

Biais de rotation de l'enquête Emploi en continu : l'impact des entrées-sorties d'échantillon

Pascale BREUIL¹

Comme beaucoup de panels tournants, l'enquête Emploi en continu est affectée d'un « biais de rotation » : alors que les six sous-échantillons enquêtés un trimestre donné sont tous supposés « représentatifs », ils peuvent conduire à des estimations systématiquement différentes d'une même grandeur. Ces biais sont particulièrement marqués pour le taux de chômage au sens du Bureau International du Travail (BIT). Ainsi, en moyenne annuelle pour 2006, le taux de chômage estimé à partir des sous-échantillons en première interrogation était de 10,5%, contre 9,3% à partir des échantillons en quatrième ou cinquième interrogation. Un écart de cette ampleur (1,2 point) constitue un élément de fragilité dans l'estimation du taux de chômage global, et il est donc important d'essayer d'en comprendre les mécanismes (pour ensuite tenter de les corriger).

On propose ici de décomposer le biais de rotation observé sur données brutes en deux effets : d'une part, les réponses recueillies pour une même personne peuvent varier de manière systématique selon le rang d'interrogation (et le mode de collecte associé : téléphone, visite, proxy...), d'autre part, l'échantillon de répondants peut lui-même se déformer.

Pour analyser ces effets, l'étude retient une approche longitudinale, consistant à étudier les transitions individuelles entre deux vagues successives. Les données utilisées sont celles des 6 dates d'observations des 12 sous-échantillons introduits entre le premier trimestre 2003 et le quatrième trimestre 2005. Parmi les 14 300 personnes d'un sous-échantillon qui répondent à un rang d'interrogation donné, environ 1000 sortent de l'échantillon à la vague suivante, et environ autant y entrent, tandis qu'elles ne sont qu'un peu plus de 600 à changer de statut d'activité au sens du BIT.

Ce premier résultat suggère que les entrées-sorties, particulièrement nombreuses, sont susceptibles de jouer un rôle important, et ce sont elles qui sont donc analysées ici. Chaque transition est étudiée de manière descriptive, puis décomposée en distinguant notamment probabilité d'être dans le champ et de répondre, et enfin estimée à partir d'une modélisation paramétrique simple.

L'étude confirme d'abord que les différences « volontaires » de protocole entre vagues (pas de repérage des nouveaux logements entre les vagues 1 et 6, imputations des réponses des logements d'inactifs en vagues intermédiaires) ont un impact négligeable sur le biais de rotation. Même après correction de ces effets « volontaires » de protocole, la dernière transition (débouchant sur une enquête en visite, plus longue) reste très atypique. De manière inattendue, les entrées et sorties du champ y sont environ deux fois plus nombreuses, ce qui suggère une légère sous-estimation des changements de situation vis-à-vis du champ lors des enquêtes téléphoniques. Mais surtout, l'attrition liée aux autres motifs de non-réponses (refus, impossible à joindre...) est supérieure de 2,1 points chez les chômeurs pour les vagues téléphoniques, ce qui

¹ Unité des Études démographiques et sociales, DSDS, Insee - pascale.breuil@insee.fr

crée un écart de taux de chômage entre les vagues extrêmes de l'ordre de 1 %. Cet écart est proche de celui observé, suggérant que cette attrition différentielle est un élément d'explication important du biais de rotation.