

# Salaires et salariés en Île-de-France

---

Pierre-Philippe Combes\*  
Gilles Duranton\*\*  
Laurent Gobillon\*\*\*

*Ce travail compare les distributions de salaires et de qualifications de l'Île-de-France et du reste du pays. Les salaires sont plus élevés en Île-de-France de 24 %. Ces différences relèvent pour le quart de qualifications plus élevées et pour le reste d'une plus grande productivité francilienne. Nous montrons aussi que les disparités de salaires sont beaucoup plus fortes à l'intérieur de l'Île-de-France et reflètent alors principalement de plus grandes disparités de qualifications entre salariés franciliens que dans le reste de la France. Les choix migratoires faits par les salariés n'expliquent qu'une faible part de ces différences. Les différences entre l'Île-de-France et le reste du pays se retrouvent à un degré moindre quand nous comparons Paris et sa petite couronne à la grande couronne parisienne.*

## WORKERS AND WAGES IN THE GREATER PARIS REGION

*We compare the distributions of wages and worker skills for the Greater Paris region to the rest of France. Wages are 24% higher in Greater Paris relative to the rest of the country. Greater skills of Parisian workers account for about a quarter of this difference while the greater productivity of the capital region accounts for the rest. We also show that wage disparities are higher within the Greater Paris region. This reflects mainly a greater dispersion of skills among Parisian workers than in the rest of the country. In turn, little of this greater dispersion of skills appears to be caused by worker mobility. The differences between the Greater Paris region and the rest of the country are also found, albeit to a lesser extent, when comparing Central Paris and its first ring of suburbs to the rest of the Greater Paris region.*

Classification JEL : J31, J61, R12, R23

---

\* Aix-Marseille Université (Aix-Marseille School of Economics), CNRS & EHESS et Sciences Po Paris. Aussi affilié au Center for Economic Policy Research. *Correspondance* : GREQAM, 2 rue de la Charité, 13236 Marseille cedex 02, France. *Courriel* : ppcombes@gmail.com. *Site internet* : <http://www.vcharite.univ-mrs.fr/pp/combes/>.

\*\* Wharton School, University of Pennsylvania. Aussi affilié au Center for Economic Policy Research. *Correspondance* : Wharton School, University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104, USA. *Courriel* : duranton@wharton.upenn.edu. *Site internet* : <https://real-estate.wharton.upenn.edu/profile/21470/>.

\*\*\* Institut national d'études démographiques – Paris School of Economics. Aussi affilié au Center for Economic Policy Research et à l'Institute for the Study of Labor. *Correspondance* : INED, 133 boulevard Davout, 75980 Paris cedex 20, France. *Courriel* : laurent.gobillon@ined.fr. *Site internet* : <http://laurent.gobillon.free.fr/>.

Nous remercions Thierry Kamionka et trois rapporteurs anonymes pour leurs commentaires. Ce travail a bénéficié d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du programme Investissements d'avenir portant la référence ANR-10-EQPX-17 (Centre d'accès sécurisé aux données – CASD).

## INTRODUCTION

Depuis la fin de la Seconde Guerre mondiale jusqu'au milieu des années 2000, une constante de la politique territoriale en France a été de chercher à réduire l'influence du « géant » francilien pour développer le « désert français » (Gravier et Dautry [1947], Gravier [1958]). L'aménagement du territoire et la décentralisation ont été, durant ces années, les deux instruments de choix des gouvernements français successifs. Toutefois, au milieu des années 2000, un tournant important a été pris avec la politique dite des pôles de compétitivité (Duranton, Martin, Mayer et Mayneris [2008]) qui a annoncé le retour d'un objectif d'efficacité et non plus seulement d'équité pour la politique territoriale en France. Ce retour a été affirmé avec le projet du Grand Paris qui prévoit des investissements très importants dans les transports publics franciliens. Ce regain d'intérêt des politiques publiques en France pour la région parisienne s'est aussi accompagné de tentatives de réorganisations administrative et institutionnelle afin de mieux faire fonctionner l'ensemble francilien. Dans le même temps, depuis plus de trente ans, le problème des « banlieues » se pose. Il constitue la manifestation la plus visible d'une grande hétérogénéité sociale au sein des régions métropolitaines françaises. Cette question se pose avec une acuité particulière pour Paris et sa banlieue.

Afin d'éclairer ces débats, nous décrivons dans ce travail les disparités salariales entre la région francilienne et le reste du pays ainsi qu'au sein de la région francilienne. Plus précisément, nous comparons les distributions de salaires et de qualifications au lieu de travail entre l'Île-de-France et le reste du pays. Les salaires sont plus élevés en Île-de-France d'environ 24 %. Ces différences reflètent pour le quart environ de meilleures qualifications (que nous mesurons comme un effet individuel fixe au court du temps dans une équation de salaire) des salariés franciliens. Pour le reste, ces différences sont le produit d'une plus grande productivité de la région Île-de-France. Les disparités de salaires sont aussi beaucoup plus fortes au sein de l'Île-de-France et reflètent principalement les plus grandes disparités de qualifications entre salariés franciliens qu'entre ceux du reste de la France. Les salariés les plus qualifiés sont fortement surreprésentés en Île-de-France de même que les moins qualifiés.

Ces différences se retrouvent dans les processus migratoires. Les migrants du reste de la France vers l'Île-de-France sont en moyenne plus qualifiés que ceux qui n'y migrent pas, mais on observe encore parmi ces migrants une surreprésentation des moins qualifiés comme des plus qualifiés. Ce processus est renforcé par le processus en direction opposée, puisque les migrants qui quittent l'Île-de-France pour le reste de la France sont en moyenne moins qualifiés que ceux qui y restent même si les salariés très peu qualifiés sont encore sous-représentés parmi ceux qui quittent l'Île-de-France. Nous montrons aussi que les natifs d'Île-de-France qui travaillent en Île-de-France ont des qualifications plus élevées que les natifs du reste de la France qui travaillent en Île-de-France. Les natifs d'Île-de-France qui travaillent dans le reste de la France ont aussi des qualifications plus élevées que les natifs du reste de la France qui y travaillent. De plus, les faibles écarts de distribution entre les migrants qui quittent l'Île-de-France et ceux qui y arrivent suggèrent que le tri spatial observé entre l'Île-de-France et le reste du pays ne peut être expliqué qu'en faible partie par les migrations.

Nos résultats montrent aussi que les écarts de salaires et de qualifications sont plus importants pour les salariés plus âgés que pour les salariés plus jeunes. Autrement dit, être né (et sans doute avoir grandi) en Île-de-France n'offre qu'un avantage faible. Avoir travaillé en Île-de-France est, en revanche, associé avec des effets beaucoup plus forts. Cela suggère que le tri spatial observé en France a une composante majeure – avoir travaillé en Île-de-France – et deux composantes mineures – être né en Île-de-France et avoir migré vers et depuis l'Île-de-France.

Nous effectuons aussi un certain nombre de comparaisons à l'intérieur de la région Île-de-France. Lorsque Paris et sa petite couronne sont comparés avec la grande couronne, des différences similaires à celles entre l'Île-de-France et le reste de la France sont observées. Ces différences sont toutefois plus faibles. Les salaires sont d'environ 9 % plus élevés dans le centre de l'Île-de-France (Paris et petite couronne) que dans sa périphérie. La dispersion des salaires est aussi plus importante à Paris et dans sa petite couronne que dans la grande couronne. Pour les qualifications, en revanche, l'écart moyen entre Paris et petite couronne et la grande couronne est quasiment le même qu'entre l'Île-de-France et le reste de la France. Ainsi, le tri spatial des qualifications semble concerner principalement la partie centrale de l'Île-de-France. Toutefois, le même phénomène ne s'observe pas quand on compare Paris à sa petite couronne. Au contraire, les salaires et les qualifications sont plus bas à Paris. La petite couronne est donc l'épicentre du tri spatial en France.

Notre étude confirme donc une spécificité de la région Île-de-France et plus particulièrement de sa partie centrale en termes de salaires et de qualifications. Cette spécificité concerne à la fois des niveaux de salaires et de qualifications plus élevés et leur plus grande dispersion. Bien que l'Île-de-France offre en moyenne de meilleurs salaires et recèle des salariés plus qualifiés, elle accueille aussi les salariés les moins qualifiés de façon disproportionnée.

Notre travail, qui est centré sur la comparaison entre l'Île-de-France et le reste de la France et sur des comparaisons au sein de l'Île-de-France, suit celui de Combes, Duranton, Gobillon et Roux [2012b] qui comparent plus globalement les zones d'emploi plus denses aux moins denses en France métropolitaine. Cette approche s'inscrit plus généralement dans la vaste littérature sur les disparités spatiales au niveau infranational et celle sur les inégalités interindividuelles à un niveau infranational. Cette dernière littérature connaît actuellement un renouveau (Glaeser, Resseger et Tobio [2009], Baum-Snow et Pavan [2014], Behrens et Robert-Nicoud [2014]). Nous confirmons, à partir du cas particulier de Paris et de sa région, certaines conclusions obtenues dans des travaux antérieurs comme, par exemple, l'existence de plus grandes inégalités interindividuelles dans les grandes villes. Toutefois, la plus grande disparité de l'Île-de-France en termes de qualifications individuelles semble expliquer pour l'essentiel sa plus grande disparité de salaires. Cela est en contraste avec les résultats de Combes *et al.* [2012b] sur la France entière et ceux de Glaeser *et al.* [2009] sur les villes américaines qui n'attribuent qu'une partie des plus grandes disparités de salaires aux écarts de qualification.

Comme Combes *et al.* [2012b], nous faisons le lien entre les disparités de qualifications et la migration des salariés ainsi que leur origine. En cela, notre étude contribue à la littérature sur le tri spatial entre villes et régions. Alors que ce phénomène semble prendre de l'ampleur (au moins aux États-Unis comme le montrent Berry et Glaeser [2005], Baum-Snow et Pavan [2014], ou encore

Diamond [2013]), il reste mal documenté et mal compris. Comme d'autres contributions antérieures, nous montrons dans le cas francilien que la prime de salaire pour les villes les plus grandes provient en partie de qualifications plus élevées (Combes, Duranton et Gobillon [2008], Mion et Naticchioni [2009], Matano et Naticchioni [2012], Baum-Snow et Pavan [2012]). Toutefois, la part expliquée des différences de salaires par les différences de qualifications est plus faible quand on compare l'Île-de-France au reste de la France (environ le quart) que quand on compare les zones d'emploi plus denses et les zones d'emploi moins denses en France (où cette part est d'environ 40 % selon Combes *et al.* [2012b]).

La complémentarité entre villes et qualifications (Glaeser et Resseger [2010]) implique que les plus qualifiés devraient migrer vers les villes plus grandes (Behrens, Duranton et Robert-Nicoud [2014]). Il a aussi été suggéré que les villes les plus grandes abritent une plus grande proportion d'individus parmi les moins qualifiés (Glaeser, Kahn et Rappaport [2008], Eeckhout, Pinheiro et Schmidheiny [2014]). Nous confirmons ces conjectures dans le cas francilien, ce qui semble suivre une même logique observée plus généralement en France entre zones d'emploi denses et moins denses (Combes *et al.* [2012b]).

Enfin, notre travail contribue à la littérature sur les migrations à l'échelle infranationale (voir les revues de Greenwood [1997], Etzo [2008]). La possibilité de phénomènes de sélection dans le choix de migrer ou non est une difficulté fondamentale rencontrée par cette littérature. Les individus décidant de migrer ne sont pas identiques à ceux qui décident de ne pas le faire et présentent des caractéristiques qui ont une influence sur les phénomènes étudiés, et notamment les salaires et les qualifications. Cette sélection est une contrepartie naturelle du tri spatial<sup>1</sup>. Nous documentons plusieurs faits intéressants sur ce sujet, telles la sélection principalement positive des migrants vers l'Île-de-France et la sélection principalement négative des migrants depuis l'Île-de-France vers le reste de la France. Toutefois, les choix migratoires n'expliquent qu'une très faible part du tri spatial entre l'Île-de-France et le reste du pays.

Le reste de cet article est organisé de la manière suivante. La section 2 présente les données utilisées et la section 3 introduit notre méthodologie. Nos principales comparaisons entre l'Île-de-France et le reste de la France apparaissent dans la section 4. La question du tri spatial entre Franciliens et non-Franciliens est explorée en détails dans la section 5. La section 6 examine un certain nombre de comparaisons au sein même de l'Île-de-France. Enfin, la section 7 conclut.

## DONNÉES

Nous utilisons le département comme unité spatiale dans notre étude. Tout d'abord, ceux-ci nous permettent de reprendre la décomposition usuelle de

---

1. Néanmoins, il ne s'agit pas d'une condition nécessaire. Nous définissons le tri spatial comme une différence de qualification des salariés entre zones. La sélection des migrants concerne le choix de localisation, qui peut dépendre de leurs qualifications. Ces deux aspects sont bien évidemment liés, mais la sélection n'est qu'une condition suffisante pour le tri spatial. Ce dernier peut avoir d'autres sources, telles la transmission familiale des qualifications, les différences de performance des systèmes éducatifs locaux ou des effets d'apprentissage localisés.

l'Île-de-France en Paris, petite couronne (qui comprend les Hauts-de-Seine, le Val-de-Marne et la Seine Saint-Denis) et grande couronne (qui comprend le Val-d'Oise, les Yvelines, l'Essonne et la Seine-et-Marne).

Pour les salaires, nous utilisons le panel des Déclarations annuelles des données sociales (DADS) de l'Insee comme dans Combes *et al.* [2008], Combes, Duranton, Gobillon et Roux [2010] et Combes *et al.* [2012b]<sup>2</sup>. Les DADS sont collectées pour le calcul des retraites, des charges sociales et des taxes payées par les entreprises. Les établissements doivent remplir un formulaire pour chacun de leur salarié chaque année. Une observation correspond ainsi à une combinaison salarié-établissement-année. Les données sont restreintes aux salariés des secteurs de l'industrie et des services qui travaillent en France, hors secteur public. Afin d'obtenir un panel cohérent au cours du temps, seuls les individus nés en octobre d'une année paire doivent être conservés. De plus, pour des raisons de faisabilité informatique, nous ne conservons que cinq points dans le temps séparés par des intervalles de trois ans : 1995, 1998, 2001, 2004 et 2007. Nous restreignons notre attention aux années postérieures à 1993 afin de disposer du nombre d'heures travaillées, ce qui nous permet de construire un salaire horaire.

Pour chaque observation, nous disposons de l'âge, du sexe, du lieu de naissance (au niveau du département), de la catégorie socioprofessionnelle (deux chiffres), et du salaire. Les données ne contiennent pas d'information sur les études. Afin d'obtenir un échantillon homogène, nous ne gardons que les salariés à temps plein. Les données contiennent aussi des informations au niveau de l'établissement comme le secteur (NAF rev.1) et le département. À ce propos, on peut souligner que notre étude se focalise entièrement sur le lieu de travail. La distinction entre lieu de travail et lieu de résidence n'est que de peu d'importance dans la comparaison entre l'Île-de-France et le reste du pays. Il est, en revanche, important de la garder à l'esprit lorsqu'on discute du tri spatial à l'intérieur de la région francilienne. Il s'agit du tri spatial sur le marché du travail et non d'un tri résidentiel.

Les données brutes contiennent 3 529 297 observations pour les cinq points dans le temps que nous conservons. Lorsque seuls les emplois occupés par des salariés âgés de 20 à 59 ans sont considérés, il ne reste plus que 3 308 205 observations. L'exclusion supplémentaire des emplois pour lesquels le salaire ou le nombre d'heures travaillées est manquant, ou pour lesquels il est mentionné que l'identifiant du salarié est mal renseigné, fait diminuer la taille de l'échantillon à 3 159 867 observations. Certains salariés occupent plusieurs emplois durant une même année. Nous ne conservons que leur emploi principal, défini comme l'emploi qui leur assure les revenus totaux les plus élevés cette année-là. La taille de notre échantillon atteint alors 2 773 484 observations. Après la suppression des emplois hors France métropolitaine, ceux ayant un identifiant du secteur ou de la zone d'emploi manquant ou mal renseigné, il reste 2 733 655 observations. Nous supprimons aussi les emplois dans les petits secteurs (comme le transport spatial), dans les secteurs à but non lucratif, et dans les secteurs dans lesquels les entreprises renseignent leur emploi uniquement au niveau régional (comme

---

2. La préparation des données suit celle de Combes *et al.* [2012b] auquel le lecteur peut se référer pour plus de détails.

la finance)<sup>3</sup>. Cette sélection conduit à un échantillon de 2 525 131 observations. Enfin, nous supprimons 0,1 % des observations à chaque extrémité de la distribution des salaires chaque année afin d'éviter les valeurs extrêmes générées par des erreurs de codage<sup>4</sup>. Notre échantillon final sur lequel est basée l'estimation des effets individuels contient 2 520 069 observations. Les distributions de salaires en 2007 sont fondées sur 566 104 observations pour la France entière. Les distributions de salaires relatives aux migrants ne conservent que les individus présents à la fois en 2004 et en 2007, soit 378 645 observations pour la France entière.

Tableau 1. *Statistiques descriptives sur les disparités de salaire au sein de différentes zones*

	Effectif	Moyenne	Écart type	1 <sup>er</sup> décile	1 <sup>er</sup> quartile	Médiane	3 <sup>e</sup> quartile	9 <sup>e</sup> quartile
France entière	566 104	15,07	9,09	9,08	10,23	12,32	16,53	24,03
Île-de-France	140 034	18,32	12,23	9,27	10,96	14,63	21,24	31,13
Reste de la France	426 070	14,00	7,48	9,04	10,12	11,87	15,23	21,10
Paris	40 295	18,33	12,41	9,21	10,84	14,61	21,25	31,11
Petite couronne	52 651	19,44	13,09	9,33	11,32	15,60	22,90	33,34
Grande couronne	47 088	17,07	10,89	9,26	10,76	13,75	19,33	28,63

Statistiques sur les salaires des salariés au sein de zones géographiques en 2007. Les salaires horaires sont en euros de 2007. « Reste de la France » fait référence à la France métropolitaine entière sans l'Île-de-France. À titre d'exemple, 10,96 dans la ligne 2 et la colonne 5 correspond au salaire des salariés se trouvant au 1<sup>er</sup> quartile de la distribution de salaires en Île-de-France.

Des statistiques descriptives sur les effectifs et la distribution des salaires au sein de nos zones d'étude sont reportées dans le tableau 1. En termes d'effectifs, l'Île-de-France contient 24,7 % des salariés français. Sans surprise, le salaire horaire moyen en Île-de-France (18,32 €) est supérieur au salaire horaire moyen dans le reste de la France (14,00 €). La différence de salaires augmente avec les quantiles le long des distributions. Ainsi, la différence n'est que de 0,23 € au 1<sup>er</sup> décile, mais elle atteint 10,00 € au dernier décile. Au sein de l'Île-de-France, on peut noter que le salaire moyen en petite couronne est supérieur à celui de Paris (19,44 € contre 18,33 €). Là encore, la différence de salaires augmente le long des distributions. Le gradient de la différence reste cependant assez faible.

3. La Nomenclature des activités françaises révision 1 (NAF rev1) est tout d'abord agrégée au niveau 3 de la Nomenclature économique de synthèse (NES) qui comprend 114 secteurs. 15 d'entre eux sont ensuite éliminés, soit par un manque de couverture des DADS (comme pour l'agriculture, la chasse et leurs services annexes (A01, A02 et A03), l'administration publique (R10), les activités associatives (R21) ou extraterritoriales (R22)), soit parce qu'ils ont trop peu d'observations (comme pour l'extraction de minerais métalliques (F11), l'extraction de minerais d'uranium (G13), le transport spatial (K06)), soit parce que les entreprises sont autorisées à centraliser leurs déclarations au niveau régional (comme pour les activités financières (L01, L02 et L03), et les postes et télécommunications (N11 et N12)), soit enfin parce qu'ils incluent une partie de production domestique (comme pour les services domestiques (P32)).

4. Une autre fraction, très faible, des observations extrêmes est supprimée lors de l'implémentation de la procédure comparant les distributions qui est décrite ci-dessous.

## MÉTHODOLOGIE

### Estimation de la productivité des salariés

Nous présentons tout d'abord notre approche permettant d'estimer le niveau de qualification de chaque salarié. Nous considérons, comme c'est courant dans la littérature (*e.g.* Combes *et al.* [2008]), que le salaire horaire d'un salarié  $i$  à chaque date  $t$  est le produit d'un terme local spécifique au secteur et de ses qualifications individuelles tel que :

$$w_{it} = B_{z(it)k(it)t} s_{it}. \quad (1)$$

où  $z(it)$  est la zone dans laquelle travaille le salarié  $i$  à la date  $t$ ,  $k(it)$  est le secteur dans lequel il travaille,  $B_{zkt}$  est un terme spécifique à la zone  $z$  et au secteur  $k$  lors de l'année  $t$  et  $s_{it}$  représente les qualifications du salarié.

Une implication directe de cette équation est que les différences de salaires entre zones reflètent à la fois les différences spatiales de qualifications entre salariés et les différences de productivité locale propres à chaque secteur. Les différences de productivité locale sont à leur tour dues à des différences de prix de vente des biens, de coût des facteurs (hors travail) et biens intermédiaires de production, et de technologie. Il existe, par exemple, des différences de prix de vente des biens de consommation quand les producteurs des zones ayant un meilleur accès aux marchés peuvent facturer des prix plus élevés que les producteurs des zones ayant un moins bon accès. Les différences de coût des facteurs locaux autres que le travail, comme le loyer à payer pour l'utilisation des terrains, peuvent être le résultat de différences de demande locale pour le sol<sup>5</sup>. Enfin, les différences de technologie peuvent refléter des différences d'avantages naturels ou d'interactions locales entre salariés. Ces interactions sont plus généralement qualifiées d'économies d'agglomération et peuvent être engendrées par différents mécanismes (comme présentés dans la revue de littérature de Duranton et Puga [2004]).

Afin de dériver une spécification pouvant être estimée avec nos données à partir de l'équation (1), il est nécessaire de spécifier tant les qualifications individuelles ( $s_{it}$ ) que la productivité des zones pour chaque secteur ( $B_{zkt}$ ). Nous supposons que la productivité pour un secteur, une zone et une année peut s'écrire :

$$\log B_{zkt} = \beta_{zt} + \mu_{kt}, \quad (2)$$

où  $\beta_{zt}$  est un effet fixe zone-année et  $\mu_{kt}$  est un effet fixe secteur-année. Une grande partie des études proposée par la littérature essaie d'identifier séparément les différents déterminants de la productivité des zones. Nous appréhendons l'ensemble des différences de productivité entre zones *via* le terme  $\beta_{zt}$  mais leurs causes et leurs déterminants exacts ne sont pas ici notre objet d'étude<sup>6</sup>. Nous

5. Cela peut avoir lieu pour des raisons qui ne sont pas nécessairement liées à la production. Par exemple, les zones avec de meilleures aménités de consommation font face à une demande de sol plus élevée et, en conséquence, se caractérisent par des prix du sol plus élevés.

6. Par exemple, Combes *et al.* [2010] régressent  $\beta_{zt}$ , après l'avoir moyenné sur la période d'étude, sur un ensemble de caractéristiques locales incluant la densité d'emploi, l'accès au marché, et des mesures variées d'avantages naturels locaux. Ils établissent que la densité d'emploi a un effet causal sur la productivité locale. Plus de résultats sur l'identification des effets d'agglomération sont présentés dans Combes, Duranton et Gobillon [2011].

considérons deux zones du territoire (par exemple, Paris et sa petite couronne) et notre objectif est de comparer les distributions de salaires et de qualifications entre ces deux zones, une fois que cet effet local est retiré.

Concernant les qualifications individuelles, nous supposons que, pour un salarié  $i$ , elles vérifient :

$$\log s_{it} = X_{it}\varphi + \delta_i + \varepsilon_{it}, \quad (3)$$

où  $X_{it}$  est un ensemble de caractéristiques individuelles qui dépend du temps,  $\delta_i$  est un effet fixe individuel et  $\varepsilon_{it}$  est un terme aléatoire, les termes aléatoires étant supposés être indépendamment et identiquement distribués entre périodes et salariés. En combinant les équations (1), (2) et (3), nous obtenons une équation de demande (inverse) de travail :

$$\log w_{it} = \beta_{z(it)t} + \mu_{k(it)t} + X_{it}\varphi + \delta_i + \varepsilon_{it}. \quad (4)$$

Ainsi, nous spécifions le salaire comme une fonction des caractéristiques observables et inobservables des salariés (le carré de l'âge et un effet fixe individuel), le type de zone dans lequel ils sont employés (l'effet fixe zone-temps, qui correspond à la productivité locale dont bénéficient tous ses salariés), et le secteur (l'effet fixe secteur-temps). Nous interprétons les effets fixes individuels comme les effets des qualifications inobservées<sup>7</sup>. Comme nous ne disposons que du carré de l'âge (l'âge étant aussi implicitement contrôlé grâce aux divers effets fixes) comme caractéristique individuelle observable, nous nous attendons à ce qu'une grande partie des écarts de salaires entre salariés soit captée par les effets fixes individuels, puisque ceux-ci capturent, entre autres, le rôle de l'éducation. Pour simplifier la rédaction, nous parlons dans ce qui suit de qualifications pour évoquer ces effets fixes<sup>8</sup>.

Afin de pouvoir estimer séparément des effets fixes salariés, secteurs et zones, il est nécessaire de disposer d'observations multiples pour chaque salarié et suffisamment de mobilité entre les secteurs et les zones. Les données que nous utilisons vérifient aisément ces conditions. L'identification des effets secteur-temps et zone-temps nécessite aussi la normalisation d'un effet secteur chaque année et d'un seul effet zone que nous fixons ici à zéro. Notons, par ailleurs, que les variations temporelles des effets zones sont identifiées à partir des salariés qui restent dans la même zone. L'identification est assurée par le fait que, pour ces salariés, l'effet individuel reste le même alors que l'effet zone change. Les différences d'effets entre zones sont, quant à elles, identifiées à partir des salariés qui changent de zone. En effet, leur effet individuel reste le même, alors que l'effet zone-temps change. Cette identification séparée des qualifications individuelles et de la productivité des zones requiert cinq hypothèses que nous évoquons maintenant sans entrer dans les détails<sup>9</sup>.

7. La spécification du logarithme du salaire est additive. Prendre l'exponentielle de la spécification conduit à une spécification du salaire en niveau dans laquelle les effets zones et les effets individuels entrent de manière multiplicative. Le rendement d'un effet individuel inclut donc l'effet fixe zone et dépend donc de la zone. L'effet fixe zone capte en particulier l'effet de la densité. Notre spécification permet donc implicitement à la densité d'influencer le rendement des qualifications.

8. Pour simplifier la présentation, les résultats de l'estimation de l'équation (4) ne seront pas commentés. Les résultats d'une telle estimation sont discutés de manière détaillée par Combes *et al.* [2008].

9. Voir Combes *et al.* [2012b] pour une discussion plus approfondie.



Premièrement, nous supposons que les qualifications individuelles peuvent être ramenées à un indice uni-dimensionnel. Nous considérons ainsi que cet indice est un résumé valide de qualifications multiples dont il a été établi que l'effet varie avec la localisation (Bacolod, Blum et Strange [2009a], [2009b], [2010]).

Deuxièmement, nous supposons que les effets locaux affectent le salaire de tous les salariés de la même façon. Toutefois, il semble que les salariés les plus qualifiés bénéficient plus d'être dans des zones plus productives (Wheeler [2001], Glaeser et Resseger [2010], Bacolod, Blum et Strange [2009a]). Une manière d'éviter ce problème est de comparer des salariés tous issus de la même zone. Ainsi nous confirmons nos résultats en comparant les salariés immobiles aux salariés mobiles au sein des zones. Plus précisément, nous pouvons comparer sans risque les salariés immobiles aux salariés mobiles pour chaque quantile de la distribution de qualifications quand les effets d'agglomération varient de façon monotone avec les qualifications individuelles.

Troisièmement, notre spécification suppose une substituabilité parfaite entre salariés (en particulier, à l'intérieur des entreprises), même si des articles montrent que des complémentarités entre salariés existent (Ciccone et Peri [2006], Moretti [2004]). Nous laissons ce problème de côté, étant donné qu'il est difficile de prendre en compte la complémentarité dans l'analyse et que ce n'est pas le sujet de notre étude.

Quatrièmement, nous considérons que les qualifications des salariés ne varient pas au cours du temps. Ainsi, nous supposons implicitement qu'elles ne sont pas influencées par la productivité locale alors qu'il a été montré qu'il existe des effets d'apprentissage en milieu urbain (Glaeser et Maré [2001], De La Roca et Puga [2012]). Dans cette perspective, la comparaison des salariés immobiles aux salariés mobiles d'une même zone revêt un intérêt particulier, puisqu'il est vraisemblable que leur expérience urbaine soit comparable.

Enfin, notre dernière hypothèse est que la migration entre zones est exogène, alors que la littérature montre que les salariés réagissent aux opportunités d'emploi (Baum-Snow et Pavan [2012]). En fait, ce problème est atténué dans notre étude par l'inclusion dans notre spécification économétrique d'effets fixes zone-temps qui captent les différences spatiales moyennes d'opportunités d'emploi. Par ailleurs, les choix de localisation dépendent non seulement des salaires mais aussi d'autres facteurs comme les préférences idiosyncratiques.

Malgré ses limites, notre méthode d'estimation des effets individuels a le double mérite de la transparence et de la comparabilité avec nos études antérieures. Notons aussi que notre méthodologie de comparaison des distributions de qualification entre zones pourrait s'appliquer de la même manière à des mesures de qualification individuelles évaluées par n'importe quelle autre méthode.

## Comparaison des distributions de qualifications et de salaires entre zones

Nous décrivons maintenant la deuxième partie de notre méthodologie qui permet de comparer les distributions des effets individuels et des salaires entre deux zones notées  $i \in \{1, 2\}$ . Plus précisément, notre approche évalue dans quelle mesure la distribution des effets individuels et des salaires dans une zone

est tronquée, translatée et dilatée par rapport à la distribution de la même variable dans l'autre zone. Nous suivons la méthodologie développée dans Combes, Duranton, Gobillon, Puga et Roux [2012a] et enrichie au cas de double troncation par Combes *et al.* [2012b]. Nous nous concentrons sur les quatre transformations (translation, dilatation, troncations à gauche et à droite) parce qu'elles nous permettent de rendre compte de presque toutes les différences entre zones de distribution de salaires et d'effets individuels<sup>10</sup>. L'interprétation économique des quatre paramètres est discutée dans la section suivante.

Par mesure de simplicité, nous exposons formellement ici seulement le cas où uniquement une translation et une dilatation sont permises<sup>11</sup>. La présentation qui suit fait référence aux salaires mais son implémentation aux effets individuels est identique.

Considérons une situation où un salaire quelconque  $w$  dans la distribution 1 est transformé linéairement en un salaire  $Dw + A$  dans la distribution 2, où  $A$  est un paramètre de translation et  $D$  est un paramètre de dilatation. Le paramètre  $A$  peut être positif ou négatif, selon qu'il y a une translation vers la droite ou vers la gauche, et le paramètre  $D$  peut être supérieur ou inférieur à 1, selon qu'il y a une dilatation ou compression. Il est à noter que pour pouvoir interpréter  $D > 1$  comme une dilatation et  $D < 1$  comme une compression, il est nécessaire de normaliser la moyenne de la distribution 1 à 0, c'est-à-dire de soustraire le salaire moyen des individus de la distribution 1 aux salaires de tous les individus.

Plus formellement, notons  $\lambda_1(u)$  le quantile de la distribution 1 au rang  $u$ . La transformation de valeurs décrite ci-dessus entraîne la relation suivante entre les deux distributions :

$$\lambda_2(u) = D\lambda_1(u) + A, \quad \text{pour } u \in [0,1]. \quad (5)$$

La procédure d'estimation est basée sur cette relation entre quantiles. Notons  $\theta = (A, D)$  l'ensemble des paramètres à estimer. L'équation (5) peut être réécrite comme un continuum d'égalités de la forme  $m_\theta(u) = 0$  pour tout  $u \in [0, 1]$  où :

$$m_\theta(u) = \lambda_2(u) - D\lambda_1(u) - A. \quad (6)$$

Il est possible d'estimer  $\theta$  en minimisant la somme des carrés des contreparties empiriques de  $m_\theta(u)$ ,  $u \in [0, 1]$  obtenues en remplaçant les fonctions quantiles par leur estimateur. Cependant, rien ne garantit que le critère de minimisation, quand il est appliqué à la transformation de la distribution 2 en la distribution 1 (la transformation opposée), mène aux mêmes valeurs estimées de  $A$  et  $D$ . Ainsi, nous utilisons aussi un continuum d'égalités obtenu en intervertissant les distributions 1 et 2 :

$$\lambda_1(u) = \frac{1}{D}\lambda_2(u) - \frac{A}{D}, \quad \text{pour } u \in [0,1]. \quad (7)$$

10. Notre approche permet de comparer deux distributions différentes en utilisant un nombre restreint de paramètres qui permettent de les relier ensemble. Nous comparons ces distributions pour l'ensemble de la population salariée et pour certains sous-groupes. Les décompositions de différences entre distributions présentées par Fortin, Lemieux et Firpo [2011] et utilisées avec des équations de salaires à effets fixes par Card, Heining et Kline [2013] constituent une approche alternative et complémentaire à la nôtre. Notre approche permet de comparer directement des distributions de manière semi-paramétrique en prenant en compte tous leurs moments, alors que les approches de type décomposition permettent d'examiner le rôle d'un facteur donné pour expliquer une différence de moyenne ou de quantile entre deux groupes.

11. Voir Combes *et al.* [2012a] et [2012b] pour une présentation approfondie.

Cette équation conduit à un nouvel ensemble d'égalités :  $\tilde{m}_0(u) = 0$ , où :

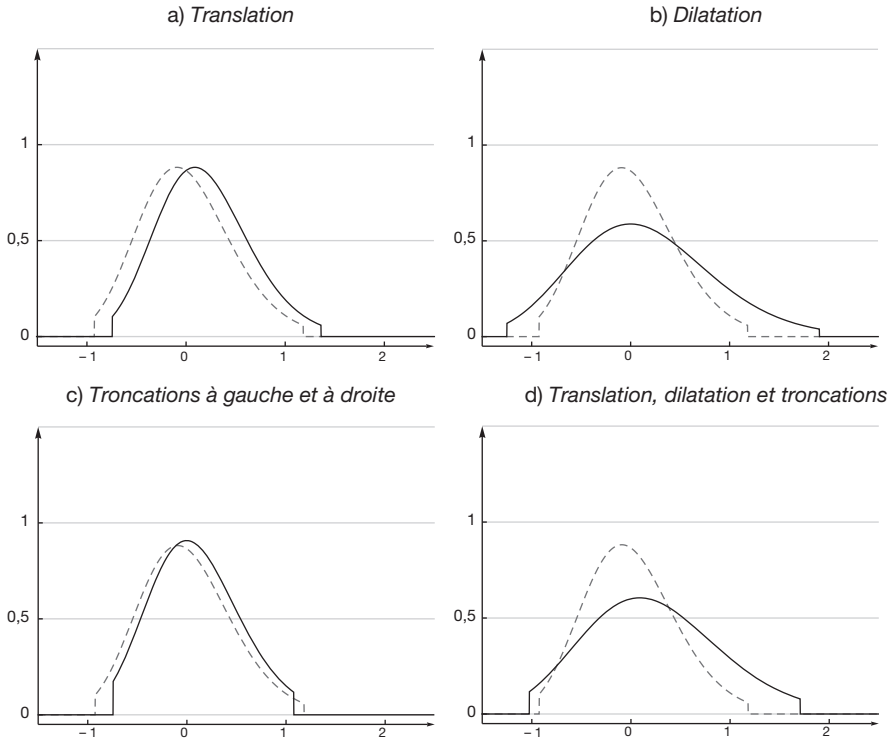
$$\tilde{m}_0(u) = \lambda_1(u) - \frac{1}{D} \lambda_2(u) + \frac{A}{D}. \quad (8)$$

Notons  $\hat{m}_0(u)$  la contrepartie empirique de  $m_0(u)$  et  $\hat{\tilde{m}}_0(u)$  la contrepartie empirique de  $\tilde{m}_0(u)$  obtenues lorsqu'on remplace les fonctions quantiles  $\lambda_i(u)$ ,  $i \in \{1, 2\}$  par leur contrepartie empirique. L'estimateur de  $\theta$  est obtenu en minimisant la somme des carrés de  $\hat{m}_0(u)$  et  $\hat{\tilde{m}}_0(u)$  :

$$\hat{\theta} = \arg \min_{\theta} M(\theta), \text{ où } M(\theta) = \int_0^1 [\hat{m}_0(u)]^2 du + \int_0^1 [\hat{\tilde{m}}_0(u)]^2 du. \quad (9)$$

Les intervalles de confiance sont obtenus par bootstrap. L'implémentation est détaillée dans Combes *et al.* [2012a]. Finalement, nous construisons une mesure du pouvoir explicatif de la spécification à partir du rapport entre la moyenne des carrés des différences de quantiles entre la distribution 1 transformée et la distribution 2, et la moyenne des carrés des différences de quantiles entre les distributions 1 et 2. Cette mesure est un pseudo- $R^2$  défini comme  $R^2 = 1 - M(\hat{\theta}) / M(0,1)$ .

Figure 1. Quatre transformations possibles de la distribution de logarithmes des salaires



Ligne en pointillés : zone 1 hypothétique.  
 Ligne continue : zone 2 hypothétique obtenue à partir d'une transformation de la distribution de la zone 1.

Afin d'expliquer les intuitions de notre approche, nous représentons par les figures 1(a) à (d) les quatre transformations discutées ci-dessus. Pour toutes les figures, la courbe en pointillés représente une distribution hypothétique du logarithme des salaires dans la zone 1, et la courbe continue représente une distribution hypothétique du logarithme des salaires dans la zone 2. Pour cette représentation, nous supposons que la distribution de la zone 1 suit une loi de Fisher tronquée, mais notre méthode d'estimation ne nécessite pas de spécifier la forme fonctionnelle des distributions. Notre approche permet d'estimer dans quelle mesure la distribution de la zone 1 (qui est utilisée comme référence) peut être transformée avec une translation et une dilatation pour approcher la distribution de la zone 2. Nous considérons aussi l'extension où la distribution de la zone 2 peut être une version tronquée à gauche et à droite de la distribution de la zone 1, et nous notons  $\underline{S}$  le rang auquel a lieu la troncation à gauche et  $1 - \bar{S}$  le rang auquel a lieu la troncation à droite.

La figure 1(a) considère le cas d'une simple translation positive ( $A > 0, D = 1, \underline{S} = \bar{S} = 0$ ) entre les deux distributions de salaires. Celles-ci ont alors la même forme. La distribution des salaires de la zone 2 se trouve à droite de celle de la zone 1. On peut noter comme repère visuel que les deux pics de densité ont la même hauteur.

La figure 1(b) considère le cas d'une dilatation ( $D > 1, A = \underline{S} = \bar{S} = 0$ ). Dans ce cas, la distribution des salaires de la zone 2 est plus étendue vers la gauche et vers la droite que celle de la zone 1. Un deuxième repère visuel est donné par le pic plus bas de la distribution de salaires de la zone 2<sup>12</sup>.

La figure 1(c) considère le cas des deux troncations à gauche et à droite simultanément ( $\underline{S} > 0, \bar{S} > 0, A = 0, D = 1$ ). Visuellement, la distribution des salaires de la zone 2 a des queues de distribution plus courtes. Si la troncation à gauche est plus importante que la troncation à droite ( $\underline{S} > \bar{S}$ ), comme c'est le cas dans notre exemple, il existe alors une asymétrie entre les deux distributions. La troncation des queues de distribution entraîne aussi un pic plus élevé. Le cas d'une troncation « négative » correspond à une situation où la distribution de la zone 1 est plus tronquée que la distribution de la zone 2.

Enfin, la figure 1(d) considère la situation où la distribution des salaires de la zone 2 est translatée à droite, dilatée, tronquée à gauche et à droite relativement à la distribution des salaires de la zone 1. Un point important à noter est que la dilatation et les troncations vont, en partie, dans des directions opposées. Alors que les troncations font remonter le pic de la distribution et réduisent la dispersion en coupant les queues de distribution, la dilatation abaisse le pic et augmente la dispersion. Cependant, ces transformations ne se compensent pas complètement. En fait, si par rapport à la distribution de la zone 1, celle de la zone 2 a des queues moins épaisses, les différences sont captées tant par une dilatation qu'une troncation. Si la distribution de la zone 2 a des queues plus épaisses, les différences sont captées par une dilatation et une troncation négative. Ainsi, quand la dilatation et les troncations sont considérées simultanément,

---

12. Il est aussi à noter que les deux pics ne sont pas exactement alignés. Cela est dû au fait que l'asymétrie de la loi de Fisher implique que le mode de la distribution est à gauche de la moyenne et est dilaté comme le reste de la distribution.

il faut garder à l'esprit que les coefficients estimés doivent être interprétés de façon conjointe. Il est aussi à noter que, visuellement, il est difficile de détecter séparément la dilatation et les troncations. Cette détection ne peut être effectuée qu'à partir des résultats des estimations.

## SALAIRES ET SALARIÉS EN ÎLE-DE-FRANCE ET DANS LE RESTE DE LA FRANCE

Dans cette section, nous effectuons une comparaison des distributions de salaires et des effets individuels entre l'Île-de-France et le reste de la France. Il est utile de garder en mémoire que les salaires reflètent à la fois le niveau de qualification et la valeur de ces qualifications, alors que les effets individuels ne reflètent que le niveau de qualification.

Pour aider à interpréter les résultats qui suivent, considérons une situation hypothétique où la distribution des effets individuels serait la même en Île-de-France et dans le reste de la France. On interprète alors une translation à droite de la distribution des salaires du reste de la France pour approcher celle de l'Île-de-France comme le résultat d'économies d'agglomération ou d'avantages naturels locaux présents en Île-de-France. Une dilatation ( $D > 1$ ), telle que celle observée dans la figure 1(a), reflète des salaires plus élevés en Île-de-France pour les salariés les mieux payés et des différences qui augmentent avec le salaire. Cela correspond à une situation pour laquelle les effets d'agglomération ou les avantages naturels ont alors des effets de plus en plus forts lorsque le niveau de qualification s'élève. Une compression ( $D < 1$ ) implique le contraire d'une dilatation.

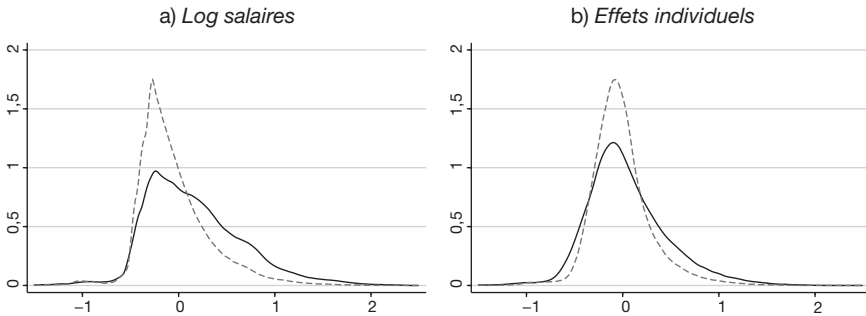
Considérons maintenant une situation où la distribution des effets individuels diffère entre l'Île-de-France et le reste de la France. L'interprétation des paramètres de translation et de dilatation n'est plus la même. Une translation vers la droite de la distribution des effets individuels dans le reste de la France pour approcher celle de l'Île-de-France reflète la sur-représentation des salariés les plus qualifiés en Île-de-France et une sous-représentation des salariés les moins qualifiés. Une dilatation indique, au contraire, une sur-représentation à la fois des salariés les plus qualifiés et des salariés les moins qualifiés. On peut déduire la part des effets d'agglomération ou d'avantages naturels locaux dans l'explication des différences de salaires entre zones en déduisant les effets des différences de qualification. Typiquement, si les distributions de qualification sont très similaires, les différences de salaires reflètent principalement des différences d'effets d'agglomération ou d'avantages naturels alors que si les différences entre distributions sont de même ampleur pour salaire et qualification, les effets d'agglomération et d'avantages naturels sont très faibles.

Enfin, nous nous intéressons aussi brièvement à de possibles troncations des queues de distribution. Une troncation à gauche (ou à droite) peut avoir lieu pour deux raisons. La première, que nous pouvons appeler « véritable troncation », reflète l'absence complète de salariés ayant des effets individuels correspondant à ceux de la queue de distribution à laquelle cette troncation s'applique. Par exemple, une véritable troncation à gauche de la distribution des qualifications traduit une absence des salariés les moins qualifiés. Cette absence complète de

certain types de salariés est une prédiction de nombreux modèles de tri spatial. Les modèles de Nocke [2006] et de Venables [2011] prédisent que la distribution des effets individuels est tronquée à gauche dans les plus grandes villes, ce qui y reflète la sélection de salariés plus qualifiés. Le tri spatial dans le modèle de Eeckhout *et al.* [2014] implique une dilatation de la distribution de qualifications dans les villes les plus grandes relativement aux plus petites, alors que le modèle de Behrens *et al.* [2012] prédit une translation vers la droite et que le modèle de Davis [2009] prédit un mélange de translation, de dilatation, et de troncation. Par ailleurs, certaines contraintes institutionnelles, tel le salaire minimum, peuvent être plus liantes dans les zones moins productives, ce qui engendre de la troncation à gauche<sup>13</sup>.

En relation avec la figure 1(d), des troncations peuvent aussi être estimées pour une raison fondamentalement différente. Imaginons, par exemple, que la dilatation de la distribution des effets individuels soit plus forte au milieu de la distribution et moins forte à ses extrémités. Cela peut arriver, par exemple, si les rendements de l'agglomération n'évoluent pas de façon monotone avec le niveau de qualification. Dans nos estimations, cela peut apparaître alors comme une dilatation et des troncations aux extrémités. Autrement dit, une troncation peut être estimée pour atténuer l'effet de la dilatation dans les queues de distribution. On peut alors parler de « troncation affaiblissante<sup>14</sup> ». Pour distinguer une véritable troncation d'une troncation affaiblissante, notons qu'une troncation affaiblissante qui accompagne une dilatation maintient une distribution plus étalée. Une véritable troncation nécessite, au contraire, que la distribution transformée soit plus resserrée.

Figure 2. Distribution des logarithmes des salaires et des effets individuels en Île-de-France et dans le reste de la France, 2007



Ligne continue : Île-de-France.  
Ligne en pointillés : Reste de la France.

13. En effet dans les zones les moins productives, un salarié avec effet individuel faible peut être dans l'impossibilité d'attendre une productivité suffisamment élevée pour atteindre même le salaire minimum. Cette contrainte est moins forte dans une zone plus productive comme le montre l'équation (1).

14. De la même manière, des troncations négatives peuvent apparaître pour atténuer dans les queues de distribution les effets d'une compression.

Tableau 2. Log des salaires en Île-de-France par rapport au reste de la France, 2007

	$\hat{A}$	$\hat{D}$	$\hat{S}$	$\hat{S}$	R <sup>2</sup>
$w_R$ et $w_{IdF}$ (translation)	0,209 (0,002)*	–	–	–	0,717
$w_R$ et $w_{IdF}$ (+ dilatation)	0,213 (0,002)*	1,420 (0,004)*	–	–	0,978
$w_R$ et $w_{IdF}$ (+ truncations à gauche et à droite)	0,286 (0,003)*	1,816 (0,014)*	0,025 (0,003)*	0,060 (0,002)*	0,998

\* : pour  $\hat{A}$ ,  $\hat{S}$ , et  $\hat{S}$  significativement différents de 0 à 5 % ; pour  $\hat{D}$ , significativement différent de 1 à 5 %.

$w_{IdF}$  : distribution des salaires en Île-de-France.

$w_R$  : distribution du logarithme des salaires dans le reste de la France.

140 034 observations en Île-de-France. 426 070 observations dans le reste de la France.

La figure 2(a) représente la distribution du logarithme des salaires en Île-de-France et dans le reste de la France en 2007.<sup>15</sup> Le tableau 2 fournit les résultats des estimations qui se rapportent à la comparaison de ces deux distributions. Dans la première ligne du tableau, nous approchons la distribution des salaires en Île-de-France à partir de celle du reste de la France uniquement au moyen d'une translation. Le coefficient estimé pour cette dernière est de 0,209. Cela indique que les salaires sont en moyenne ( $= \exp(0,209) - 1 =$ ) 23 % plus élevés en Île-de-France que dans le reste du pays. Cette prédiction est très proche de la différence empirique moyenne documentée dans le tableau 1. Notons aussi que ce paramètre explique à lui seul plus de 70 % des écarts au carré entre les quantiles de ces deux distributions.

Il apparaît toutefois clairement sur la figure 2(a) qu'une translation est insuffisante pour transformer la distribution des salaires dans le reste de la France et bien approcher la distribution des salaires en Île-de-France. Dans la deuxième ligne de résultats du tableau 2, nous ajoutons un paramètre de dilatation à notre spécification. Nous obtenons une valeur estimée de 1,420, alors que la valeur estimée pour la translation reste quasiment la même à 0,213 (ce qui implique des salaires 24 % plus élevés). Cette dilatation apparaît clairement dans la figure 2(a). La dernière colonne montre aussi que ce second paramètre améliore grandement la qualité de l'approximation puisque 98 % des écarts au carré entre les quantiles des deux distributions sont expliqués par les deux paramètres. Nous utilisons cette spécification avec une translation et une dilatation comme spécification de référence.

Ce coefficient de 1,420 estimé pour la dilatation est important. Si on le combine avec le coefficient estimé de 0,213 pour la translation, on arrive à un salaire prédit au premier décile qui est 3,0 % plus élevé en Île-de-France relativement au reste de la France. Pour un salaire au premier quartile, ces deux paramètres prédisent un écart de 8,0 %. Pour la médiane, la différence prédite est de 15,5 %. Pour le troisième quartile et le dernier décile, les différences prédites sont respectivement de 28,2 % et de 47,0 %. Dans les données, le tableau 1 indique que les différences empiriques sont de 2,5 %, 8,3 %, 23,3 %, 39,5 % et 47,5 % respectivement. Cela suggère que nos deux coefficients permettent de bien approcher les salaires

15. Sauf avis contraire, toutes les figures sont pour 2007. Par ailleurs, la moyenne pour le groupe de référence est normalisée à zéro.

d'Île-de-France par une transformation de ceux du reste de la France, avec une légère sous-estimation pour la médiane et le troisième quartile.

La troisième ligne de résultats du tableau 2 considère une spécification plus complète avec deux paramètres de troncation. L'introduction de ces deux paramètres supplémentaires modifie légèrement l'estimation du paramètre de translation qui augmente, pour s'élever à 0,286, et celui du paramètre de dilatation qui augmente aussi à 1,816. Les deux paramètres de troncation suggèrent une troncation aux centiles 2,5 et  $100 - 6 = 94$  de la distribution des salaires dans le reste de la France afin d'approcher celle de l'Île-de-France. On note qu'il s'agit là d'un cas de troncation affaiblissante. La figure 2(a) montre clairement que la distribution des salaires est plus étalée en Île-de-France. Il se trouve simplement que la dilatation est plus forte au milieu de la distribution que dans les queues. Pour le bas de la distribution, la troncation peut résulter de l'existence du salaire minimum. En effet, les salaires particulièrement bas prédits par un coefficient de dilatation très élevé ne sont pas observés en Île-de-France. On améliore alors l'approximation en tronquant cette distribution<sup>16</sup>. De la même manière, une troncation à droite de la distribution permet de capturer une queue de distribution à droite moins épaisse en Île-de-France que ne le prédirait une simple dilatation de la distribution de salaires du reste de la France. Malgré tout l'intérêt que peut avoir cette spécification plus complète avec quatre paramètres, et dans un souci de simplicité, nous ne reportons que les résultats d'estimations avec une translation et une dilatation dans la section suivante.

Tableau 3. Log des effets individuels en Île-de-France par rapport au reste de la France, 2007

	$\hat{A}$	$\hat{D}$	$\hat{S}$	$\hat{S}$	R <sup>2</sup>
$\delta_R$ et $\delta_{IdF}$ (translation)	0,059 (0,001)*	–	–	–	0,238
$\delta_R$ et $\delta_{IdF}$ (+ dilatation)	0,063 (0,001)*	1,430 (0,004)*	–	–	0,984
$\delta_R$ et $\delta_{IdF}$ (+ troncations à gauche et à droite)	0,060 (0,002)*	1,626 (0,001)*	0,037 (0,002)*	0,020 (0,001)*	0,996

\* : pour  $\hat{A}$ ,  $\hat{S}$ , et  $\hat{S}$  significativement différents de 0 à 5 % ; pour  $\hat{D}$  significativement différent de 1 à 5 %.

$\delta_{IdF}$  : distribution des salaires en Île-de-France.

$\delta_R$  : distribution du logarithme des salaires dans le reste de la France.

140 034 observations en Île-de-France. 426 070 observations dans le reste de la France.

La figure 2(b) représente les distributions des effets individuels en Île-de-France et dans le reste de la France. Le tableau 3 réplique le tableau 2 pour les effets individuels. La première ligne de résultats indique que les effets individuels

16. Le salaire minimum devrait être moins contraignant en Île-de-France (où la productivité est plus élevée) que dans le reste de la France. En l'absence de dilatation, le salaire minimum implique une troncation négative de la queue de la distribution quand on transforme la distribution des salaires du reste de la France pour approcher la distribution des salaires en Île-de-France. Avec une forte dilatation, le salaire minimum devient plus contraignant pour la distribution transformée qui approche celle des salaires en Île-de-France et cela implique alors une troncation positive telle que celle observée. Toutefois, il n'est pas possible d'interpréter les coefficients estimés pour la troncation de la queue de la distribution comme un effet du salaire minimum puisque d'autres forces peuvent être en jeu.



(en logarithmes) sont en moyenne 0,059 plus élevés en Île-de-France que dans le reste de la France. Cela correspond à une différence de  $(\exp(0,059) - 1 =) 6\%$ .

Étant donné que les salaires sont en moyenne 24 % plus élevés en Île-de-France que dans le reste du pays, ces résultats indiquent qu'environ le quart des différences de salaire entre l'Île-de-France et le reste de la France peut être expliqué par la présence de salariés plus qualifiés en Île-de-France. Ce résultat est à rapprocher de celui de Combes *et al.* [2012b] qui indiquent que les différences de qualifications expliquent environ 40 % des différences de salaires entre les zones d'emplois plus denses que la médiane et les zones d'emplois moins denses en France. Notons aussi que ce paramètre de translation n'explique que 24 % des différences au carré entre les quantiles de ces deux distributions.

La raison de cette faible performance se voit clairement sur la figure 2(b). La distribution des effets individuels en Île-de-France est beaucoup plus étalée que celle du reste du pays. Comme le montre la deuxième ligne de résultats du tableau 3, l'ajout d'un paramètre de dilatation améliore très fortement la qualité de l'approximation. Le paramètre de translation reste néanmoins quasiment le même (0,063 au lieu de 0,059). Le paramètre de dilatation estimé est de 1,430, ce qui est très proche de son homologue pour les salaires.

Dans la troisième ligne de résultats du tableau 3, l'ajout de deux paramètres de troncation laisse le paramètre de translation presque inchangé à 0,060, fait augmenter le paramètre de dilatation à 1,626 et révèle l'existence d'une troncation de quelques centiles à gauche comme à droite de la distribution d'effets individuels en Île-de-France. À nouveau, ces paramètres de troncation améliorent (faiblement) la qualité de l'approximation en affaiblissant les effets du paramètre de dilatation dans les deux queues de la distribution.

Nous tirons un certain nombre de conclusions de ces résultats. Tout d'abord, nous confirmons l'existence bien connue de salaires plus élevés, de 24 %, en Île-de-France relativement au reste du pays. Toutefois, les différences de salaires entre l'Île-de-France et le reste de la France sont loin de pouvoir se résumer à cet écart de moyenne. La distribution des salaires est beaucoup plus étalée en Île-de-France. Les salaires les plus bas sont quasiment les mêmes que dans le reste du pays, alors que les salaires au dernier décile sont presque 50 % plus élevés. Comme la dilatation de la distribution des salaires est la même que celle des effets individuels, nous confirmons que l'essentiel des différences d'étalement entre distributions des salaires entre l'Île-de-France et le reste de la France est expliqué par des différences dans la distribution des qualifications. Alors que les salariés les plus qualifiés sont très fortement sur-représentés en Île-de-France, les salariés les moins qualifiés le sont quasiment tout autant. Toutefois, les différences de qualification, si elles expliquent les différences d'étalement entre les distributions de salaires en Île-de-France et dans le reste du pays, n'expliquent qu'environ le quart des différences de niveaux. Après avoir pris en compte la plus grande dilatation des effets individuels et leur niveau plus élevé, les salaires restent plus élevés en Île-de-France de 16 %  $(= \exp(0,213 - 0,063) - 1)$ , ce qui reflète la plus grande productivité de l'Île-de-France par rapport au reste du pays<sup>17</sup>.

17. Cet écart est généralement attribué aux économies d'agglomération. Les résultats de Combes *et al.* [2012a] indiquent que les établissements (principalement manufacturiers) de l'aire urbaine de Paris sont en moyenne 13 % plus productifs que ceux des autres grandes villes qui sont eux-mêmes 4 % plus productifs que ceux des autres villes.

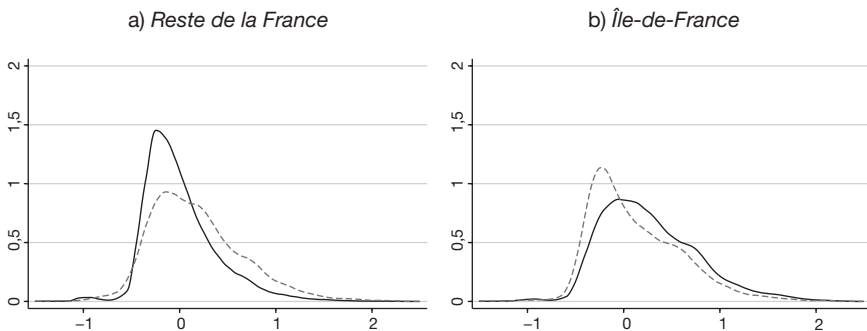
Enfin, notons que ces résultats sont cohérents en partie avec les théories de tri spatial proposées par Behrens *et al.* [2014] (existence d'une translation de la distribution des qualifications et d'une prime d'agglomération) et Eeckhout *et al.* [2014] (qui prédit une dilatation). En revanche, nous ne trouvons aucun élément en faveur d'une segmentation parfaite des salariés selon leur productivité. Tant pour les salaires que pour les effets individuels, les distributions pour l'Île-de-France et le reste de la France se chevauchent bien plus qu'elles ne se séparent.

## COMPARAISON DÉTAILLÉE ENTRE L'ÎLE-DE-FRANCE ET LE RESTE DE LA FRANCE : MIGRANTS ET NATIFS

### Migrants et non-migrants, Île-de-France et reste de la France, 2004-2007

Après avoir éliminé les différences de salaires entre zones qui s'appliquent de manière systématique à tous les salariés, on peut imaginer trois raisons pour lesquelles la distribution des effets individuels peut différer entre l'Île-de-France et le reste de la France. La première concerne la migration différenciée des salariés selon leur niveau de qualification. Il se peut que les salariés, en fonction de leur niveau de qualification, aient des préférences différentes en matière de localisation ou qu'ils subissent des coûts de la vie différents. La seconde raison pour laquelle les qualifications peuvent varier entre zones est liée à la façon dont les qualifications sont acquises au sein de ces zones. Par exemple, des qualifications plus élevées à la génération précédente amènent à des qualifications plus élevées pour les natifs de la génération actuelle pour autant que les qualifications soient en partie transmises au sein des familles. Il se peut aussi que les performances du système éducatif, ou les dotations en universités ou grandes écoles, varient. Enfin, la troisième raison, qui est sans doute importante, est que des effets d'apprentissage peuvent affecter des employés de qualification différente de manière

Figure 3. Distribution du logarithme des salaires en Île-de-France et dans le reste de la France pour les migrants et les non-migrants, 2004-2007



Ligne continue : Non-migrants.  
Ligne en pointillés : Migrants vers l'autre zone.

différente selon les zones Glaeser et Maré [2001], De La Roca et Puga [2012]. Dans ce dernier cas, les effets individuels estimés par notre approche ne reflètent plus uniquement les qualifications des salariés mais un mélange de qualification et d'effet local.

Pour pouvoir aller plus loin dans notre analyse, nous distinguons maintenant les migrants et les non-migrants tout en restant dans la comparaison entre l'Île-de-France et le reste du pays. Au sein de chaque zone, migrants et non-migrants font face à des effets locaux similaires au sein de chaque quantile. Si l'on fait l'hypothèse que les effets locaux sont monotones avec le niveau de qualification, on peut identifier quels sont les quantiles affectés par le tri spatial et dans quelle mesure<sup>18</sup>.

Tableau 4. Log des salaires des migrants et des non-migrants d'Île-de-France et du reste de la France, 2004-2007

	$\hat{A}$	$\hat{D}$	R <sup>2</sup>	Obs. 1	Obs. 2
$w_{RR}$ et $w_{R \rightarrow IdF}$	0,171 (0,006)*	1,325 (0,015)*	0,976	281 764	6 105
$w_{IdFIdF}$ et $w_{R \rightarrow IdF}$	-0,091 (0,063)	0,957 (0,110)	0,986	84 279	6 105
$w_{IdFIdF}$ et $w_{IdF \rightarrow R}$	-0,145 (0,006)*	0,946 (0,010)*	0,961	84 279	6 497
$w_{RR}$ et $w_{IdF \rightarrow R}$	0,117 (0,006)*	1,311 (0,014)*	0,952	281 764	6 497

\* : pour  $\hat{A}$  significativement différent de 0 à 5 % ; pour  $\hat{D}$  significativement différent de 1 à 5 %.

$w_{ii}$  : distribution du logarithme des salaires des non-migrants dans le reste de la France ( $I = R$ ) ou en Île-de-France ( $i = IdF$ ).

$w_{i \rightarrow j}$  : distribution du logarithme des salaires des migrants venant d'une zone  $i$  ( $i \in \{R, IdF\}$ ) vers une zone  $j$  ( $j \in \{R, IdF\}$ ).

Obs. 1 et Obs. 2 : nombre d'observations dans le premier groupe (référence) et dans le second.

Pour identifier migrants et non-migrants, nous confrontons pour chaque salarié sa localisation (Île-de-France ou reste de la France) en 2004 à celle de 2007. La figure 3(a) représente la distribution du logarithme des salaires des non-migrants du reste de la France et des migrants qui se déplacent du reste de la France vers l'Île-de-France. La figure 3(b) représente les distributions correspondantes pour les non-migrants et les migrants d'origine francilienne.

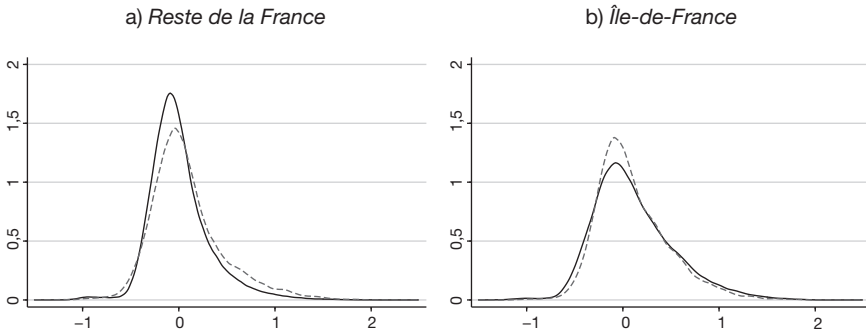
Le tableau 4 reporte les résultats de différentes comparaisons entre ces distributions en utilisant un paramètre de translation et un paramètre de dilatation. La première ligne de résultats évalue comment la distribution des salaires des migrants depuis le reste de la France vers l'Île-de-France peut être approchée en traduisant et en dilatant la distribution des salaires des non-migrants du reste de la France. Les résultats indiquent que les migrants reçoivent des salaires plus élevés et plus dispersés. Toutefois, l'écart de salaire entre les migrants vers l'Île-de-France et les non-migrants du reste de la France est plus faible que l'écart

18. Selon les résultats de Wheeler [2001], Bacolod *et al.* [2009a] et Glaeser et Resseger [2010], il s'agit d'une hypothèse raisonnable.

de salaire moyen entre l'Île-de-France et le reste de la France tel qu'il est documenté dans le tableau 2. Cela est confirmé par la comparaison entre les migrants vers l'Île-de-France et les non-migrants de l'Île-de-France. Par rapport à ces derniers, les migrants ont un salaire plus faible et la distribution de leurs salaires est comprimée.

La troisième ligne du tableau 4 reporte des résultats concernant la comparaison entre la distribution des salaires des non-migrants en Île-de-France et celle des migrants vers le reste de la France. Les résultats montrent que les salaires des migrants sont plus faibles que ceux des non-migrants. Ils sont aussi plus comprimés. Toutefois, le salaire de ces migrants vers le reste de la France est plus élevé que celui des non-migrants du reste de la France. La distribution des salaires des migrants est aussi plus étalée.

Figure 4. Distribution du logarithme des effets individuels en Île-de-France et dans le reste de la France pour les migrants et non-migrants, 2004-2007



Ligne continue : Non-migrants.  
Ligne en pointillés : Migrants vers l'autre zone.

Tableau 5. Log des effets individuels des migrants et des non-migrants d'Île-de-France et du reste de la France, 2004-2007

	$\hat{A}$	$\hat{D}$	R <sup>2</sup>	Obs. 1	Obs. 2
$\delta_{RR}$ et $\delta_{R \rightarrow IdF}$	0,075 (0,005)*	1,252 (0,018)*	0,995	281 764	6 105
$\delta_{IdFIdF}$ et $\delta_{R \rightarrow IdF}$	-0,031 (0,017)	0,880 (0,135)	0,813	84 279	6 105
$\delta_{IdFIdF}$ et $\delta_{IdF \rightarrow R}$	-0,013 (0,005)*	0,881 (0,010)*	0,964	84 279	6 497
$\delta_{RR}$ et $\delta_{IdF \rightarrow R}$	0,093 (0,005)*	1,252 (0,014)*	0,975	281 764	6 497

\* : pour  $\hat{A}$  significativement différent de 0 à 5 % ; pour  $\hat{D}$  significativement différent de 1 à 5 %.  
 $\delta_{ij}$  : distribution du logarithme des effets individuels des non-migrants dans le reste de la France ( $i = R$ ) ou en Île-de-France ( $i = IdF$ ).  
 $\delta_{i \rightarrow j}$  : distribution du logarithme des effets individuels des migrants venant d'une zone  $i$  ( $i \in \{R, IdF\}$ ) vers une zone  $j$  ( $j \in \{R, IdF\}$ )

Obs. 1 et Obs. 2 : nombre d'observations dans le premier groupe (référence) et dans le second.

Les figures 4(a) et (b) reproduisent les figures 3(a) et (b) pour les effets individuels au lieu des salaires. La figure 4(a) représente la distribution des effets individuels pour les non-migrants du reste de la France et les migrants du reste de la France vers l'Île-de-France. La figure 4(b) représente les distributions correspondantes pour les salariés initialement employés en Île-de-France.

Le tableau 5 reporte les résultats des estimations qui correspondent aux comparaisons entre ces distributions en utilisant un paramètre de translation et un paramètre de dilatation. La première ligne de résultats montre que la distribution des effets individuels des migrants depuis le reste de la France vers l'Île-de-France est traduite vers la droite et dilatée relativement à la même distribution pour les non-migrants du reste de la France. En revanche, quand on compare ces mêmes migrants vers l'Île-de-France aux non-migrants de l'Île-de-France, on peut voir que leurs effets individuels sont maintenant plus faibles (même si l'écart est marginalement non significatif à 5 %) et proviennent d'une distribution plus comprimée.

La troisième ligne de résultats du tableau 5 montre que les migrants de l'Île-de-France vers le reste de la France ont en moyenne un effet individuel plus faible dont la distribution est compressée relativement à celle des non-migrants de l'Île-de-France. Toutefois, la quatrième ligne montre aussi que ces migrants de l'Île-de-France vers le reste de la France ont un effet individuel en moyenne plus élevé que celui des non-migrants dans leur zone de destination. La distribution des effets individuels pour les migrants est aussi dilatée par rapport à celle des non-migrants.

Nous tirons un certain nombre de conclusions de ces résultats. Tout d'abord, les migrants vers l'Île-de-France reçoivent des salaires plus élevés que leur population d'origine mais plus bas que ceux de leur population d'accueil, alors que le contraire est vrai pour les migrants vers le reste de la France. Cela contraste quelque peu avec les résultats de Combes *et al.* [2012b] sur les zones denses et moins denses où les migrants ont toujours un salaire moyen plus faible. Cela vient en partie du fait que les migrants vers l'Île-de-France ont en moyenne un effet individuel bien plus élevé que ceux qui restent dans le reste de la France. De même, les migrants qui viennent d'Île-de-France ont aussi un effet individuel en moyenne plus élevé que celui des non-migrants du reste de la France. Cet effet individuel des migrants vers le reste de la France est aussi en moyenne plus faible que celui de leur population d'origine. Ces processus de migration avec sélection positive vers l'Île-de-France et sélection négative vers le reste de la France sont cohérents avec les résultats établis plus haut dans la comparaison générale entre l'Île-de-France et le reste de la France. On trouve que les plus qualifiés et les moins qualifiés sont sur-représentés dans la migration vers l'Île-de-France, alors que c'est l'opposé dans la migration vers le reste du pays. Ainsi, les processus migratoires contribuent au tri spatial en Île-de-France par rapport au reste de la France. Toutefois, si on compare les résultats sur les effets individuels des migrants deux à deux, on peut voir que les migrants qui partent ne sont pas très différents de ceux qui arrivent. Cela limite le pouvoir des processus migratoires dans l'explication du tri spatial.

## Natifs et non-natifs de l'Île-de-France et du reste de la France

Pour compléter les résultats précédents, il est utile de prendre une perspective de plus long terme et de définir le statut de migrant par le lieu de naissance. Pour

éviter toute confusion sur les termes, nous parlons alors de « natifs » et de « non-natifs ». Pour l'Île-de-France, nous définissons comme natif tout salarié né en Île-de-France et comme non-natif tout salarié né dans le reste de la France, et réciproquement pour le reste de la France<sup>19</sup>.

Tableau 6. Log des salaires des natifs et des non-natifs d'Île-de-France et du reste de la France, 2007

	$\hat{A}$	$\hat{D}$	R <sup>2</sup>	Obs. 1	Obs. 2
$w_{RR}$ et $w_{R \rightarrow IdF}$	0,227 (0,002)*	1,485 (0,005)*	0,981	397 876	75 966
$w_{IdFIdF}$ et $w_{R \rightarrow IdF}$	0,023 (0,002)*	1,068 (0,005)*	0,818	64 068	75 966
$w_{IdFIdF}$ et $w_{IdF \rightarrow R}$	-0,142 (0,003)*	0,869 (0,006)*	0,907	64 068	28 194
$w_{RR}$ et $w_{IdF \rightarrow R}$	0,061 (0,003)*	1,208 (0,008)*	0,969	397 876	28 194

\* : pour  $\hat{A}$  significativement différent de 0 à 5 % ; pour  $\hat{D}$  significativement différent de 1 à 5 %.  
 $w_{ii}$  : distribution du logarithme des salaires des natifs dans le reste de la France ( $i = R$ ) ou en Île-de-France ( $i = IdF$ ).  
 $w_{i \rightarrow j}$  : distribution du logarithme des salaires des migrants venant d'une zone  $i$  ( $\in \{R, IdF\}$ ) vers une zone  $j$  ( $\in \{R, IdF\}$ ).  
 Obs. 1 et Obs. 2 : nombre d'observations dans le premier groupe (référence) et dans le second.

La première ligne du tableau 6 reporte les résultats de la comparaison de la distribution des salaires pour les salariés nés et travaillant dans le reste de la France avec la distribution des salaires des salariés nés hors de l'Île-de-France mais qui y travaillent. On trouve que cette dernière distribution est à la fois translatée vers la droite et dilatée par rapport à la première. De façon intéressante, les coefficients sont très proches de ceux de la deuxième ligne du tableau 2 qui compare plus généralement tous les salariés qui travaillent dans le reste de la France à ceux qui travaillent en Île-de-France. Cela suggère que les natifs et les non-natifs en Île-de-France sont très similaires en ce qui concerne leurs salaires. La deuxième ligne de résultats du tableau 6 apporte une confirmation directe de cette conjecture. Les salaires des non-natifs qui travaillent en Île-de-France ne sont que marginalement plus élevés que ceux des natifs et leur distribution plus dilatée.

La troisième ligne de résultats du tableau 6 compare les natifs de l'Île-de-France qui y travaillent à ceux qui travaillent dans le reste de la France. Les salaires de ces derniers sont plus faibles et plus comprimés. Toutefois, l'écart de salaire moyen entre les natifs de l'Île-de-France qui sont restés et ceux qui sont partis est plus faible que l'écart moyen de salaires entre tous les salariés de l'Île-de-France et tous les salariés du reste de la France. Ce résultat est cohérent avec les résultats de la quatrième ligne du tableau 6 qui compare les natifs du reste de la France qui y travaillent avec les non-natifs. Ces derniers ont un salaire en moyenne supérieur et leur distribution est dilatée par rapport à celle des natifs.

19. Il est difficile d'obtenir de l'information sur la localisation avant l'entrée sur le marché du travail. Nous utilisons le lieu de naissance qui, dans une grande majorité des cas, correspond au lieu de scolarisation et permet, plus généralement, de capturer les avantages liés à la localisation avant l'entrée sur le marché du travail.

Tableau 7. Log des effets individuels des natifs et des non-natifs d'Île-de-France et du reste de la France, 2007

	$\hat{A}$	$\hat{D}$	R <sup>2</sup>	Obs. 1	Obs. 2
$\delta_{RR}$ et $\delta_{R \rightarrow IdF}$	0,059 (0,001)*	1,547 (0,005)*	0,974	397 876	75 966
$\delta_{IdFIdF}$ et $\delta_{R \rightarrow IdF}$	-0,014 (0,002)*	1,166 (0,006)*	0,883	64 068	75 966
$\delta_{IdFIdF}$ et $\delta_{IdF \rightarrow R}$	-0,025 (0,003)*	0,913 (0,007)*	0,813	64 068	28 194
$\delta_{RR}$ et $\delta_{IdF \rightarrow R}$	0,048 (0,002)*	1,212 (0,008)*	0,934	397 876	28 194

\* : pour  $\hat{A}$  significativement différent de 0 à 5 % ; pour  $\hat{D}$  significativement différent de 1 à 5 %.

$\delta_{ij}$  : distribution du logarithme des effets individuels des natifs dans le reste de la France ( $i = R$ ) ou en Île-de-France ( $i = IdF$ ).

$\delta_{i \rightarrow j}$  : distribution du logarithme des effets individuels des natifs venant d'une zone  $i$  ( $i \in \{R, IdF\}$ ) vers une zone  $j$  ( $j \in \{R, IdF\}$ ).

Obs. 1 et Obs. 2 : nombre d'observations dans le premier groupe (référence) et dans le second.

La première ligne du tableau 7 reporte les résultats de la comparaison des effets individuels des natifs du reste de la France qui y travaillent encore avec ceux qui travaillent en Île-de-France. Les effets individuels de ces derniers sont plus élevés et leur distribution plus dilatée. Les résultats obtenus pour la comparaison plus générale de tous les salariés du reste de la France avec ceux de l'Île-de-France dans la deuxième ligne du tableau 3 sont quasiment identiques. Cela suggère à nouveau une grande similarité entre les natifs et les non-natifs en Île-de-France. Cette conjecture est confirmée par les résultats de la deuxième ligne du tableau 7 qui compare directement natifs et non-natifs en Île-de-France. On trouve que les effets individuels des non-natifs sont marginalement plus faibles et proviennent d'une distribution plus dilatée.

La troisième ligne de résultats du tableau 7 compare les effets individuels des natifs d'Île-de-France qui y travaillent encore avec les effets individuels de ceux qui travaillent maintenant dans le reste de la France. Les effets individuels de ces derniers sont un peu plus faibles et leur distribution un peu plus compressée. Enfin, la quatrième ligne de résultats compare les effets individuels des natifs et des non-natifs pour le reste de la France. On trouve que les non-natifs ont des effets individuels en moyenne plus élevés et qui proviennent d'une distribution plus dilatée relativement aux natifs du reste de la France qui y travaillent.

Nous tirons un certain nombre de conclusions de ces résultats. Tout d'abord, les non-natifs ont des salaires en moyenne plus élevés que les natifs. Toutefois, ces différences sont limitées. On trouve aussi plus de dispersion dans les salaires des non-natifs, ce qui reflète aussi la plus grande dispersion de leurs effets individuels. Surtout, en Île-de-France, ces différences d'effets individuels sont assez faibles, et les différences globales observées plus haut entre l'Île-de-France et le reste de la France ne peuvent pas être entièrement expliquées par les non-natifs. Il est aussi intéressant de remarquer qu'en terme d'effets individuels les comparaisons de « court terme » entre migrants et non-migrants donnent des résultats très proches de la comparaison de « long terme » entre natifs et non-natifs. Ainsi, les processus migratoires contribuent au tri spatial, mais ils sont sans doute loin d'en être l'unique origine.

## Jeunes et moins jeunes en Île-de-France et dans le reste de la France

Afin d'examiner une autre source potentielle de tri spatial, il est utile de répéter notre analyse pour différentes classes d'âges. Pour éviter de multiplier les résultats, nous considérons les salariés de moins de 35 ans et les salariés de plus de 45 ans.

Tableau 8. Log des salaires des migrants et des non-migrants de moins de 35 ans d'Île-de-France et du reste de la France, 2004-2007

	$\hat{A}$	$\hat{D}$	R <sup>2</sup>	Obs. 1	Obs. 2
$w_R$ et $w_{IdF}$	0,150 (0,002)*	1,529 (0,007)*	0,979	176 338	59 883
$w_{RR}$ et $w_{R \rightarrow IdF}$	0,166 (0,006)*	1,394 (0,018)*	0,957	91 110	3 533
$w_{IdFIdF}$ et $w_{R \rightarrow IdF}$	-0,049 (0,006)*	0,943 (0,013)*	0,799	26 790	3 533
$w_{IdFIdF}$ et $w_{IdF \rightarrow R}$	-0,096 (0,007)*	0,945 (0,013)*	0,959	26 790	3 262
$w_{RR}$ et $w_{IdF \rightarrow R}$	0,120 (0,007)*	1,399 (0,020)*	0,958	91 110	3 262

\* : pour  $\hat{A}$  significativement différent de 0 à 5 % ; pour  $\hat{D}$  significativement différent de 1 à 5 %.  
 $w_i$  : distribution du logarithme des salaires en Île-de-France ( $i = IdF$ ) ou dans le reste de la France ( $i = R$ ).  
 $w_{ij}$  : distribution du logarithme des salaires des non-migrants d'une zone  $i$  ( $i \in \{R, IdF\}$ ).  
 $w_{i \rightarrow j}$  : distribution du logarithme des salaires des migrants venant d'une zone  $i$  ( $i \in \{R, IdF\}$ ) vers une zone  $j$  ( $j \in \{R, IdF\}$ ).

Obs. 1 et Obs. 2 : nombre d'observations dans le premier groupe (référence) et dans le second.

Tableau 9. Log des salaires des migrants et des non-migrants de plus de 45 ans d'Île-de-France et du reste de la France, 2004-2007

	$\hat{A}$	$\hat{D}$	R <sup>2</sup>	Obs. 1	Obs. 2
$w_R$ et $w_{IdF}$	0,268 (0,003)*	1,350 (0,008)*	0,982	132 414	41 310
$w_{RR}$ et $w_{R \rightarrow IdF}$	0,331 (0,019)*	1,485 (0,034)*	0,980	104 415	1 066
$w_{IdFIdF}$ et $w_{R \rightarrow IdF}$	0,050 (0,020)*	1,112 (0,026)*	0,907	30 830	1 066
$w_{IdFIdF}$ et $w_{IdF \rightarrow R}$	-0,061 (0,019)*	1,098 (0,029)*	0,819	30 830	1 216
$w_{RR}$ et $w_{IdF \rightarrow R}$	0,220 (0,019)*	1,467 (0,039)*	0,969	104 415	1 216

\* : pour  $\hat{A}$  significativement différent de 0 à 5 % ; pour  $\hat{D}$  significativement différent de 1 à 5 %.  
 $w_i$  : distribution du logarithme des salaires en Île-de-France ( $i = IdF$ ) ou dans le reste de la France ( $i = R$ ).  
 $w_{ij}$  : distribution du logarithme des salaires des non-migrants d'une zone  $i$  ( $i \in \{R, IdF\}$ ).  
 $w_{i \rightarrow j}$  : distribution du logarithme des salaires des migrants venant d'une zone  $i$  ( $i \in \{R, IdF\}$ ) vers une zone  $j$  ( $j \in \{R, IdF\}$ ).

Obs. 1 et Obs. 2 : nombre d'observations dans le premier groupe (référence) et dans le second.



La première ligne du tableau 8 reporte les résultats de la comparaison de la distribution des salaires pour les salariés de moins de 35 ans en Île-de-France avec la distribution des salaires pour les salariés de moins de 35 ans dans le reste de la France. La première ligne du tableau 9 reporte les résultats de la même comparaison pour les salariés de plus de 45 ans. Les quatre lignes suivantes de ces deux tableaux reportent les résultats des comparaisons de distributions de salaires entre migrants et non-migrants pour les deux mêmes groupes d'âge.

Relativement aux résultats du tableau 2, on peut voir que les écarts de salaires entre l'Île-de-France et le reste de la France sont beaucoup plus faibles pour les moins de 35 ans et beaucoup plus forts pour les plus de 45 ans. En ce qui concerne la dispersion, les différences sont beaucoup moins marquées et vont dans le sens opposé avec plus de dispersion pour les salariés plus jeunes. Ainsi la prime de salaire des salariés de l'Île-de-France par rapport au reste de la France est presque deux fois plus forte à  $(\exp(0,268) - 1) = 31\%$  pour les salariés de plus de 45 ans que pour ceux de moins de 35 ans pour lesquels elle est de  $(\exp(0,150) - 1) = 16\%$ . Les écarts entre migrants et non-migrants sont aussi différents dans les deux classes d'âge. Les salariés migrants de plus de 45 ans ont des primes de salaires plus importantes (ou parfois moins négatives) que celles des salariés de moins de 35 ans. Ainsi, par exemple, les salariés de moins de 35 ans qui migrent vers l'Île-de-France reçoivent un salaire inférieur à celui des salariés déjà résidents en Île-de-France, alors que l'inverse est vrai pour les salariés de plus de 45 ans.

Tableau 10. Log des effets individuels des migrants et des non-migrants de moins de 35 ans d'Île-de-France et du reste de la France, 2004-2007

	$\hat{A}$	$\hat{D}$	R <sup>2</sup>	Obs. 1	Obs. 2
$\delta_R$ et $\delta_{IdF}$	-0,027 (0,002)*	1,426 (0,006)*	0,984	176 338	59 883
$\delta_{RR}$ et $\delta_{R \rightarrow IdF}$	0,017 (0,004)*	1,267 (0,022)*	0,995	91 110	3 533
$\delta_{IdFIdF}$ et $\delta_{R \rightarrow IdF}$	0,001 (0,004)	0,884 (0,015)*	0,796	26 790	3 533
$\delta_{IdFIdF}$ et $\delta_{IdF \rightarrow R}$	0,023 (0,006)*	0,953 (0,016)*	0,966	26 790	3 262
$\delta_{RR}$ et $\delta_{IdF \rightarrow R}$	0,039 (0,005)*	1,364 (0,022)*	0,952	91 110	3 262

\* : pour  $\hat{A}$  significativement différent de 0 à 5 % ; pour  $\hat{D}$  significativement différent de 1 à 5 %.

$\delta_i$  : distribution du logarithme des effets individuels en Île-de-France ( $i = IdF$ ) ou dans le reste de la France ( $i = R$ ).

$\delta_{ii}$  : distribution du logarithme des effets individuels des non-migrants d'une zone  $i$  ( $i \in \{R, IdF\}$ ).

$\delta_{i \rightarrow j}$  : distribution du logarithme des effets individuels des migrants venant d'une zone  $i$  ( $i \in \{R, IdF\}$ ) vers une zone  $j$  ( $j \in \{R, IdF\}$ ).

Obs. 1 et Obs. 2 : nombre d'observations dans le premier groupe (référence) et dans le second.

La première ligne du tableau 10 reporte les résultats de la comparaison de la distribution des effets individuels pour les salariés de moins de 35 ans en Île-de-France avec la distribution des effets individuels pour les salariés de moins de 35 ans dans le reste de la France. La première ligne du tableau 11 reporte les résultats de la même comparaison pour, cette fois, les salariés de plus de

Tableau 11. Log des effets individuels des migrants et non-migrants de plus de 45 ans d'Île-de-France et du reste de la France, 2004-2007

	$\hat{A}$	$\hat{D}$	R <sup>2</sup>	Obs. 1	Obs. 2
$\delta_R$ et $\delta_{IdF}$	0,152 (0,003)*	1,372 (0,007)*	0,961	132 414	41 310
$\delta_{RR}$ et $\delta_{R \rightarrow IdF}$	0,234 (0,017)*	1,522 (0,032)*	0,968	104 415	1 066
$\delta_{IdFIdF}$ et $\delta_{R \rightarrow IdF}$	0,070 (0,018)*	1,120 (0,024)*	0,866	30 830	1 066
$\delta_{IdFIdF}$ et $\delta_{IdF \rightarrow R}$	0,044 (0,016)*	1,045 (0,026)	0,865	30 830	1 216
$\delta_{RR}$ et $\delta_{IdF \rightarrow R}$	0,208 (0,016)*	1,420 (0,036)*	0,964	104 415	1 216

\* : pour  $\hat{A}$  significativement différent de 0 à 5 % ; pour  $\hat{D}$  significativement différent de 1 à 5 %.  
 $\delta_i$  : distribution du logarithme des effets individuels en Île-de-France ( $i = IdF$ ) ou dans le reste de la France ( $i = R$ ).  
 $\delta_{ii}$  : distribution du logarithme des effets individuels des non-migrants d'une zone  $i$  ( $i \in \{R, IdF\}$ ).  
 $\delta_{i \rightarrow j}$  : distribution du logarithme des effets individuels des migrants venant d'une zone  $i$  ( $i \in \{R, IdF\}$ ) vers une zone  $j$  ( $j \in \{R, IdF\}$ )  
 Obs. 1 et Obs. 2 : nombre d'observations dans le premier groupe (référence) et dans le second.

45 ans. Les quatre lignes suivantes de ces deux tableaux reportent les résultats des comparaisons de distributions d'effets individuels entre migrants et non-migrants pour les deux mêmes groupes d'âge.

Le premier résultat important est que le tri spatial est beaucoup plus fort chez les plus de 45 ans que chez les moins de 35 ans. Pour ces derniers, les effets individuels sont en moyenne plus élevés dans le reste de la France qu'en Île-de-France. En revanche, les effets individuels sont en moyenne beaucoup plus élevés en Île-de-France que dans le reste de la France pour les plus de 45 ans. Ils sont même entre deux et trois fois plus élevés que pour l'ensemble des salariés tel que rapporté dans le tableau 3. Par ailleurs, la distribution des effets individuels reste plus hétérogène en Île-de-France que dans le reste de la France et les différences entre les deux groupes d'âge sont faibles.

En ce qui concerne les migrants, on remarque aussi qu'ils ont des effets individuels en moyenne plus élevés dans les deux groupes d'âge. Cela contraste avec les résultats pour l'ensemble de la population du tableau 5. Dans une large mesure, les différences observées dans ce tableau pour tous les migrants vers l'Île-de-France ou venant de l'Île-de-France relativement aux non-migrants du reste de la France sont expliquées par les différences très marquées qu'on observe chez les plus de 45 ans pour ces deux groupes de migrants.

Si l'on compare les résultats des tableaux 8 et 9 pour les salaires et des tableaux 10 et 11 pour les effets individuels, on peut voir que la prime de salaire nette des effets individuels est plus importante pour les moins de 35 ans ( $0,150 - (-0,027) = 0,177$  ou 19 %) que pour les plus de 45 ans ( $0,268 - 0,152 = 0,116$  ou 12 %). Il est intéressant de constater que les migrants vers l'Île-de-France ont une prime de salaire nette des effets individuels d'environ 10 % pour les deux groupes d'âge, alors que la situation est plus contrastée pour les migrants vers le reste de la France. Les plus jeunes migrent avec une prime de salaire nette des effets individuels un peu inférieure à 8 %, alors que les plus anciens migrent depuis l'Île-de-France vers le reste de la France avec une prime

proche de zéro. En revanche, la plus grande hétérogénéité des effets individuels des salariés en Île-de-France explique l'essentiel de la plus grande hétérogénéité des salaires dans cette région.

Ces résultats sont compatibles avec une situation où l'apprentissage sur le marché du travail joue un rôle fondamental. En effet, l'apprentissage sur le marché du travail peut expliquer comment les différences d'effets individuels entre l'Île-de-France et le reste de la France sont faibles pour les salariés plus jeunes mais peuvent être beaucoup plus importantes pour les salariés plus vieux. Des effets importants d'apprentissage sur le marché du travail francilien peuvent aussi expliquer pourquoi la prime de salaire francilienne corrigée des effets individuels est plus faible pour les migrants que pour ceux qui travaillent déjà en Île-de-France. En revanche, cette explication cadre moins bien avec le fait que les salariés plus jeunes (à qui il reste à apprendre) reçoivent une prime de salaire nette des effets individuels plus importante que les salariés plus vieux (qui ont déjà appris)<sup>20</sup>.

## DISPARITÉS DE SALAIRES ET TRI SPATIAL À L'INTÉRIEUR DE L'ÎLE-DE-FRANCE

Nous nous tournons maintenant vers les comparaisons à l'intérieur de l'Île-de-France. La section précédente montre l'existence d'une forte hétérogénéité des salaires et des qualifications entre l'Île-de-France et le reste de la France. Nous cherchons maintenant à savoir si l'Île-de-France constitue un bloc uniforme ou si le tri spatial avec le reste de la France est expliqué par une partie de l'Île-de-France seulement.

### Paris plus petite couronne et grande couronne

Le tableau 12 reporte les résultats de la comparaison des distributions de salaires de Paris et sa petite couronne avec la grande couronne. La première ligne de résultats montre que les salaires sont plus élevés et que leur distribution est plus dilatée à Paris et en petite couronne qu'en grande couronne. L'écart de salaire entre Paris plus petite couronne et la grande couronne est de  $(\exp(0,084) - 1) = 9\%$ . Cet écart est légèrement supérieur à celui qui sépare, en dehors de l'Île-de-France, les trois départements des Bouches-du-Rhône, du Nord et du Rhône du reste du pays, mais il reste bien plus faible que celui entre l'Île-de-France et le reste du pays.

Les lignes 2 à 5 du tableau 12 reportent les résultats des comparaisons entre migrants et non-migrants pour Paris plus la petite couronne et la grande couronne. Pour toutes les comparaisons, les salaires des migrants sont inférieurs à ceux des non-migrants. Les valeurs estimées pour le coefficient D indiquent soit une modeste compression, soit une modeste dilatation.

---

20. Cela peut être néanmoins rationalisé si le surcoût de vivre en Île-de-France affecte plus les salariés les plus jeunes qui sont susceptibles de consacrer une part plus importante de leur salaire au logement. Le différentiel de salaire net des effets individuels et du coût de la vie pourrait être alors négatif pour les plus jeunes qui payeraient ainsi indirectement leur apprentissage et positifs pour les plus vieux qui ont déjà appris.

Tableau 12. Log des salaires des migrants et des non-migrants de Paris plus la petite couronne et de la grande couronne d'Île-de-France, 2004-2007

	$\hat{A}$	$\hat{D}$	R <sup>2</sup>	Obs. 1	Obs. 2
$w_{GC}$ et $w_{PPC}$	0,084 (0,003)*	1,122 (0,006)*	0,958	47 088	92 946
$w_{GGC}$ et $w_{GC \rightarrow PPC}$	-0,016 (0,007)*	1,050 (0,014)*	0,841	25 681	4 048
$w_{PPCPPC}$ et $w_{GC \rightarrow PPC}$	-0,118 (0,007)*	0,933 (0,012)*	0,981	50 376	4 048
$w_{PPCPPC}$ et $w_{PPC \rightarrow GC}$	-0,160 (0,008)*	0,910 (0,015)*	0,985	50 376	4 174
$w_{GGC}$ et $w_{PPC \rightarrow GC}$	-0,058 (0,008)*	1,024 (0,016)	0,956	25 681	4 174

\* : pour  $\hat{A}$  significativement différent de 0 à 5 % ; pour  $\hat{D}$  significativement différent de 1 à 5 %.  
 $w_i$  : distribution du logarithme des salaires dans la grande couronne de l'Île-de-France ( $i = GC$ ) ou dans Paris et sa petite couronne ( $i = PPC$ ).  
 $w_{ii}$  : distribution du logarithme des salaires des non-migrants d'une zone  $i$  ( $i \in \{GC, PPC\}$ ).  
 $w_{i \rightarrow j}$  : distribution du logarithme des salaires des migrants venant d'une zone  $i$  ( $i \in \{GC, PPC\}$ ) vers une zone  $j$  ( $j \in \{GC, PPC\}$ ).  
 Obs. 1 et Obs. 2 : nombre d'observations dans le premier groupe (référence) et dans le second.

Tableau 13. Log des effets individuels des migrants et des non-migrants de Paris plus la petite couronne et de la grande couronne d'Île-de-France, 2004-2007

	$\hat{A}$	$\hat{D}$	R <sup>2</sup>	Obs. 1	Obs. 2
$\delta_{GC}$ et $\delta_{PPC}$	0,037 (0,003)*	1,142 (0,007)*	0,954	47 088	92 946
$\delta_{GGC}$ et $\delta_{GC \rightarrow PPC}$	-0,027 (0,006)*	0,991 (0,016)	0,975	25 681	4 048
$\delta_{PPCPPC}$ et $\delta_{GC \rightarrow PPC}$	-0,079 (0,006)*	0,873 (0,013)*	0,958	50 376	4 048
$\delta_{PPCPPC}$ et $\delta_{PPC \rightarrow GC}$	-0,097 (0,006)*	0,853 (0,014)*	0,983	50 376	4 174
$\delta_{GGC}$ et $\delta_{PPC \rightarrow GC}$	-0,045 (0,006)*	0,969 (0,016)	0,967	25 681	4 174

\* : pour  $\hat{A}$  significativement différent de 0 à 5 % ; pour  $\hat{D}$  significativement différent de 1 à 5 %.  
 $\delta_i$  : distribution du logarithme des effets individuels dans la grande couronne de l'Île-de-France ( $i = GC$ ) ou de Paris et sa petite couronne ( $i = PPC$ ).  
 $\delta_{ii}$  : distribution du logarithme des effets individuels des non-migrants d'une zone  $i$  ( $i \in \{GC, PPC\}$ ).  
 $\delta_{i \rightarrow j}$  : distribution du logarithme des effets individuels des migrants venant d'une zone  $i$  ( $i \in \{GC, PPC\}$ ) vers une zone  $j$  ( $j \in \{GC, PPC\}$ ).  
 Obs. 1 et Obs. 2 : nombre d'observations dans le premier groupe (référence) et dans le second.

Le tableau 13 réplique l'analyse du tableau 12 pour les effets fixes individuels. La première ligne de résultats, qui compare la distribution des effets individuels pour tous les salariés de Paris et la petite couronne à la distribution correspondante pour la grande couronne, indique que les effets individuels sont en moyenne plus élevés à Paris et en petite couronne qu'en grande couronne

d'environ  $(\exp(0,037) - 1) = 4\%$ . Cet écart est plus faible que celui séparant l'Île-de-France du reste de la France mais reste néanmoins économiquement significatif. La distribution des effets individuels est aussi plus dilatée à Paris et en petite couronne qu'en grande couronne.

Les comparaisons entre migrants et non-migrants indiquent que les migrants ont des effets individuels en moyenne plus faibles que les non-migrants de leur zone d'origine. Les effets individuels des migrants sont aussi plus faibles en moyenne que ceux des non-migrants de leur zone d'arrivée. Toutefois, malgré cette sélection négative des migrants dans les deux zones, la migration contribue à nouveau au tri spatial. À Paris et en petite couronne, les migrants qui partent vers la grande couronne ont des effets individuels plus faibles que les migrants qui arrivent. On observe le contraire en grande couronne. Toutefois, les faibles écarts qu'on observe entre partants et arrivants suggèrent, encore une fois, une faible contribution de la migration au tri spatial.

Même si nous ne reportons pas les résultats ici pour éviter de multiplier les tableaux de chiffres, nous avons aussi comparé les distributions de salaires et d'effets individuels de Paris et sa petite couronne avec celles de la grande couronne pour les moins de 35 ans et les plus de 45 ans comme dans les tableaux 8-9 et 10-11. Bien que les écarts de salaires soient les mêmes pour ces deux groupes d'âge, nous observons néanmoins que les effets individuels sont relativement plus élevés pour les salariés de plus de 45 ans. Cela suggère que le même type d'effet d'apprentissage mis en évidence plus haut dans notre comparaison entre l'Île-de-France et le reste de la France est aussi à l'œuvre à l'intérieur de l'Île-de-France.

Nous tirons un certain nombre de conclusions de ces résultats. Tout d'abord, quand on compare tous les salariés, on retrouve à l'intérieur de l'Île-de-France le même type de différences qu'entre l'Île-de-France et le reste de la France. La spécificité de l'Île-de-France soulignée plus haut provient principalement de sa partie centrale, Paris et la petite couronne. En termes de salaires, la grande couronne est toutefois plus proche de Paris et de la petite couronne que du reste de la France. La ressemblance entre la comparaison Île-de-France et reste de la France et celle au sein de l'Île-de-France peut être poussée plus loin puisque, dans les deux cas, les mouvements migratoires semblent avoir un effet comparable sur le tri spatial. Par ailleurs, des effets d'apprentissage semblent aussi expliquer les différences de salaires et d'effets individuels dans les deux cas.

## Paris et petite couronne

Pour affiner notre analyse de l'intérieur de l'Île-de-France, nous comparons maintenant la ville de Paris et la petite couronne. Les résultats de cette comparaison sont donnés dans le tableau 14. La première ligne montre que les salaires sont plus élevés dans la petite couronne (qui contient bien évidemment les départements des Haut-de-Seine et du Val-de-Marne) que dans Paris intra-muros. La valeur estimée pour le coefficient A indique que les salaires sont environ 6 % plus élevés en petite couronne que dans Paris. Bien que non négligeable, cet écart de salaire est toutefois un peu inférieur à celui qui sépare le centre de l'Île-de-France à la grande couronne et beaucoup plus faible que celui qu'on observe entre l'Île-de-France et le reste de la France. La distribution des salaires est aussi marginalement plus compressée à Paris qu'en petite couronne.

Tableau 14. Log des salaires des migrants et des non-migrants de Paris et de la petite couronne, 2004-2007

	$\hat{A}$	$\hat{D}$	R <sup>2</sup>	Obs. 1	Obs. 2
$w_{PC}$ et $w_{Paris}$	-0,056 (0,003)*	0,962 (0,005)*	0,943	52 651	40 295
$w_{PCPC}$ et $w_{PC \rightarrow Paris}$	-0,110 (0,010)*	0,971 (0,014)*	0,988	26 007	2 992
$w_{ParisParis}$ et $w_{PC \rightarrow Paris}$	-0,051 (0,009)*	0,973 (0,016)	0,942	17 518	2 992
$w_{ParisParis}$ et $w_{Paris \rightarrow PC}$	-0,026 (0,009)*	1,016 (0,016)	0,926	17 518	3 859
$w_{PCPC}$ et $w_{Paris \rightarrow PC}$	-0,085 (0,008)*	1,014 (0,015)	0,973	26 007	3 859

\* : pour  $\hat{A}$  significativement différent de 0 à 5 % ; pour  $\hat{D}$  significativement différent de 1 à 5 %.  
 $w_i$  : distribution du logarithme des salaires dans la petite couronne de l'Île-de-France ( $i = PC$ ) ou à Paris ( $i = Paris$ ).  
 $w_{ij}$  : distribution du logarithme des salaires des non-migrants d'une zone  $i$  ( $\in \{PC, Paris\}$ ).  
 $w_{i \rightarrow j}$  : distribution du logarithme des salaires des migrants venant d'une zone  $i$  ( $\in \{PC, Paris\}$ ) vers une zone  $j$  ( $\in \{PC, Paris\}$ ).  
 Obs. 1 et Obs. 2 : nombre d'observations dans le premier groupe (référence) et dans le second.

L'analyse des migrants dans les lignes 2 à 5 du tableau 14 montre que les migrants ont en moyenne un salaire inférieur aux non-migrants, que ce soit ceux de leur zone d'origine ou ceux de leur zone de destination. Les valeurs estimées pour le coefficient D indiquent aussi soit une modeste compression, soit une modeste dilatation.

Tableau 15. Log des effets individuels des migrants et des non-migrants de Paris et de la petite couronne, 2004-2007

	$\hat{A}$	$\hat{D}$	R <sup>2</sup>	Obs. 1	Obs. 2
$\delta_{PC}$ et $\delta_{Paris}$	-0,047 (0,003)*	0,967 (0,005)*	0,983	52 651	40 295
$\delta_{PCPC}$ et $\delta_{PC \rightarrow Paris}$	-0,095 (0,010)*	0,923 (0,014)*	0,988	26 007	2 992
$\delta_{ParisParis}$ et $\delta_{PC \rightarrow Paris}$	-0,039 (0,009)*	0,933 (0,016)*	0,968	17 518	2 992
$\delta_{ParisParis}$ et $\delta_{Paris \rightarrow PC}$	-0,018 (0,009)*	0,975 (0,016)	0,912	17 518	3 859
$\delta_{PCPC}$ et $\delta_{Paris \rightarrow PC}$	-0,074 (0,008)*	0,965 (0,015)*	0,982	26 007	3 859

\* : pour  $\hat{A}$  significativement différent de 0 à 5 % ; pour  $\hat{D}$  significativement différent de 1 à 5 %.  
 $\delta_i$  : distribution du logarithme des salaires dans la petite couronne de l'Île-de-France ( $i = PC$ ) ou à Paris ( $i = Paris$ ).  
 $\delta_{ij}$  : distribution du logarithme des salaires des non-migrants d'une zone  $i$  ( $\in \{PC, Paris\}$ ).  
 $\delta_{i \rightarrow j}$  : distribution du logarithme des salaires des migrants venant d'une zone  $i$  ( $\in \{PC, Paris\}$ ) vers une zone  $j$  ( $\in \{PC, Paris\}$ ).  
 Obs. 1 et Obs. 2 : nombre d'observations dans le premier groupe (référence) et dans le second.

Comme pour les comparaisons précédentes, le tableau 15 réplique les estimations du tableau 14 pour les effets individuels. La première ligne de résultats, qui compare la distribution des effets individuels pour tous les salariés de Paris à la distribution correspondante pour la petite couronne, indique que les effets individuels sont en moyenne plus élevés en petite couronne qu'à Paris d'environ  $(\exp(0,047) - 1 =) 5 \%$ . Cet écart est assez important puisqu'il concerne deux aires géographiques voisines et de petite taille. Par ailleurs, la distribution des effets fixes est marginalement plus compressée à Paris qu'en petite couronne.

Comme précédemment lorsque nous avons contrasté Paris plus la petite couronne et la grande couronne, les comparaisons entre migrants et non-migrants indiquent que les migrants ont des effets individuels en moyenne plus faibles que les non-migrants de leur zone d'origine. Les effets individuels des migrants sont aussi plus faibles en moyenne que ceux des non-migrants de leur zone d'arrivée. L'analogie avec le cas précédent peut être poussée plus loin encore puisqu'on observe à nouveau un cas de sélection négative des migrants pour Paris allant de pair avec une situation où les migrants contribuent au tri spatial. Les migrants qui quittent Paris vers la petite couronne ont des effets individuels plus élevés que les migrants qui arrivent. On observe le contraire en petite couronne. À nouveau, les écarts entre partants et arrivants sont faibles.

La comparaison des distributions de salaires entre Paris et sa petite couronne pour les moins de 35 ans et les plus de 45 ans révèle à nouveau des effets d'apprentissage intéressants puisque l'écart d'effet individuel entre Paris et sa petite couronne pour les moins de 35 ans n'est en moyenne que de  $(\exp(0,015) - 1 =) - 1,5 \%$ , alors qu'il est de  $(\exp(0,070) - 1 =) - 6,8 \%$  pour les plus de 45 ans.

Nous tirons un certain nombre de conclusions de ces résultats. Tout d'abord, même à un niveau spatial relativement fin, l'Île-de-France ne forme pas un bloc homogène. On distingue trois ensembles : la ville de Paris, la petite couronne, et la grande couronne. C'est en petite couronne que les salaires sont les plus élevés et en grande couronne qu'ils sont les plus bas. Dans une large mesure, ce classement reflète la distribution des effets individuels entre ces zones. Après avoir pris en compte les effets individuels, les salaires sont  $(\exp(0,084 - 0,037) - 1 =) 5 \%$  plus élevés à Paris et en petite couronne qu'en grande couronne. Quand on distingue Paris de la petite couronne, on constate que les salaires sont  $(\exp(0,056 - 0,047) - 1 =) 1 \%$  plus élevés en petite couronne qu'à Paris après avoir pris en compte les différences d'effets individuels. La petite couronne est donc non seulement l'épicentre du tri spatial en Île-de-France, mais c'est aussi le lieu où les salaires sont les plus élevés même après avoir pris en compte les effets individuels. Enfin, on trouve que les écarts entre les distributions des salaires et des effets individuels entre zones au sein de l'Île-de-France sont principalement le fait de translation. Les coefficients estimés pour la dilatation sont en général proches de 1, contrairement à ce qui se produit lorsqu'on compare l'Île-de-France au reste de la France. Comme pour les comparaisons précédentes, on trouve aussi des différences marquées entre les salariés plus âgés et les salariés moins âgés qui soulignent l'importance des effets d'apprentissage.

## CONCLUSIONS

Notre analyse des différences de salaires et de qualifications des salariés entre la région Île-de-France et le reste de la France et au sein de la région francilienne quantifie un certain nombre des faits stylisés déjà connus et met en évidence d'autres faits nouveaux.

Les salaires en Île-de-France sont supérieurs à ceux du reste de la France d'environ un quart. Un quart de cette différence provient de qualifications plus élevées de la force de travail francilienne. Les trois quarts restants sont, en revanche, causés par des salaires plus élevés en Île-de-France à qualifications égales. Cela suggère une plus grande productivité francilienne causée soit par des économies d'agglomérations, soit par des avantages productifs d'une autre nature tels que de meilleures infrastructures par exemple.

Cette plus grande richesse et ces meilleures qualifications ne sont toutefois que des moyennes. Les inégalités salariales sont beaucoup plus fortes en Île-de-France et reflètent pour l'essentiel la surreprésentation en Île-de-France des salariés les plus qualifiés ainsi que celle des moins qualifiés. La région la plus riche de France est aussi une région où les salariés les moins qualifiés avec les salaires les plus faibles sont présents de façon disproportionnée. Toute politique territoriale qui a un objectif d'équité ne doit pas ignorer ce fait.

Dans le même temps, il apparaît aussi que la surreprésentation des plus qualifiés ainsi que celle des moins qualifiés en Île-de-France n'est que pour partie causée par les flux migratoires vers l'Île-de-France ou depuis l'Île-de-France ou par le fait d'être né en Île-de-France. Au contraire, le tri spatial semble fortement lié à l'âge et au fait d'avoir acquis une expérience professionnelle en Île-de-France. Bien que ces résultats demandent à être approfondis et confirmés par d'autres approches, ils suggèrent qu'il existe des bénéfices importants à pouvoir résider et travailler en Île-de-France. Faciliter l'accès au marché du travail francilien semble donc fondamental. Cela implique de pouvoir accompagner la croissance francilienne par une plus grande offre de logements et une amélioration des transports et de la qualité de vie dans cette région. La réduction des autres barrières à la mobilité sont aussi importantes pour donner aux jeunes salariés qui le souhaitent l'opportunité de venir améliorer leurs qualifications sur le marché du travail francilien.

Enfin, la région francilienne est loin de constituer un bloc homogène en terme de marché du travail. Pour les distributions de salaires et de qualifications, ses franges extérieures sont les moins bien loties, à mi-chemin du reste de l'Île-de-France et du reste de la France. À l'opposé, la petite couronne autour de Paris est la partie de l'Île-de-France qui offre les salaires les plus élevés. Comme pour l'Île-de-France par rapport au reste de la France, cet avantage de la petite couronne par rapport au reste de l'Île-de-France reflète à la fois un avantage productif et des qualifications plus élevées. Pour autant que ces qualifications soient acquises par une expérience sur le marché du travail local (comme pour l'Île-de-France par rapport au reste de la France), pouvoir accéder plus spécifiquement au marché du travail de la petite couronne est important pour tous les Franciliens. Pour cela, et là encore, une plus grande offre de logements dans les abords immédiats de Paris et une amélioration des transports pour pouvoir accéder à la petite couronne depuis le reste de l'Île-de-France sont sans doute fondamentales.



RÉFÉRENCES BIBLIOGRAPHIQUES

- BACOLOD M., BLUM B.S. et STRANGE W.C. [2009a], « Skills in the city », *Journal of Urban Economics*, 65 (2), p. 136-153.
- BACOLOD M., BLUM B.S. et STRANGE W.C. [2009b], « Urban interactions: soft skills versus specialization », *Journal of Economic Geography*, 9 (2), p. 227-262.
- BACOLOD M., BLUM B.S. et STRANGE W.C. [2010], « Elements of skill: Traits, intelligences, education, and agglomeration », *Journal of Regional Science*, 50 (1), p. 245-280.
- BAUM-SNOW N. et PAVAN R. [2012], « Understanding the city size wage gap », *Review of Economic Studies*, 79 (1), p. 88-127.
- BAUM-SNOW N. et PAVAN R. [2014], « Inequality and city size », *Review of Economics and Statistics* (à paraître).
- BEHRENS K., DURANTON G. et ROBERT-NICOUD F. [2014], « Productive cities: Sorting, selection, and agglomeration », *Journal of Political Economy* (à paraître).
- BEHRENS K. et ROBERT-NICOUD F. [2014], « Survival of the fittest in cities: Agglomeration, polarization, and income inequality », *The Economic Journal* (à paraître).
- BERRY C.R. et GLAESER E.L. [2005], « The divergence of human capital levels across cities », *Papers in Regional Science*, 84 (3), p. 407-444.
- CARD D., HEINING J. et KLINE P. [2013], « Workplace heterogeneity and the rise of West German wage inequality », *Quarterly Journal of Economics*, 128 (3), p. 967-1015.
- CICCONE A. et PERI G. [2006], « Identifying human capital externalities: Theory with an applications », *Review of Economic Studies*, 73 (2), p. 381-412.
- COMBES P.-P., DURANTON G. et GOBILLON L. [2008], « Spatial wage disparities: Sorting matters! », *Journal of Urban Economics*, 63 (2), p. 723-742.
- COMBES P.-P., DURANTON G. et GOBILLON L. [2011], « The identification of agglomeration economies », *Journal of Economic Geography*, 11 (2), p. 253-266.
- COMBES P.-P., DURANTON G., GOBILLON L., PUGA D. et ROUX S. [2012a], « The productivity advantages of large cities: Distinguishing agglomeration from firm selection », *Econometrica*, 80 (6), p. 2543-2594.
- COMBES P.-P., DURANTON G., GOBILLON L. et ROUX S. [2010], « Estimating agglomeration economies with history, geology, and worker effects », dans GLAESER E.L. (eds), *The Economics of Agglomeration*, Cambridge (Mass.), National Bureau of Economic Research, p. 15-65.
- COMBES P.-P., DURANTON G., GOBILLON L. et ROUX S. [2012b], « Sorting and local wage and skill distributions in France », *Regional Science and Urban Economics*, 42 (6), p. 913-930.
- DAVIS D.R. [2009], *A spatial knowledge economy*, Miméo, Columbia University.
- DE LA ROCA J. et PUGA D. [2012], « The dynamic earnings premium of dense cities », *Miméo*, CEMFI et IMDEA Social Sciences.
- DIAMOND R. [2013], « The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000 », *Miméo*, Harvard University.
- DURANTON G., MARTIN P., MAYER T. et MAYNERIS F. [2008], *Les Pôles de compétitivité : que peut-on en attendre ?*, CEPREMAP, Paris.
- DURANTON G. et PUGA D. [2004], « Micro-foundations of urban agglomeration economies » dans HENDERSON J.V. et THISSE J.-F. (eds), *Handbook of Regional and Urban Economics*, vol. 4, Amsterdam, North-Holland, p. 2063-2117.
- ECKHOUT J., PINHEIRO R. et SCHMIDHEINY K. [2014], « Spatial sorting », *Journal of Political Economy*, 122 (3), p. 554-620.
- ETZO I. [2008], « Internal migration: A review of the literature », Processed, University of Cagliari.
- FORTIN N., LEMIEUX T. et FIRPO S. [2011], « Decomposition methods in economics », dans ASHENFELTER O. et CARD D. (eds), *Handbook of Labor Economics*, volume 4A, Amsterdam, North-Holland, p. 1-102.

- GLAESER E.L., KAHN M.E. et RAPPAPORT J. [2008], « Why do the poor live in cities? The role of public transportation », *Journal of Urban Economics*, 63 (1), p. 1-24.
- GLAESER E.L. et MARÉ D. [2001], « Cities and skills », *Journal of Labor Economics*, 19 (2), p. 316-342.
- GLAESER E.L., RESSEGER M. et TOBIO K. [2009], « Inequality in cities », *Journal of Regional Science*, 49 (4), p. 617-646.
- GLAESER E.L. et RESSEGER M.R. [2010], « The complementarity between cities and skills », *Journal of Regional Science*, 50 (1), p. 221-244.
- GRAVIER J.-F. [1958], *Paris et le désert français*, Paris, Flammarion.
- GRAVIER J.-F. et DAUTRY R. [1947], *Paris et le désert français : décentralisation, équipement, population*, Paris, Le Portulan.
- GREENWOOD M.J. [1997], « Internal migrations in developed countries », dans ROSEN-ZWEIG M.R. et STARK O. (eds), *Handbook of Population and Family Economics*, volume 1B, Amsterdam, North-Holland, p. 647-720.
- MATANO A. et NATICCHIONI P. [2012], « Wage distribution and the spatial sorting of workers », *Journal of Economic Geography*, 12 (2), p. 379-408.
- MION G. et NATICCHIONI P. [2009], « The spatial sorting and matching of skills and firms », *Canadian Journal of Economics*, 42 (1), p. 683-701.
- MORETTI E. [2004], « Human capital externalities in cities », dans HENDERSON V. et THISSE J.-F. (dir.), *Handbook of Regional and Urban Economics*, vol. 4, Amsterdam, North-Holland, p. 2243-2291.
- NOCKE V. [2006], « A gap for me: Entrepreneurs and entry », *Journal of the European Economic Association*, 4 (5), p. 929-956.
- VENABLES A.J. [2011], « Productivity in cities: Self-selection and sorting », *Journal of Economic Geography*, 11 (2), p. 241-251.
- WHEELER C.H. [2001], « Search, sorting, and urban agglomeration », *Journal of Labor Economics*, 19 (4), p. 879-899.