate-forme MyGuichet et en profitant du cadre et de l'expertise proposé par le CTIE.

Cette mise en place d'un nouveau mode de collecte CAWI a nécessité la refonte des étapes de validation/d'imputation des données, ce qui a posé des défis d'intégration de ces nouveaux procédés avec les processus déjà existants dans les workflows de l'enquête EBM traditionnelle. La mise en place de ces nouveaux processus se sont faits en parallèle du maintien de l'enquête EBM PAPI qui a continué à être proposée aux enquêtés avec l'ensemble de ses propres procédés de gestion d'enquête et de validation.

3. Une stratégie sur plusieurs fronts pour répondre à ces défis

Cette présentation sera l'occasion d'aborder les moyens techniques et non-techniques qui ont été mis en œuvre pour répondre à ces défis.

Au niveau non-technique, nous montrerons comment le recrutement d'un analyste métier (*business-analyst*) dès la phase de conception du projet a permis de définir au mieux les besoins, les contraintes techniques et les ressources nécessaires à la réalisation du projet.

Cette analyse de besoin a aussi été l'occasion de mettre en place un cadre méthodologique (*AGILE*) centré sur la communication entre les différentes parties prenantes (chef de projet, donneur d'ordre, statisticiens, développeurs informatiques, ...). Au-delà du développement

de savoir-faire techniques pour mettre en place des solutions techniques répondant aux besoins métiers, les *softs-skills* ont été au centre de la réflexion. L'intégration progressive des développeurs informatiques et de nouvelles techniques dans les processus métier et le travail en équipe sur des éléments les plus complexes ont notamment été à la base de notre stratégie.

Au niveau technique, nous présenterons les outils mis en place au niveau de la collecte des données, de la gestion de l'enquête, de la gestion des bases des données par le métier, pour la codification des données ainsi que pour leur validation.

Cela a permis de travailler sur plusieurs fronts :

1. La numérisation de la collecte des données sur le guichet unique MyGuichet,
2. Le développement d'outils de gestion d'enquêtes adaptés au CAWI
3. La mise en place de nouveaux processus d'imputation/de validation des données

tout en étant capable de continuer à gérer l'enquête EBM PAPI.

Conclusion

Cette présentation sera aussi l'occasion de présenter en conclusion certains indicateurs relatifs aux performances de ce pilote, notamment le taux de participation ventilé par âge, nationalité, par mode de convocation.

Les étapes de validation des données

En conclusion : le pilote a été mis en Production et les ménages ont été invités. Chiffres : taux de participation, taux de participation/modes, taux de participation/ménage, /canton /age /nationalité Ouverture : analyses qualité des données, effets de modes,

Bibliographie

E. Sala, P. Lynn (2007). The potential of a multi-mode data collection design to reduce non response bias. The case of a survey of employers, Springer Science+Business Media B.V. 2007

De leeuw, Edith. (2018). Mixed-Mode: Past, Present, and Future. Survey Research Methods. 12. 75-89. 10.18148/srm/2018.v12i2.7402.