

PRÉSENTATION DU NUMÉRO

Le Courrier des statistiques, dans sa nouvelle formule, publie son numéro « N4 » : N comme nouveau, N comme numérique. L'ambition de 2018 est maintenue, avec pour objectif d'explorer les différents aspects des métiers et méthodes de la statistique, qu'elle soit publique ou non, et avec plusieurs volontés. Une exigence de pédagogie tout d'abord : les sujets sont vite complexes, techniques, or les articles visent non seulement les statisticiens publics, mais aussi, potentiellement, les étudiants, universitaires, et plus généralement le « citoyen éclairé », qui doit être à même de comprendre les principes des méthodes utilisées. Une indispensable ouverture ensuite, en invitant des instituts statistiques étrangers, des producteurs de données intéressantes pour un usage statistique, des universitaires concernés par les questions de traitement et de qualité de la donnée. Dans la même veine, nous veillons à ce que les articles sortent de leur zone de confort, par exemple en insistant sur la nécessité de références, notamment au niveau international.

Ce quatrième numéro du Courrier des statistiques s'ouvre sur un sujet qui n'a pas été abordé jusque-là par la revue, avec trois articles consacrés à la microsimulation. On peut se demander si les méthodes de la microsimulation relèvent bien du domaine de la statistique publique, et c'est justement une des questions que soulève le premier article. **Didier Blanchet** y définit les notions de microsimulation, statique ou dynamique, la façon de procéder en pratique, et les principales composantes d'un modèle de microsimulation. Il y explicite les raisons pour lesquelles en France, cette méthode s'est imposée naturellement dans la statistique publique. Il s'agit donc d'un article général qui, même s'il cite souvent le modèle Destinie, pose en quelque sorte un cadre, indispensable pour la suite. Les deux autres papiers portent en effet à chaque fois sur un modèle de microsimulation spécifique.

Ainsi, **Pierre Cheloudko et Henri Martin** racontent l'histoire du modèle TRAJECTOIRE, qui modélise le système de retraite et inclut ainsi inévitablement une forte dimension temporelle. Le modèle est dynamique, et avec lui le mot simulation recouvre ainsi à la fois la simulation (déterministe) d'une législation et la simulation (aléatoire) d'une carrière future. L'article montre les étapes ayant permis d'aboutir au dispositif actuel, en commençant par un modèle agrégé portant sur des cas-types, puis un deuxième bloc de modélisation des départs à la retraite, apportant la dimension dynamique, et enfin TRAJECTOIRE qui modélise l'intégralité des carrières, avec des données au niveau individuel.

Simon Fredon et Michaël Sicsic présentent Ines, modèle statique qui permet de simuler les politiques sociales et fiscales. Ils mettent en évidence les différents axes de travail lorsqu'on élabore et maintient un modèle de microsimulation : suivre de près et prendre en considération l'évolution de la législation, se constituer des bases de données solides à partir de plusieurs sources, caractériser les dispositifs à simuler. On constate également la puissance de tels modèles, dont les usages sont variés : effet d'une hausse ou baisse des transferts sociaux et fiscaux, effets des réformes sur les inégalités, etc.

Les autres articles de ce numéro ont un point commun : d'une façon ou d'une autre, ils portent sur des « références » de la statistique publique.

La nomenclature des professions et des catégories socio-professionnelles tout d'abord : cet instrument majeur pour les statistiques sociales vient de faire l'objet d'une vaste refonte, ce qui constitue un événement rare. **Thomas Amossé** en narre la genèse, après avoir rappelé l'historique des classifications socio-professionnelles et posé le concept de nomenclature. Il ne s'agit pas juste d'une nouvelle version, mais aussi d'une nouvelle logique, qui tient

compte du fait que les usages sont multiples : elle doit donc être suffisamment plastique pour s'adapter à ces différents usages.

Le répertoire Sirene est aussi un référentiel, un pivot de la production statistique d'entreprise, mais aussi de la simplification administrative, avec son identifiant unique qui structure tout l'écosystème administratif autour des entreprises. Contrairement à l'article précédent où un résultat abouti était présenté, l'article de **Christophe Alviset** se situe en amont, en présentant le programme Sirene 4. Le projet ne vise pas à revoir Sirene de fond en comble, mais plutôt à l'améliorer sur un certain nombre de sujets essentiels en vue d'une industrialisation des traitements : identification automatique d'établissement, codification automatique d'activité, normalisation des échanges de données, contrôles. L'enjeu sera aussi d'accompagner la révolution numérique de l'administration au service des entreprises.

La nécessité de références vaut non seulement pour les données, mais aussi pour les processus de production statistique. Depuis une dizaine d'années, le GSBPM s'est imposé comme une norme au niveau européen, désormais largement partagée. Dans ce numéro N4, le choix a été fait de donner la parole à un institut statistique étranger, en l'occurrence le *Statistiska centralbyrån* (SCB) en Suède, car il a appliqué de manière très vaste et systématique une démarche très voisine de celle portée par la communauté statistique internationale avec le GSBPM. **Johan Erikson** y explique les riches enseignements tirés de cette pratique : les facteurs de réussite, les freins et les évolutions envisagées.

Les enquêtes auprès des ménages, lorsqu'elles sont effectuées en face-à-face, requièrent des techniques de tirages d'échantillon très sophistiquées liées à la nécessité de faciliter au plan pratique le déplacement des enquêteurs, tout en respectant les contraintes de maîtrise de probabilités d'inclusion et d'optimisation de la précision. Pour éviter de se poser de complexes questions méthodologiques lors de chaque tirage, on se constitue une référence, que l'on conserve dix ans, et au sein duquel on tire les échantillons pour toute nouvelle enquête en face-à-face : l'échantillon-maître. Comme l'expliquent **Patrick Sillard et alii**, ce dernier est désormais opérationnel depuis 2020. La nouveauté réside dans l'utilisation des sources fiscales et la coordination avec l'échantillon de l'enquête Emploi.

Le dernier article, contrairement aux précédents qui portent sur des référentiels transversaux, s'intéresse à un domaine spécifique : le logement. Et là aussi il s'agit de s'interroger sur la référence : quelles sont les données et statistiques de référence ? **Jérôme Harnois et Pierre Lamarche** abordent la question sous toutes ses facettes : la construction, l'équipement en logement, le prix, le coût d'occupation, l'état des logements et le mal-logement, etc. L'article ne se limite donc pas à pointer telle ou telle enquête ou source administrative : il adopte un point de vue original en visant une démarche systémique, structurée par le cycle de vie du logement, mais également contrainte par les limites des sources de données disponibles.

Odile Rascol
Rédactrice en chef, Insee
