

La collecte par internet est-il l'avenir des enquêtes Génération du Céreq ?

C. Barret - C. Dzikowski

Le dispositif Génération a pour objectif l'étude des parcours d'insertion professionnelle des jeunes débutants. À cet effet un panel de jeunes sortants du système éducatif est interrogé par téléphone à 3, 5 et 7 ans après leur entrée sur le marché du travail. Ce dispositif permet de rendre compte des interactions entre la formation initiale et le parcours professionnel. Une expérimentation de collecte par internet (Cawi : Computer-Assisted Web Interviewing) a été menée en parallèle de l'enquête 2013 auprès de la Génération 2010, sur un échantillon disjoint de celui de la collecte principale par téléphone (Cati : Computer-Assisted Telephone Interviewing).

Le questionnaire Cawi est, à quelques modifications marginales, identique au questionnaire Cati. Les abandons en cours de collecte plus importants pour la collecte internet sont vraisemblablement liés au manque d'ergonomie du questionnaire web, à sa longueur mais aussi de manière plus marquée à certaines questions difficiles, telles que celles comportant des consignes longues.

A l'aide d'un calcul de précision, une comparaison des structures des réponses Cawi et Cati sur un champ comparable nous permet de conclure à une égalité statistique sur 40% des variables du questionnaire. En revanche, 30% des variables ne présentent aucune égalité statistique sur aucune des modalités. Les différences observées sont vraisemblablement liées à un effet de sélection et/ou à un effet de mesure.

Pour essayer de distinguer ces deux effets, une méthode de matching, habituellement utilisée dans l'évaluation de dispositifs publics, a été mobilisée. Ces méthodes permettent d'estimer les effets d'un dispositif, dans le cas d'expérimentations non préalablement contrôlées, en comparant les mesures faites sur les individus traités et leurs contrefactuels non traités ayant des caractéristiques similaires avant l'accès au dispositif. Dans le cas présent, le dispositif étudié est la passation par internet. Les contrefactuels Cati d'un individu répondant par Cawi sont déterminés par la proximité de leur probabilité de répondre par internet avec celle de l'individu Cawi. Les écarts constatés entre le groupe traité (Cawi) et celui de leur contrefactuel (Cati) sont interprétés ici comme de l'effet de mesure. Ces méthodes présentent cependant une limite importante. Les effets de mesure isolés à partir de cette méthode dépendent de la spécification du modèle corrigeant la sélection.

Face aux difficultés rencontrées pour distinguer les effets de sélection des effets de mesure, d'autres stratégies seront mises en œuvre lors de la prochaine expérimentation. D'une part des questions permettant de mieux prendre en compte la propension des répondants à utiliser internet seront posées afin d'améliorer la spécification du modèle de sélection dans le cadre de la méthode de matching. D'autre part, un échantillon embarqué avec affectation aléatoire du mode de collecte aux individus qui acceptent de répondre à l'enquête sera mis en place afin de limiter l'effet de sélection et d'évaluer l'effet de mesure plus directement. L'étude présente une réflexion sur les stratégies d'agrégation des données Cati et Cawi en présence de biais de mesure.