

Pour qui sont les bons salaires ? Une estimation débiaisée du rôle des effets d'appariements dans les inégalités de salaire

Damien BABET (*), Olivier GODECHOT (**), Marco Guido PALLADINO (***)

(*) Insee, Direction des études et synthèses économiques

(**) CNRS, Sciences-Po

(***) Sciences-Po

damien.babet@insee.fr

Mots-clés. : Données de panel, effets fixes, split sampling, surapprentissage, formes quadratiques.

Domaines. Économétrie.

Résumé

En Allemagne et aux États-Unis notamment, la croissance des inégalités de salaires se manifeste surtout entre entreprises plutôt qu'au sein des entreprises. Cela s'explique par les différences entre entreprises, mais aussi par l'effet de *sorting* ou d'appariement : certaines entreprises payent mieux, et ces employeurs ont de plus en plus tendance à embaucher des salariés qui gagnent plus (Card, Heining, Kline 2013, Song et al. 2019). Cette augmentation de l'effet d'appariement se manifeste-t-elle également en France ?

La mesure du *sorting* se fonde sur un modèle de panel à doubles effets fixes de type « AKM » (Abowd, Kramarz et Margolis, 1999) où le logarithme du salaire d'un salarié i est la somme de l'effet entreprise $\psi_{j(i,t)}$ de l'employeur j de i et d'un effet individuel θ_i , capturant l'hétérogénéité inobservée, ainsi que de l'effet de covariables x qui ne sont pas constantes dans le temps.

$$y_{it} = \beta x_{it} + \theta_i + \psi_{j(i,t)} + u_{it}$$

Le *sorting* est alors la covariance entre les effets entreprises et les effets individus $cov(\theta_i, \psi_j)$. Or l'identification de ces effets fixes s'appuie sur les salariés mobiles et la variation de leur salaire lorsqu'ils passent d'un employeur à l'autre. Il existe alors un biais « de mobilité limitée », un phénomène de surapprentissage déjà connu, du fait que la connectivité du réseau des entreprises, et l'estimation des effets fixes relatifs, peut reposer sur un faible nombre de salariés mobiles : la variance des effets fixes est biaisée à la hausse et leur corrélation à la baisse. De manière générale, ce sont toutes les formes quadratiques des estimateurs qui sont biaisées.

Deux principales solutions sont aujourd’hui proposées dans la littérature : le *clustering* des entreprises ou des individus pour réduire la dimension de l’hétérogénéité des effets fixes (Bonhomme, Lamadon, Manresa 2017), qui demande des hypothèses additionnelles, et l’estimation de l’erreur d’estimation de chaque effet fixe par *leave-out* (Kline, Saggio Sølrvsten, 2020), qui reste complexe et coûteuse en temps de calcul.

Nous introduisons deux innovations. D’une part, nous estimons pour la première fois l’évolution du *sorting* sur des données françaises, grâce à la constitution d’un pseudo-panel en chaînant les DADS exhaustives (Godechot et al. 2020). D’autre part, nous utilisons une méthode de *split-sampling* dont nous montrons qu’elle est valide sous les mêmes hypothèses que la méthode de *leave-out*, mais plus simple et rapide. On peut en effet écrire l’estimateur *split-sampling* d’une forme quadratique des paramètres comme $\hat{\omega}^{SP} = \hat{\alpha}'_0 A \hat{\alpha}_1$ avec $\hat{\alpha}_s = (\hat{\beta}, \hat{\theta}, \hat{\psi})_s$, $s = 1, 0$ indiquant la moitié d’échantillon utilisée et A étant une matrice symétrique appropriée. Avec ϵ l’erreur d’estimation des paramètres, $\hat{\alpha}_s = \alpha + \epsilon_s$ on a alors en espérance un biais égal à la forme quadratique des erreurs entre chaque moitié de l’échantillon, qui est nulle sous des conditions raisonnables :

$$\begin{aligned} \mathbf{E}[\hat{\omega}^{SP}] - \omega &= \text{trace}(A\mathbf{E}[\epsilon_0\epsilon_1']) \\ &= \mathbf{E}[\epsilon_0'A\epsilon_1] \end{aligned}$$

Cette correction augmente très sensiblement la mesure de l’effet de *sorting*. Estimé ainsi, l’appariement explique 11,6% de la variance des log-salaires en France sur la période 2002-2006, et 13,5% sur la période 2012-2016. Toutes les estimations, corrigées ou non, s’accordent pour décrire une augmentation du poids du *sorting*.

Bibliographie

- [1] Abowd J. M., Kramarz F., Margolis D. N., « High wage workers and high wage firms. » *Econometrica* 67.2, 1999.
- [2] Bonhomme S., Lamadon T., Manresa E., « Discretizing unobserved heterogeneity » Working Paper, IFAU, 2017.
- [3] Card D., Heining J., Kline P., “Workplace heterogeneity and the rise of West German wage inequality,” *The Quarterly journal of economics*, 128, 2013.
- [4] Godechot O., et al., « The great separation : Top earner segregation at work in high-income countries », MaxPo Discussion Paper, 2020.
- [5] Kline P., Saggio R., Sølrvsten M., « Leave-out estimation of variance components », *Econometrica*, 88, 2020.
- [6] Song J., Price D. J., Guvenen F., Bloom N., VonWachter T., “Firming up inequality,” *The Quarterly journal of economics*, 134, 2019.