

Estimation des revenus des ménages à partir de régressions quantiles

Auteur

Eric Durieux - *Psar Analyses territoriales*

Résumé

La mesure et la qualification de la précarité énergétique deviennent un enjeu majeur des politiques publiques depuis l'inscription de la lutte contre cette forme de précarité dans la loi dite "Grenelle-2" et le récent débat sur la transition énergétique. La difficulté est de pouvoir réaliser de telles estimations aux niveaux géographiques les plus fins possibles, notamment au niveau communal. La mesure de cette précarité énergétique passant par la mesure d'un taux d'effort énergétique, défini comme la part des revenus disponibles consacrés au chauffage pour ce qui est du logement, il nous est ainsi nécessaire de connaître à l'échelle des ménages les revenus disponibles d'une part et les dépenses énergétiques d'autre part.

Il n'existe pas de sources statistiques permettant de réunir ces différentes composantes, c'est pourquoi nous avons décidé d'utiliser le recensement de la population, qui nous fournit des caractéristiques sur la population et les logements, puis d'estimer le revenu de chaque ménage.

La très grande majorité des études économiques empiriques portent sur les effets moyens des variables auxquelles on s'intéresse. Cet état de fait s'explique par la diffusion d'un outil méthodologique, la régression linéaire, simple à utiliser et bien maîtrisé. La moyenne apporte une information essentielle, mais cependant limitée : le revenu moyen n'informe pas, par exemple, sur leur répartition plus ou moins inégale. Or dans le cadre de notre étude il est important de bien estimer les revenus en queue de distribution.

Les régressions de quantile sont une réponse à ce type de préoccupation. Elles permettent notamment d'étudier l'impact de différents facteurs sur l'ensemble de la distribution d'une variable d'intérêt. Si leur principe est ancien, elles se sont récemment développées, par exemple dans le cadre d'évaluation de politiques publiques (voir La régression de quantiles en pratique, Pauline GIVORD et Xavier D'HAULTFŒUILLE 2001). Un ensemble de packages préprogrammés, notamment sous R, en font aujourd'hui un outil simple d'utilisation.

Cet article propose une méthode d'estimation du revenu des ménages à partir de régressions quantiles. On estime des régressions quantiles sur différentes strates à partir de l'enquête sur les revenus fiscaux et sociaux (ERFS). Ces régressions vont nous permettre ensuite d'estimer pour chaque individus du recensement de la population les quantiles issus des différents modèles. L'appartenance à un intervalle inter quantile est ensuite estimé à partir d'un modèle de régression logistique multinomial du type réseau de neurones.

Cette procédure empirique se rapproche par certains aspects de la méthode M-Quantile décrite dans l'article de Chambers et Tzavidis (M-Quantile Models for Small Area, 2006). Néanmoins, elle s'en différencie par l'approche empirique pour atteindre l'objectif d'estimation de données individuelles. Les différents tests réalisés à partir d'échantillons de contrôles montrent une stabilité de la méthode et une très bonne restitution de la distribution des revenus, ce que permet moins bien de faire un modèle linéaire.