

Multiplication des données disponibles et évaluation des politiques publiques

par François Ecalé, économiste, Université Paris 1

Les données disponibles pour estimer l'impact des mesures de politique économique sont déjà très nombreuses, mais les résultats de ces estimations sont souvent très imprécis constate François Ecalé. L'évaluation des politiques publiques pose de difficiles problèmes méthodologiques que le développement des grandes boîtes noires du Big Data et de l'Open Data ne résoudra pas.

Le manque de données disponibles a longtemps été invoqué par de nombreux économistes pour délaissé les travaux empiriques, notamment l'évaluation des politiques publiques, et faire porter leurs efforts sur la recherche théorique. Le Big Data et l'Open Data permettront-ils de mieux connaître l'impact des politiques publiques et d'en améliorer l'efficacité ? Je n'en suis pas convaincu.

Les économistes disposent déjà depuis longtemps de l'accès à de nombreuses bases de données, rendues anonymes lorsqu'elles sont individuelles, qui résultent d'enquêtes des services statistiques ou de la gestion des administrations. Ils peuvent désormais assez facilement les exploiter à distance et les apparier. Certes, quelques administrations résistent au développement de l'open data, comme les services fiscaux, mais elles devraient rapidement se laisser aussi entraîner dans ce mouvement.

Les évaluations économiques des politiques publiques sont ainsi beaucoup plus fréquentes dans de nombreux domaines, comme celui de l'emploi. L'analyse des séries macroéconomiques a été complétée, voire remplacée, par le traitement d'informations individuelles sur les ménages et entreprises. La multiplication des données a permis de leur appliquer des méthodes d'analyse statistiques beaucoup plus sophistiquées.

L'envers de la médaille est une diminution de la lisibilité de ces études pour ceux qui ne sont pas initiés aux dernières techniques économétriques. A la lecture de certains articles, il apparaît que l'auteur préfère la beauté mathématique de la méthode mise en œuvre à la compréhension des résultats par ceux qui contribuent à définir les politiques économiques. Il n'est donc pas étonnant que ces derniers utilisent très peu les résultats de ces études et ni le Big Data ni l'Open Data n'y changeront quoi que ce soit.

En outre, la multiplication des données disponibles et des méthodes d'analyse ne conduit pas nécessairement à des résultats plus précis, comme le montrent les évaluations empiriques des allègements de charges sociales sur les bas salaires. Il s'agit certainement d'une des mesures de politique économique ayant donné lieu au plus grand nombre d'études, mettant en œuvre les méthodes d'analyse les plus diverses, ce qui est justifié au regard de son coût (20 Md€). Il résulte de ces travaux que le nombre d'emplois créés, ou sauvés, se situe dans une fourchette très large : de 200 000 à 1 000 000. Au moins ces études concluent-elles toutes à un impact significativement positif de ce dispositif sur l'emploi alors que, pour la plupart des mesures de politique économique, le signe même de cet impact est indéterminé ou ne fait pas consensus parmi les économistes. Cette indétermination de l'impact des politiques économiques a des causes profondes qui subsisteront malgré le développement du Big Data et de l'Open data.

L'impact d'une politique publique au regard d'un objectif tel que l'emploi doit en effet être mesuré en faisant la différence entre le nombre d'emplois constaté en mettant en œuvre cette politique et le nombre d'emplois qui aurait été constaté si elle n'avait pas été mise en œuvre. Or, par définition, ce « contrefactuel », comme disent les économistes, est difficilement observable. Il est certes possible de comparer les effectifs de deux échantillons d'entreprises tirées au hasard, les unes soumises à cette politique et les autres non soumises, mais de telles expérimentations seront toujours difficiles à mettre en œuvre en pratique. A défaut, les méthodes statistiques permettent de comparer des entreprises de mêmes caractéristiques, soumises ou non à cette politique, mais il n'est jamais sûr que les entreprises comparées sont vraiment comparables (par définition, les entreprises concernées par une mesure de politique économiques diffèrent de celles qui ne sont pas concernées) et les résultats sont

nécessairement imprécis.

De plus, les effets des politiques économiques sont rarement immédiats, car il faut que les ménages ou les entreprises en prennent connaissance puis ajustent leurs comportements. Or, ces politiques sont souvent modifiées avant de produire tous leurs résultats et de pouvoir être évaluées. Il arrive aussi que plusieurs mesures soient prises à peu près au même moment et qu'il soit difficile de distinguer leurs effets. L'imprécision des analyses de l'impact des allègements de charges sociales intervenus dans les années 1998 à 2005 tient ainsi pour beaucoup à la difficulté de séparer leurs effets de ceux de la réduction de la durée du travail.

Enfin, à supposer qu'il soit possible d'estimer précisément l'impact d'une mesure de politique de l'emploi en montrant que les effectifs ont augmenté plus fortement dans les entreprises concernées, il faudrait encore tenir compte de ce que les économistes appellent le « bouclage macroéconomique » : ces nouveaux emplois génèrent des revenus supplémentaires qui entraînent une augmentation de la consommation mais aussi des importations... ; la baisse du chômage tend à accroître les pressions à la hausse des salaires, ce qui réduit la compétitivité des entreprises et les exportations... Pour tenir compte de ce bouclage macroéconomique, il faut un modèle de l'économie française. Or la fiabilité de ces modèles a été fortement remise en cause par la crise de 2008-2009 et les difficultés de sortie de cette crise.

La multiplication des données disponibles et la croissance des capacités de traitement informatique pourraient laisser croire que de nouvelles générations de modèles, beaucoup plus fiables, pourront être développées dans les prochaines années. Je crains que ce soit une illusion, semblable à celle qu'entretenaient les macro-économistes des années soixante et soixante-dix pour les mêmes raisons (développement de l'appareil statistique et des capacités de traitement des ordinateurs). Ils ont ainsi construit de très grands modèles avec des centaines d'équations qui ont été abandonnés pour des modèles beaucoup plus simples, mais inévitablement réducteurs de la réalité, parce que personne, à part leurs concepteurs, ne pouvait comprendre leur fonctionnement et en interpréter correctement les résultats.

Je pressens que le Big Data va conduire à construire de nouvelles grandes boîtes noires d'où sortiront des résultats très difficiles à interpréter. Extrapolant les comportements passés, ces outils pourront avoir de bonnes capacités prédictives à court terme mais leurs performances ne se maintiendront pas dans la durée.

Au total, la multiplication des données mises à la disposition des chercheurs ne résoudra pas les difficultés de l'évaluation des politiques publiques résultant de l'absence de contrefactuel, de l'instabilité des politiques et des limites des outils statistiques et économiques, ceux-ci devant à la fois être simples pour être compréhensibles et prendre en compte toute la complexité du réel.

 [Télécharger le PDF de l'article](#)

<< [Retour au sommaire](#)

PRES@ JE.COM

Une publication de l'Institut PRESAJE

(Prospective, Recherche et Etudes Sociétales Appliquées à la Justice et à l'Economie)

Siège social : 2 avenue Hoche 75008 Paris - Courrier : 30 rue Claude Lorrain 75016 Paris

Tél. 01 46 51 12 21 - E-mail : contact@presaje.com - www.presaje.com

Directeur de la publication : Michel Rouger

Pour ne plus recevoir d'e-mails de la part de Presaje, [cliquez ici](#)

[>> CONSULTER LES PRECEDENTS NUMEROS](#)