

Les besoins des économistes en matière d'appariements sécurisés

Nathalie PICARD¹, Benoît RIANDEY² et Anne SOLAZ²

Les économètres peinent à la recherche de données longitudinales. Les panels sont rares en France malgré les avancées récentes appuyées par le rapport au Cnis de Stéfan Lollivier et Mylène Chaleix. Aussi sont-ils souvent conduits à se limiter aux panels étrangers, notamment anglais ou américains. Lorsqu'ils analysent des événements rares non directement échantillonnables, les panels montrent rapidement leurs limites et les chercheurs voudraient faire appel aux sources administratives.

L'usage général du NIR dans la sphère sociale ouvre en théorie une très large palette de données appariables. Mais l'exigence protectrice d'un décret en Conseil d'État a longtemps limité ces appariements à quelques opérations phares comme le vieux panel DAS, les échantillons permanents de retraités ou de cotisants ou le panel des bénéficiaires de minima sociaux. Les techniques anonymes d'appariements sécurisés (par cryptage irréversible, le hachage) évitent cette lourde démarche. Elles permettraient d'envisager le suivi anonyme de carrières individuelles en rapprochant les données des DADS et celles de l'Unedic ou de l'ANPE. Ainsi les économistes étofferaient leurs analyses de la dynamique de l'emploi et de la recherche d'emploi.

Les traitements précédents reposent strictement sur des données administratives anonymisées. Ce ne semble plus possible en France quand on élargit la recherche à la dynamique de l'emploi dans le couple, contrairement aux pays disposant de registres de population. Les fichiers fiscaux sont les seuls fichiers administratifs exhaustifs à identifier les couples, mais qui oserait y voir la racine de biographies professionnelles de couples ?

L'enquête Emploi en continu remplit répond à ce besoin, mais seulement dans une perspective de court terme ne couvrant même pas la période de dégressivité des allocations chômage. Il faudrait pouvoir la prolonger par des années de carrières observées dans les fichiers administratifs. Si l'échantillon démographique permanent relève la composition du ménage, il semble inadapté à s'enrichir de la biographie professionnelle du conjoint. La requête des économistes n'est pas simple à satisfaire.

Trop souvent les économistes ne trouvent pas leur compte auprès des enquêtes socio-démographiques. Ils en déplorent fréquemment l'absence des données économiques essentielles comme le revenu ou la présence de données économiques de faible qualité. Pourtant, des données administratives de qualité seraient mobilisables, notamment celles des DADS. Les données fiscales, d'abord mobilisées par le feu Cerc, viennent maintenant enrichir

¹ Université de Cergy

² Ined

les enquêtes de l'Insee. Leur usage pourrait être beaucoup plus large comme le mentionnait Jean-Michel Charpin, moyennant le recours aux plus strictes méthodes de la confidentialité.

Les moyens techniques permettent donc d'élargir ces appariements de données issues de sources ou d'institutions distinctes dans le respect de l'anonymat. C'est le débat que vise à susciter cette communication à l'issue de la présentation technique de Catherine Quantin sur les nouvelles méthodes d'appariements sécurisés proposées à la statistique publique.