

L'apport des Big Data pour les prévisions macroéconomiques à court terme et « en temps réel » : une revue critique*

Pete Richardson

Question clé

Avec la croissance exponentielle des capacités de stockage et de traitement de données au cours des dernières années, l'utilisation des ensembles dits de « Big Data » est devenue de plus en plus réalisable pour les économistes et d'autres analystes. Ces ensembles de Big Data englobent notamment les données issues de la recherche sur Internet, des médias sociaux et des transactions financières. Un certain nombre d'études empiriques récentes, principalement postérieures à la crise, ont donc exploré l'utilisation potentielle de ces données, généralement plus volumineuses et plus rapides que celles fournies traditionnellement par les autorités statistiques nationales, en tant qu'outils de prévision macroéconomique. Néanmoins, à ce jour, il semble y avoir relativement peu d'examen systématiques des travaux empiriques associés à cette problématique.

La revue proposée dans l'article vise à rétablir l'équilibre avec une analyse de la pertinence du Big Data pour les prévisions économiques et un examen critique de plusieurs études empiriques publiées à ce jour, en s'appuyant sur différentes sources, dont les recherches sur Internet (en particulier sur Google), les médias sociaux (Twitter) et les statistiques relatives aux transactions financières. L'article est principalement mené du point de vue de la prévision économique pratique et examine des études portant sur divers sujets, notamment les marchés du travail, les dépenses de consommation, le commerce, le tourisme, l'inflation, les marchés financiers et l'activité macroéconomique dans son ensemble.

Principaux résultats et message

L'éventail des études empiriques examinées fournit des informations intéressantes et des preuves de corrélations significatives et de performances prédictives dans divers domaines. De manière générale, les résultats sont toutefois assez mitigés, reflétant à la fois la simplicité relative des modèles utilisés et les limitations importantes en termes de qualité, de forme, de taille d'échantillon et de leur nature qualitative. Les applications les plus réussies semblent être celles qui cherchent à intégrer cette classe d'informations dans un cadre économique cohérent, par opposition à une approche statistique simpliste, de type boîte noire. À cet égard, il reste encore beaucoup à faire :

- pour affiner et améliorer les normes de qualité des ensembles de Big Data disponibles et leur accessibilité ;
- pour proposer de meilleures méthodes d'extraction des informations pertinentes pour des domaines spécifiques de la recherche économique ;
- pour améliorer les moyens de comparaison et de test entre les différentes méthodes de mesures ;
- pour poursuivre l'adaptation et l'amélioration des cadres pertinents de tests et de modélisation permettant d'intégrer des informations relatives à un futur assez proche dans les prévisions macroéconomiques à court terme.

Néanmoins, il existe quelques exemples évidents pour lesquels de tels indicateurs pourraient utilement compléter les approches existantes fondées sur la prévision immédiate et d'autres approches basées sur des indicateurs, dans le cadre de la sélection générale des variables à analyser et, en tant que tels, ces exemples constituent un ajout utile aux outils dont disposent l'économiste et le statisticien pour l'analyse économique à court terme.