

Complément C

Origine et ampleur des inégalités spatiales de salaire en France^(*)

Pierre-Philippe Combes

CERAS-ENPC, Boston University et CNRS (URA 2036)

Gilles Duranton

Department of Geography and Environment (London School of Economics)

Laurent Gobillon

CREST-INSEE et CNRS (URA 2200)

Introduction

Les différences de salaires entre régions, ou entre ville et campagne, ou encore entre Paris et Province, sont souvent perçues comme inéquitables et injustes. Elles ne seraient dues qu'à des avantages exogènes de certains lieux, comme de meilleures dotations en infrastructures publiques, des ressources naturelles plus abondantes ou même simplement une localisation physique plus avantageuse. Ainsi, au nom d'un principe d'égalité, il serait légitime d'effectuer des transferts des grandes villes, et notamment de Paris, vers les villes plus petites ou les zones rurales, afin de compenser ces différentiels de salaires non liés à de véritables différences de compétence. Le but de ce travail est tout d'abord de clarifier le débat théorique relatif aux inégalités salariales spatiales grâce aux concepts de l'économie de la croissance et de la production, de l'économie du travail et de l'économie géographique et urbaine. Ensuite, nous constatons qu'effectivement, les disparités salariales sont importantes entre les 341 zones d'emploi françaises. Ces disparités sont également persistantes dans le temps et fortement corrélées avec la « densité du tissu économique ». Enfin, une analyse économétrique plus fine montre comment l'on peut distinguer et pondérer

(*) Nous remercions Thierry Magnac, Henry Overman, Diego Puga et Jean-Marc Robin pour leurs conseils avisés, ainsi que Francis Kramarz et Sébastien Roux pour leur aide efficace quant à l'utilisation des données.

les différentes explications de ces différences de productivité locale. Cela permet de reprendre sur des bases clarifiées le débat relatif, d'une part, à d'éventuelles iniquités dues à l'espace et, d'autre part, aux instruments susceptibles ou non d'affecter les différentiels locaux de salaires.

1. Les explications économiques des différences de productivité locale

On peut regrouper les explications des différences spatiales de salaires et de productivité du travail en trois grandes familles.

La première famille d'explications repose sur les différences de dotations « physiques » entre villes ou régions. Ces différences de dotations physiques impliqueraient un différentiel de productivité qui lui-même se traduirait en salaires plus élevés dans les zones favorisées. Que faut-il entendre par dotations physiques ? Il s'agit tout d'abord des différences de dotations en ressources naturelles. De plus, être localisé le long d'une côte, à proximité d'une rivière ou à l'inverse d'être enclavé dans une région montagneuse peut influencer sur les coûts de production. Bénéficier ou non d'un bon climat peut aussi agir directement sur la productivité des travailleurs. Ces différences de dotations peuvent naturellement être également relatives au capital public ou privé. Finalement, les différences en dotations, telles que nous les définissons, peuvent aussi inclure des différences institutionnelles ou de technologie entre régions. Ce type d'explications est souvent favorisé par les géographes, mais aussi les économistes de la croissance, dont la littérature à ce sujet est volumineuse (voir par exemple Durlauf et Quah, 1999 ou Temple, 1999 pour des *surveys* récents).

En France, les différences institutionnelles sont faibles. On peut aussi penser que le capital privé est *a priori* relativement mobile. Par conséquent, il est peu probable que les différences de dotations en capital par tête soient à la source des inégalités territoriales en France. La littérature insiste aussi sur l'existence de différences technologiques exogènes et sur le rôle de la géographie physique. Ce type de constat est cependant établi empiriquement pour le moment surtout au niveau des pays (Hall et Jones, 1999 et Landes, 1999) et semble plus difficile à justifier au sein du territoire français. En revanche, les infrastructures publiques sont à la fois importantes et très inégalement réparties sur le territoire français si l'on pense par exemple au train à grande vitesse, aux aéroports internationaux ou encore au réseau autoroutier.

En termes de politique économique, il est difficile de jouer sur la géographie physique ou les dotations naturelles, mais, en revanche, les décideurs maîtrisent les dotations en capital public local et peuvent jouer indirectement sur les différences technologiques, par le biais des politiques de recherche et d'éducation. Par ailleurs, il est intéressant de noter qu'avec ce premier

type d'explications, l'espace est bien directement responsable d'une certaine iniquité. Les travailleurs localisés dans les zones les mieux dotées bénéficient, uniquement de par leur localisation, d'une meilleure rémunération.

La deuxième famille d'explications relie directement les différences de composition de la main d'œuvre locale aux différences spatiales de productivité et de salaire. En d'autres termes, les localités dans lesquelles le travail est plus qualifié ont un salaire moyen plus élevé de ce fait. En effet, les différents secteurs d'activité ne nécessitent pas *a priori* les mêmes parts de chaque qualification et, d'autre part, ces différents secteurs ne sont pas distribués de la même façon dans l'espace. Ainsi, il est clair que si dans une zone donnée les secteurs utilisant une main d'œuvre plus qualifiée ou plus expérimentée sont surreprésentés, le salaire moyen local y est plus élevé. De plus, même pour un secteur donné, si dans une zone d'emploi les entreprises de ce secteur emploient de la main d'œuvre plus qualifiée, le salaire y est aussi plus élevé. Ce type d'explication est souvent préféré par les économistes du travail. Notons que les chercheurs travaillant sur les inégalités régionales ont quant à eux plutôt délaissé ce type d'explication, aux exceptions notables et récentes de Glaeser et Maré (2001) pour les États-Unis et de Duranton et Monastiriotis (2002) pour le Royaume-Uni. Ces deux études confirment que les différences de qualification jouent un rôle important dans la détermination des inégalités spatiales de salaires.

Avec cette deuxième famille d'explications, les problèmes d'iniquité relatifs aux différences salariales se posent de façon complètement différente. L'espace n'est plus directement responsable des inégalités spatiales. Ces dernières ne font que refléter des différences de rémunérations entre qualifications. Par conséquent, c'est au niveau du marché du travail qu'il faudrait se demander si ce phénomène est légitime ou non. On se situe donc en amont de l'analyse spatiale. De plus, jouer sur les dotations locales en facteurs, en améliorant par exemple les infrastructures publiques locales, n'a alors que peu de chances d'affecter les inégalités salariales spatiales, puisque la cause en est autre. Enfin, notons que si les structures sectorielles s'homogénéisaient entre zones d'emploi et que chaque secteur employait les mêmes proportions de chaque qualification quelle que soit sa localisation, les salaires moyens des zones d'emploi s'égaliseraient. Cependant, les inégalités interindividuelles dues aux différences de qualification et d'expérience, elles, persisteraient.

Enfin, la troisième famille d'explications, plus proprement spécifique à l'économie urbaine ou géographique, repose sur l'existence d'interactions entre travailleurs et entreprises localisés dans un même lieu. Ces interactions engendreraient des gains de productivité conduisant à des salaires plus élevés. Là encore, les interactions évoquées peuvent être de nature diverse. L'explication la plus ancienne remonte à la notion de division du travail proposée par Adam Smith. Étant donné que les entreprises ne peuvent embaucher que des personnes résidant dans un certain périmètre autour de leur localisation, le marché du travail est dans une large mesure « local ». Plus la

densité de population est forte, plus la division du travail, c'est-à-dire la spécialisation des tâches, peut être poussée et engendrer ainsi des gains de productivité.

Dans une optique voisine, un argument plus récent, proposé par Helsley et Strange (1990), repose sur le fait que l'appariement (*matching*) entre travailleurs et entreprises s'améliore avec la taille du marché local du travail. Plus le nombre de travailleurs est élevé, plus la probabilité de trouver un employé qui corresponde bien aux besoins d'une entreprise donnée augmente. Inversement, plus le nombre d'entreprises est élevé, plus il est facile pour les travailleurs de trouver un emploi pour lequel ils sont idéalement qualifiés.

Les liens entre producteurs de biens finaux et de biens intermédiaires sont également susceptibles d'être à l'origine d'externalités locales améliorant la productivité. Comme souligné dans les modèles de l'économie urbaine (Abdel-Rahman, 1988 et Fujita, 1988) ou d'économie géographique (Krugman et Venables, 1995) récents, la présence d'un plus grand nombre de fournisseurs locaux d'*inputs* intermédiaires améliore l'efficacité productive du fait de la coexistence de rendements croissants internes aux entreprises et de coûts de transport. Le phénomène est cumulatif, puisqu'en retour ces fournisseurs ont des incitations à se localiser près de leurs clients.

Enfin, des externalités locales « pures » (ou « *spillovers* technologiques »), notamment dues aux échanges d'information et de connaissances plus intenses dans les zones denses, peuvent également améliorer la productivité locale. Comme souligné par Lucas (1988), de telles interactions entre agents localisés dans un même lieu pourraient même plus généralement être à la source du progrès technique et de la croissance. Dans cette perspective, il est important de déterminer si ces *spillovers* sont plus intenses au sein d'un même secteur (« externalités de localisation ») ou, à l'inverse, s'ils sont principalement intersectoriels (« externalités d'urbanisation »).

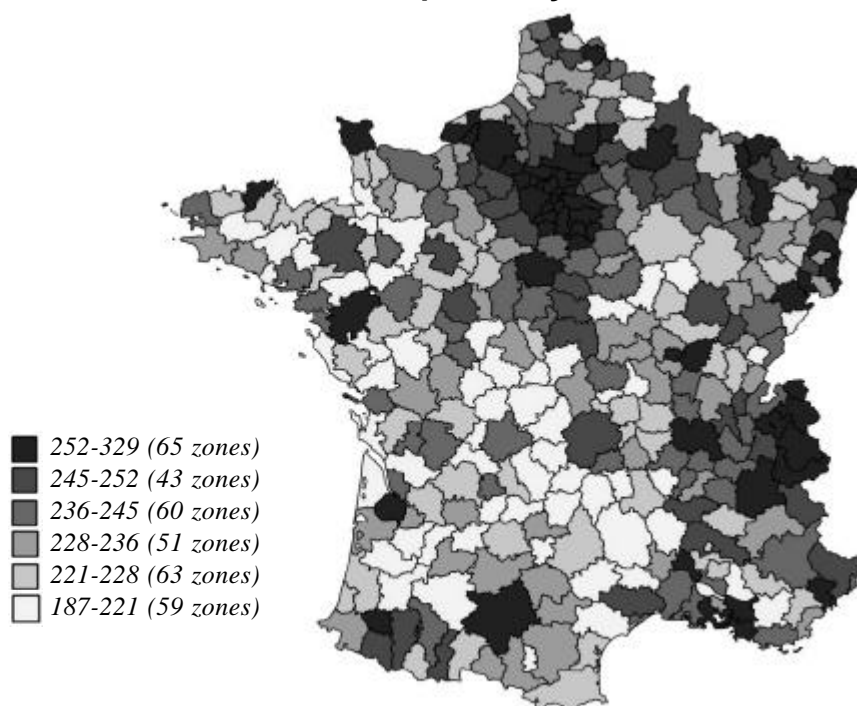
De même que pour le courant d'explications précédent, les gains d'efficacité de cette nature, qui sont liés au regroupement des agents en un même lieu, ne dépendent pas directement des dotations locales. En revanche, ils sont tributaires de la taille et de la composition sectorielle de l'activité économique locale ainsi que des coûts de transport interrégionaux. Ces paramètres sont partiellement contrôlés par les pouvoirs publics. Cependant, les politiques industrielles visant à développer certaines industries n'ont pas donné les résultats espérés au niveau national. Le même type de politique au niveau local risque de se révéler tout aussi décevant. Par ailleurs, l'effet d'une baisse des coûts de transport est, selon l'économie géographique et contrairement à certaines intuitions fausses, de renforcer les avantages dont disposent les zones centrales et développées et de renforcer les incitations à s'y localiser. Ce n'est que pour des coûts de transport faibles que des phénomènes de congestion redonneraient certains avantages aux zones périphériques.

D'un point de vue empirique, Henderson (1988) montre l'existence d'externalités de localisation, alors que Ciccone et Hall (1996) soulignent la présence d'externalités de densité. Ces deux études sur les États-Unis ont été répliquées avec succès un grand nombre de fois. En ce qui concerne la nature des externalités liées aux interactions entre agents, Dumais, Ellison et Glaeser (1997) mettent l'accent sur les éléments relatifs au fonctionnement du marché du travail alors que Jaffe, Trajtenberg et Henderson (1993) montrent que même longtemps après leur dépôt, les brevets restent plus fortement cités à proximité de leur lieu de dépôt, soulignant ainsi l'existence d'externalités technologiques. En revanche, le débat entre externalités de localisation et d'urbanisation n'est pas tranché. Sur données américaines, Glaeser, Kallal, Schleifer et Scheinkman (1992) obtiennent des résultats contraires à ceux d'Henderson, Kuncoro et Turner (1995) quant à l'importance d'un tissu économique spécialisé ou diversifié. D'après Glaeser et *al.* (1992), la diversité renforcerait la croissance locale de l'emploi alors que la spécialisation aurait des effets néfastes. D'après Henderson et *al.* (1995), la spécialisation n'aurait que peu d'effets alors que la diversité n'aurait d'importance que pour les secteurs de haute technologie. En France, Combes (2000) montre que la spécialisation ne favorise que rarement la croissance de l'emploi local, mais que la diversité peut le faire pour certains secteurs de services et de haute technologie.

Ainsi, la distinction entre les trois catégories d'éléments locaux affectant la productivité du travail, dotations, qualifications et interactions est claire. D'un point de vue empirique, il apparaît que chaque type d'explication a reçu un certain nombre de validations. Néanmoins, la littérature n'a jusqu'à présent exploré ces explications que séparément. De plus, le poids relatif de chacune d'elles sur les inégalités salariales reste à déterminer, ce que nous essayons de faire sur données françaises. Cela apparaît comme un besoin important, puisque, comme nous l'avons souligné, l'efficacité des politiques économiques cherchant à agir sur la productivité des agents, et par suite sur l'équité sociale, dépend directement de la source des inégalités spatiales⁽¹⁾.

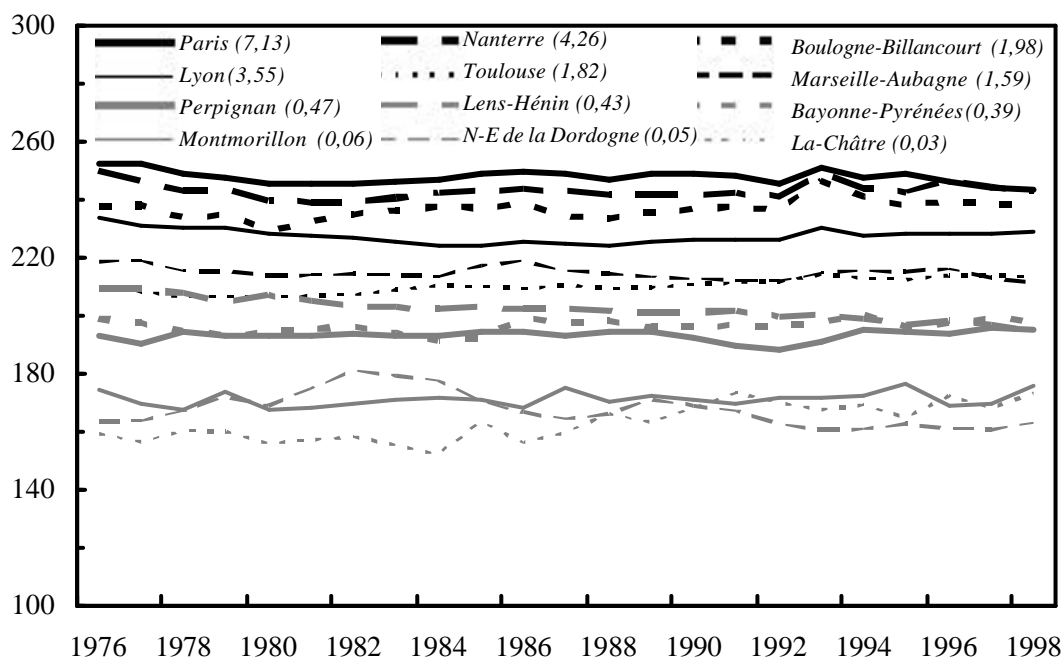
(1) L'objet de ce complément n'est cependant pas d'étudier systématiquement tous les facteurs d'iniquité spatiale ni de déterminer quelles seraient les politiques publiques optimales, mais de simplement en quantifier certains de leurs déterminants. Il est possible de se référer à la revue de littérature de Gérard-Varet et Mougeot (2000) pour plus de détails sur les problèmes d'iniquité spatiale.

1. Coût salarial moyen journalier dans les zones d'emploi françaises en 1998



Source : Calcul des auteurs.

2. Coût salarial moyen détrendé dans quelques zones d'emploi françaises



Note : Les salaires des trois années manquantes sont remplacés par la moyenne des deux années contiguës. Le chiffre entre parenthèses après le nom de la zone d'emploi correspond à la part de la zone d'emploi dans l'emploi total français.

Source : Calcul des auteurs.

2. Les inégalités salariales spatiales en France

Maintenant qu'un certain nombre d'éléments théoriques ont été clarifiés, commençons par établir quelques faits stylisés relatifs aux inégalités spatiales salariales en France. À cette fin, nous utilisons les Déclarations annuelles de données sociales de l'INSEE au 25^e, relatives aux salariés des entreprises publiques et privées nés en octobre d'une année paire. Chaque employeur doit remplir pour chacune de ces personnes et pour chacun de ses établissements une déclaration contenant, entre autres, diverses mesures de salaire et de coût salarial, la catégorie socioprofessionnelle de l'individu ainsi que le code d'activité de l'établissement et sa localisation.

Sébastien Roux de l'INSEE a assemblé ces données sous forme de panel pour les années 1976 à 1998. Ce panel a été utilisé pour la première fois par Abowd, Kramarz et Margolis (1999) dont le travail contient une description détaillée de ces données. Pour certaines années proches des recensements (1981, 1983 et 1990), les données sont indisponibles. Les agriculteurs sont par ailleurs exclus du panel. De plus, nous ne considérons pas les employés de certains secteurs pour lesquels la localisation infrarégionale n'est pas disponible ou pour lesquels le nombre d'observations est trop faible.

Nous travaillons au niveau des 341 zones d'emploi françaises, en considérant 99 secteurs industriels ou de services et 5 types de catégories socioprofessionnelles pour des salariés à temps plein du secteur privé. Après nettoyage des données (valeurs manquantes, secteurs peu représentés, etc.), ce panel contient 8 826 422 observations pour les vingt années observées. Ces données servent de base pour les statistiques descriptives qui suivent et pour les estimations sur données agrégées. En revanche, pour des raisons de capacité informatique, les estimations sur données individuelles ne sont basées que sur six années régulièrement espacées dans le temps, 1976, 1980, 1984, 1988, 1992 et 1996, soit 2 664 474 observations. Les résultats rapportés ci-dessous proviennent du travail en cours de Combes, Duranton et Gobillon (2002).

Le graphique 1 représente le coût salarial moyen des zones d'emploi⁽²⁾. Une grande diversité, avec de fortes variations spatiales, est d'ores et déjà observée. La région parisienne et les grandes villes apparaissent également clairement comme des lieux où les salaires sont plus élevés.

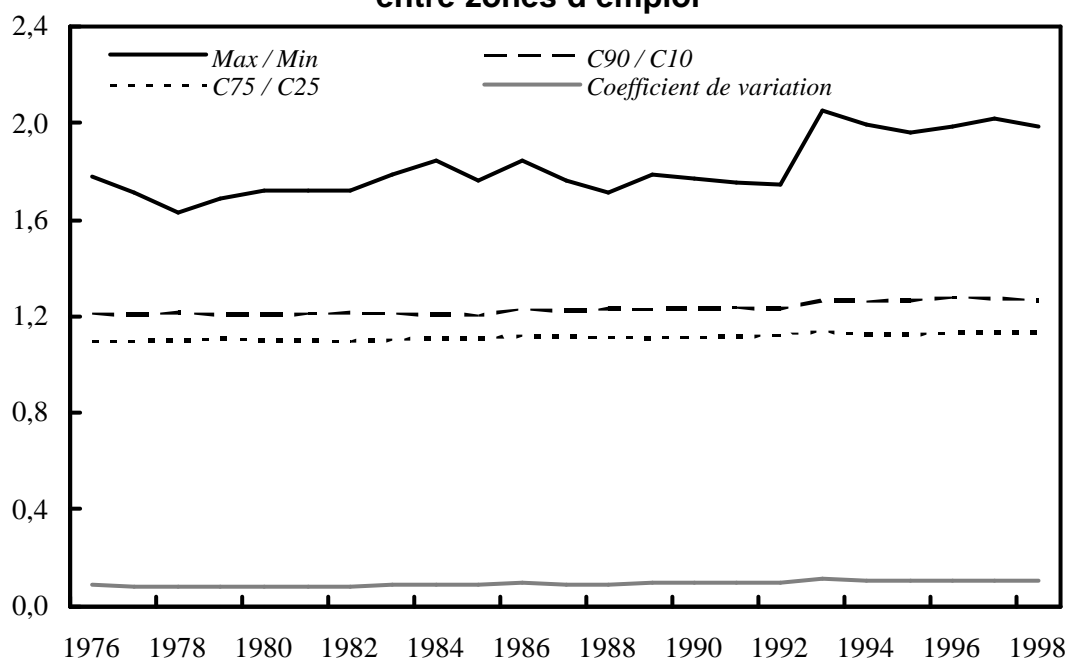
Le graphique 2 représente ce coût salarial dans quelques zones d'emploi françaises au cours des vingt-deux dernières années.

L'ampleur des inégalités salariales en France est confirmée sur ces courbes. Les salaires sont plus de 60 % plus élevés dans certaines zones d'emploi de région parisienne que dans d'autres plus rurales, comme par exemple Montmorillon, le nord-est de la Dordogne ou La Chatre. De plus, au-delà de petites variations de court terme, les salaires, et ces écarts, apparaissent très persistants dans le temps. Une légère augmentation globale des inégalités semble apparaître en fin de période.

(2) Dans tout ce complément, l'unité utilisée est toujours le franc 1980.

Afin de confirmer la stabilité de ces différences locales de salaire, nous calculons pour chacune des années des indicateurs simples d'inégalité. Ainsi, le coefficient de variation, le rapport Max/Min, et les rapports des quantiles à 10 et 90 % et à 25 et 75 % des salaires des zones d'emploi sont représentés sur le graphique 3. Ces indicateurs, stables dans le temps, confirment tous la forte persistance temporelle des inégalités de salaire entre zones d'emploi. La légère hausse des inégalités au cours de la période peut y être décelée, quel que soit l'indicateur utilisé.

3. Indicateurs d'inégalité de coût salarial *détrendé* entre zones d'emploi



Source : Calcul des auteurs.

La stabilité des inégalités salariales spatiales est confirmée par le calcul d'une matrice de transition, représentée sur le tableau 1. Pour chaque année, les zones d'emploi sont alors divisées en cinq classes comprenant chacune 20 % de l'échantillon. La première ligne se lit de la façon suivante : étant dans la classe la plus faible en 1976, une zone d'emploi a 66 % de chances de toujours s'y trouver en 1998. Elle a 23 % de chances de se trouver dans la classe immédiatement au-dessus, 7 et 3 % dans les deux classes suivantes, et aucune d'atteindre la classe la plus élevée. À l'autre extrême, 76 % des zones d'emploi aux salaires les plus élevés en 1976 sont toujours dans cette classe en 1998, 16 % d'entre elles se trouvent dans la classe juste inférieure, 7 % dans la classe suivante et aucune ne se trouve dans les deux classes les plus faibles. De façon plus générale, le chiffre situé sur la diagonale est toujours le plus élevé de la ligne. Cependant, les mouvements ne sont pas totalement négligeables, mais restent le plus fréquents pour les classes voisines de la classe initiale. Cette conclusion est

corroborée par les résultats contenus dans les deux premières colonnes du tableau 2. Les salaires retardés d'une ou cinq périodes expliquent respectivement 94 et 87 % (ligne « R² ») de la variation spatiale des salaires. Le coefficient d'auto-corrélation est également très élevé.

1. Matrice de transition

En 1998 \ En 1976	Classe 1	Classe 2	Classe 3	Classe 4	Classe 5	Nombre de zones
Classe 1	0,66	0,23	0,07	0,03	0,00	69
Classe 2	0,21	0,42	0,32	0,03	0,01	68
Classe 3	0,07	0,28	0,38	0,23	0,03	68
Classe 4	0,06	0,06	0,15	0,54	0,19	68
Classe 5	0,00	0,00	0,07	0,16	0,76	68
Nombre de zones	69	68	68	68	68	341

Source : Calcul des auteurs.

2. Quelques régressions simples du salaire local 1998 sur des variables locales

	Salaire		Densité	Emploi total	Part des cadres	Diversité
	1997	1993				
Constante	0,128 ⁽¹⁾ (0,072)	0,773 ⁽³⁾ (0,100)	5,720 ⁽³⁾ (0,014)	5,147 ⁽³⁾ (0,025)	5,362 ⁽³⁾ (0,006)	5,329 ⁽³⁾ (0,037)
Élasticité	0,976 ⁽³⁾ (0,013)	0,862 ⁽³⁾ (0,018)	0,049 ⁽³⁾ (0,003)	0,049 ⁽³⁾ (0,004)	1,763 ⁽³⁾ (0,085)	0,047 ⁽³⁾ (0,012)
R ²	0,94	0,87	0,51	0,37	0,56	0,04

Notes : (1) Significatif à 10 % ; (2) Significatif à 5 % ; (3) Significatif à 1 % ; Régressions en logarithmes, sauf pour la part des cadres qui est en niveau, écart-type entre parenthèses.

Source : Calcul des auteurs.

Enfin, une étude de corrélations simples montre que ces différences spatiales de salaires sont bien liées à des éléments de la structure économique locale, comme prédit par la théorie. Le tableau 2, colonnes 3 à 6, présente les résultats de régressions, où chaque variable explicative est introduite seule. La densité de l'emploi « explique » à elle seule plus de la moitié de la variation spatiale des salaires (colonne 3). Le niveau d'emploi, quant à lui, explique 37 % de cette même variation (colonne 4). Le taux de cadres a aussi un très fort pouvoir explicatif avec un R² quasiment à 60 % (colonne 5). La diversité sectorielle joue aussi significativement sur le niveau des salaires, mais n'explique qu'une faible part (4 %) de sa variation spatiale (colonne 6).

3. Une quantification des facteurs d'inégalité : dotations, qualifications et interactions

3.1. Aspects méthodologiques

La méthodologie que nous employons pour comprendre les inégalités spatiales mises en évidence dans la partie précédente et en quantifier les différents déterminants est issue d'un modèle théorique très simple reposant sur les hypothèses suivantes. Les technologies sont supposées à rendements constants, n'utilisant que du travail. Les différentes qualifications sont supposées parfaitement substituables. La productivité du travail ne dépend ni de l'échelle de production ni d'autres *inputs*, mais varie avec la qualification. Le marché du travail est supposé concurrentiel. Le salaire est alors égal à la productivité marginale du travail. Ainsi, peut-on aussi interpréter nos résultats sur les salaires en termes de productivité. Enfin, la productivité totale des facteurs, qui s'applique à chaque établissement localisé dans la zone d'emploi a et opérant dans le secteur k à la date t , $A_{a,k,t}$, dépend des différents éléments présentés dans la première partie :

$$A_{a,k,t} = A_{k,t} (E_{a,t}, I_{a,t}, I_{a,k,t}, S_{a,k,t})$$

La productivité des facteurs, $A_{k,t}(\cdot)$ est une fonction qui dépend du secteur et de la date considérée mais pas de la zone d'emploi. Toutefois, ses arguments, eux, sont locaux. Les effets des dotations de la zone sont capturés par un vecteur de variables $E_{a,t}$ comprenant par exemple les dotations physiques, les caractéristiques géographiques, les infrastructures publiques et les différences institutionnelles. Les interactions intrazone mais intersectorielles se reflètent dans le vecteur de variables $I_{a,t}$. Ce vecteur inclut le niveau d'emploi total, la superficie de la zone d'emploi (pour capturer les effets de la densité) et la diversité sectorielle totale. Les interactions intrazone et intrasectorielle, $I_{a,k,t}$, comprennent la spécialisation locale dans le secteur et le nombre d'entreprises dans la zone-secteur. Les effets de qualification sont enfin capturés par le vecteur $S_{a,k,t}$ qui contient les parts moyennes de chaque qualification ainsi que l'âge moyen des salariés dans l'emploi sectoriel local et son carré.

Ainsi, sur données agrégées par zone d'emploi et secteur, le modèle estimé est le suivant :

$$(1) \quad \begin{cases} \log w_{a,k,t} = \mathbf{m}_k + \mathbf{b}_{a,t} + I_{a,k,t} \mathbf{g}_k + S_{a,k,t} \mathbf{d}_k + \mathbf{e}_{a,k,t} \\ \mathbf{b}_{a,t} = \mathbf{w}_0 + \mathbf{n}_t + I_{a,t} \mathbf{q} + \mathbf{h}_{a,t} \end{cases}$$

avec $\mathbf{h}_{a,t} = E_{a,t} \mathbf{a} + u_{a,t}$.

Le but est d'étudier la part de la variabilité spatiale des salaires, $w_{a,k,t}$, expliquée par chacun des éléments mentionnés, dotations, interactions et

qualifications. Ces salaires sont tous exprimés en francs constants de 1980. En première étape, nous introduisons d'abord des indicatrices sectorielles, m_k . Elles permettent de déterminer quelle est la part des inégalités spatiales de salaires expliquée par des éléments liés aux différences nationales entre secteurs. Par exemple, si un secteur emploie plus de cadres et que ceux-ci bénéficient d'un salaire plus élevé de façon générale, le salaire apparaît plus élevé dans les zones où ce secteur est surreprésenté, sans que cela soit lié à des externalités locales. Les effets fixes sectoriels captent ce type de phénomènes.

Aux côtés des ces indicatrices, nous introduisons ensuite les variables explicatives locales et sectorielles liées aux effet d'interactions, $I_{a,k,t}$, et de qualifications, $S_{a,k,t}$. Si les externalités soulignées par l'économie géographique sont pertinentes, les variables de $I_{a,k,t}$ (la spécialisation définie comme la part du secteur dans l'emploi local et le nombre d'établissements de la zone-secteur) devraient être significatives et expliquer une large part de la variance. Les parts des différentes qualifications dans le vecteur $S_{a,k,t}$ déterminent par ailleurs si au niveau local, une part plus importante qu'au niveau national de chaque qualification accroît les salaires ou non. Le même argument s'applique à l'âge moyen et à son carré aussi contenus dans le même vecteur. Notons enfin que les effets de ces variables locales peuvent dépendre du secteur considéré et que toutes les variables sont centrées par rapport à leur moyenne secteur-temps.

Les effets sectoriels, d'interactions et de qualification étant contrôlés, des indicatrices zone-temps, $b_{a,t}$, capturent enfin tous les effets locaux non sectoriels. Sous les hypothèses de centrages faites plus haut, ces indicatrices représentent un salaire local net des effets sectoriels au niveau national. Ce salaire est aussi expurgé des effets d'interactions et de qualifications et d'âge au niveau sectoriel. Ses variations spatiales sont dues aux interactions intersectorielles et aux effets de dotations. Ce salaire est donc régressé en deuxième étape sur des indicatrices annuelles, n_t , qui captent les variations temporelles et spatiales de salaire, et sur les variables d'interactions intersectorielles, $I_{a,t}$, incluant les effets de la densité, de la superficie et de la diversité sectorielle globale. Ne disposant pas de données suffisamment précises relatives aux dotations des zones d'emploi, l'effet de celles-ci, $E_{a,t} \mathbf{a}$, est inclus dans le résidu $h_{a,t}$. Si l'on ne peut distinguer l'impact des dotations d'infrastructures publiques de celles de la géographie physique locale par exemple, cela permet néanmoins de distinguer la part de la variance expliquée par ce que nous avons appelé les dotations locales de celle expliquée par les interactions locales intersectorielles.

L'approche qui vient d'être présentée peut être vue comme une généralisation de ce qui a été fait dans la littérature par le passé, en considérant cependant simultanément tous les types d'explications. Notre avantage est aussi de disposer de données relativement exhaustives couvrant une pé-

riode beaucoup plus longue avec des unités géographiques construites selon des critères économiques, les zones d'emploi, et non administratives, ce qui peut parfois introduire des artefacts statistiques. Toutefois, l'utilisation de données agrégées conduit à deux types de difficultés. Premièrement, il est difficile d'identifier correctement certains effets. En particulier, l'impact moyen des qualifications et de l'âge ne permet pas de distinguer les effets de la qualification de chaque individu d'éventuelles externalités de capital humain (la rencontre d'individus qualifiés accroît leur productivité). Le second problème a trait à l'autosélection des individus dans l'espace. Il est possible d'imaginer qu'à qualification égale, les salariés à Paris aient un salaire plus élevé, non pas à cause d'une meilleure dotation en infrastructures, mais parce que Paris sélectionne des individus plus dynamiques et productifs. Dans les deux cas, seules des données individuelles permettent de palier ces problèmes liés à l'agrégation et de contrôler beaucoup mieux pour les différences de qualifications, d'expérience et d'aptitude entre individus.

La spécification estimée sur données individuelles est la suivante, l'indice i étant relatif à un individu donné :

$$(2) \quad \begin{cases} \log w_{i,t} = \mathbf{m}_{k(i,t)} + \mathbf{b}_{a(i,t),t} + I_{a(i,t)k(i,t)} \mathbf{g}_k + X_{i,t} \mathbf{j} + \mathbf{d}_i + \mathbf{e}_{i,t} \\ \hat{w}_{a(i,t),t} = \mathbf{w}_0 + \mathbf{n}_t + I_{a(i,t)} \mathbf{q} + \mathbf{h}_{a(i,t),t} \end{cases}$$

avec $\mathbf{h}_{a,t} = E_{a,t} \mathbf{a} + u_{a,t}$.

Hormis des méthodes d'estimation différentes, la principale modification réside dans le fait que les effets de qualification et d'expérience ne sont plus contrôlés simplement au niveau agrégé par la catégorie socioprofessionnelle et l'âge moyen des employés et son carré, mais par un effet individuel propre à chacun d'entre eux, \mathbf{d}_i , et par leur propre âge et son carré (centrés comme les autres variables explicatives par rapport à la moyenne du secteur-temps), inclus dans les variables $X_{i,t}$. Ainsi, les effets d'interactions locales, sectoriels ou non, et de dotations, sont alors nettement mieux isolés des effets de qualification et d'expérience. Notons finalement que seul un panel individuel permet ce type d'estimation, puisqu'il est nécessaire d'observer plusieurs fois chaque individu dans le temps pour pouvoir inclure de tels effets fixes individuels dans les régressions tout en identifiant les autres paramètres d'intérêt.

3.2. Interactions intrasectorielles locales, qualifications et expérience, et effets purement locaux

Le tableau 3 donne les R^2 de différentes spécifications correspondant à l'estimation sur données agrégées de la première équation du modèle (1).

Les indicatrices temporelles expliquent à elles seules 21 % de la variance des salaires locaux et sectoriels de l'année. Lorsqu'on introduit en

plus les indicatrices sectorielles, la part de la variance des salaires sectoriels locaux instantanés expliquée s'élève à 51 %. Ces chiffres élevés peuvent résulter de différences technologiques dans le temps et surtout entre secteurs ou par le fait que les secteurs n'emploient pas les mêmes proportions de chaque qualification. Comme nous l'avons déjà souligné, si la fraction des qualifications fortement rémunérées est plus importante dans un secteur, les employés de ceux-ci le sont plus en moyenne, et si de plus les secteurs ne sont pas uniformément répartis dans l'espace, cela génère des différences spatiales de salaire. Nos indicatrices sectorielles permettent de contrôler cela.

3. Part de la variance des salaires expliquée, données agrégées

Régression	(1)	(2)	(3)	(4)	(5)	(5)	(6)	(7)	(8)	(9)
Indicatrices annuelles	oui	oui	oui	oui	oui	oui	oui	oui	oui	oui
Indicatrices sectorielles	non	oui	oui	oui	oui	non	oui	oui	oui	oui
Indicatrices zone-temps	non	non	non	non	non	oui	oui	oui	oui	oui
Interactions intrasectorielles	non	non	oui	non	oui	non	non	oui	non	oui
Qualifications / Âge	non	non	non	oui	oui	non	non	non	oui	oui
R ²	0,21	0,51	0,63	0,74	0,76	0,46	0,69	0,73	0,80	0,81

Note : Pour faciliter l'interprétation de cette analyse de la variance, les indicatrices temporelles sont introduites dès la première étape.

Source : Calcul des auteurs.

Afin de déterminer la part relative de la variabilité purement spatiale expliquée par chaque groupe de variable, il s'agit maintenant de comparer la variation de R² par rapport à 0,51 que chacune d'elle induit.

Le premier constat fort est que les interactions locales intrasectorielles n'expliquent qu'une part supplémentaire assez faible des inégalités de salaire, induisant une variation du R² de 0,12. Pourtant, les élasticités relatives à la spécialisation sont précisément estimées dans une grande majorité des cas. Parmi 99, 98 sont positives et significatives à 5 %. Ainsi, les externalités de localisation liées à la spécialisation sectorielle semblent significatives en France : lorsqu'un secteur a une part dans l'emploi local supérieure à sa part nationale, il bénéficie d'une productivité plus élevée. Néanmoins, ce type d'externalités ne compte que très peu pour expliquer les inégalités spatiales de salaire.

À l'inverse, les parts des différentes qualifications et les effets d'âge expliquent une fraction beaucoup plus importante des inégalités spatiales. Le gain de R² est alors de 0,23, soit une valeur deux fois plus élevée que celle due aux interactions intrasectorielles. On retrouve naturellement que les qualifications élevées obtiennent des salaires significativement plus élevés que la moyenne. Le contraire se produit pour les qualifications les plus

faibles. N'oublions pas que les effets secteurs sont déjà pris en compte et que les variables explicatives sont centrées par rapport à leur moyenne dans le secteur à la date donnée. Ainsi, c'est le fait de disposer au niveau local d'une plus grande proportion de cadres qu'au niveau national dans un secteur donné qui induit des salaires plus élevés, et non le fait d'utiliser plus de cadres dans ce secteur que dans d'autres de façon générale, effet déjà contrôlé par les indicatrices sectorielles. Ces différences de qualification intrasectorielle mais interzones d'emploi contribuent donc largement à expliquer les variations spatiales de productivité. Les effets d'interactions sectorielles locales, de qualification et d'âge font varier le R^2 de 0,25 quand elles sont introduites toutes ensemble, ce qui signifie que ces effets ne sont que très partiellement orthogonaux.

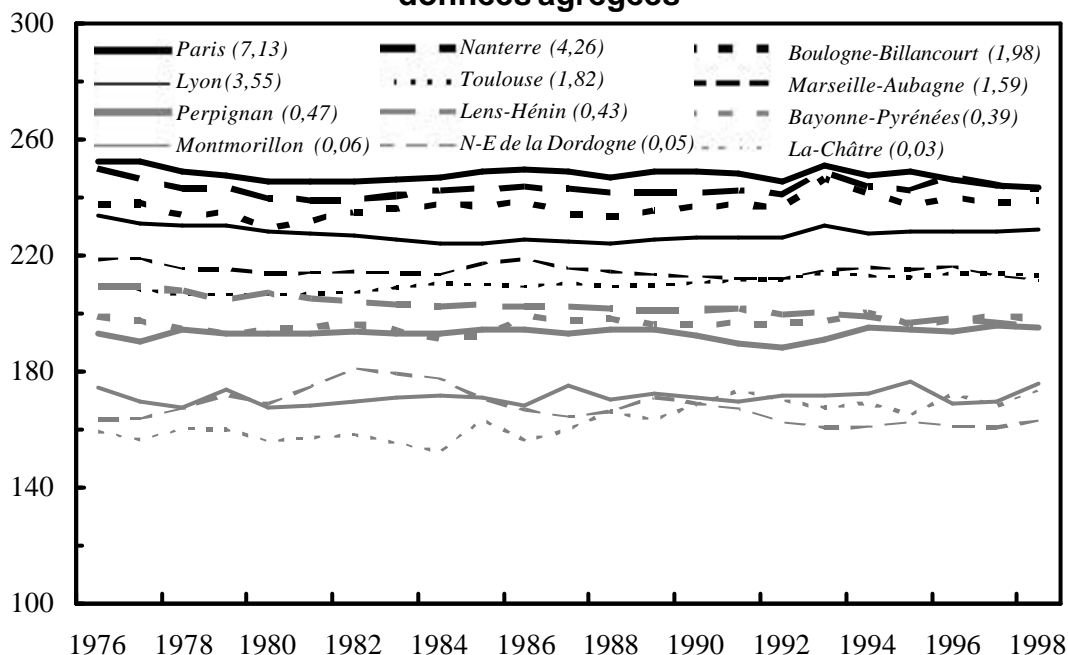
Les effets purement locaux et non sectoriels, capturés par les indicatrices zone-temps, expliquent seuls (c'est-à-dire sans que les variables d'interactions intrasectorielles et de qualification ne soient simultanément introduites) un peu moins que les qualifications, entraînant une variation de R^2 de 0,18. Ces effets sont dans une large mesure orthogonaux aux effets purement sectoriels et non locaux. En effet, la variation de R^2 qu'ils entraînent par rapport au modèle incluant uniquement les effets temporels est de 0,25 contre 0,30 pour les effets secteurs purs, mais lorsque les deux types d'effets sont simultanément introduit, le R^2 s'élève à 0,69, soit une variation de 0,48 largement supérieure à celle de chacun des effets introduits séparément. L'introduction supplémentaire des effets d'interactions locales intrasectorielles n'augmente toujours que faiblement la part de la variance expliquée, de 0,04. L'effet intrasectoriel des qualifications est quant à lui plus orthogonal aux effets purement locaux, la part de la variance expliquée augmentant de 0,11. Finalement, tous les effets considérés simultanément amènent le R^2 à 0,81, soit une augmentation de 0,30 par rapport au modèle n'incluant pas d'effets locaux.

Les résultats précédents correspondent à ce que chaque groupe de variables introduit seul explique, c'est-à-dire la variation maximale de variance des salaires qu'il induit. On peut à l'inverse étudier la diminution de R^2 résultant de l'omission d'un des groupes de variables par rapport à la régression complète. On obtient ainsi la variation minimale de variance induite par le groupe de variables. Nous pouvons alors ainsi résumer nos résultats : la variation de R^2 induite par les interactions sectorielles se situe entre 0,01 et 0,12, celle des effets de qualifications et d'âge entre 0,08 et 0,23 et celle des effets purement locaux entre 0,05 et 0,18. Ainsi, les effets de qualification et d'expérience sont les plus forts, ceux d'interactions intrasectorielles les plus faibles, alors que les effets d'interactions intersectorielles seraient intermédiaires, plus proche de l'ordre des effets de qualification et d'expérience tout de même.

Ce type d'analyse de la variance reste à poursuivre sur données individuelles, mais il est alors nettement plus complexe. En revanche, nous proposons un autre point de vue permettant d'illustrer de façon simple nos résul-

tats. Tout d'abord, le graphique 4 présente pour les douze zones d'emploi déjà considérées dans le graphique 2, le coût salarial net des effets de première étape, c'est-à-dire net des effets sectoriels, d'interactions intra-sectoriels, de qualification et d'âge et *détrendé*. Il s'agit donc de la variable dépendante de deuxième étape, nette des effets purement temporels.

4. Coût salarial net des effets de première étape et *détrendé*, données agrégées



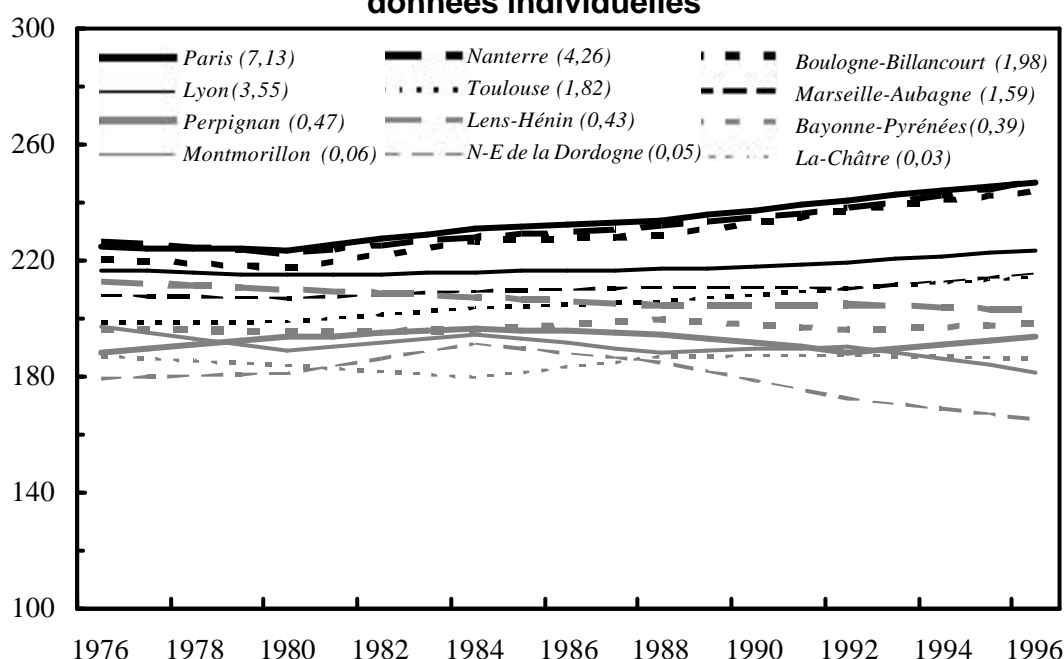
Source : Calcul des auteurs.

Il apparaît alors déjà clairement que les inégalités entre les zones d'emploi de région parisienne ou correspondant aux plus grandes villes françaises, d'une part, et les zones d'emploi des villes secondaires ou de campagne, d'autre part, se sont resserrées, surtout à partir des années quatre-vingt. Les effets non sectoriels, et notamment la taille de la zone d'emploi, ne sont donc pas les seuls facteurs jouant sur les inégalités, mais la composition de l'emploi en termes de secteur et de qualification en est également responsable. Lorsque, sur données individuelles, les caractéristiques de chaque salarié sont prises en compte, l'effet est plus marqué, et les inégalités de salaire encore plus réduites. Cela apparaît sur le graphique 5 où le salaire net des effets de première étape et *détrendé*, correspondant cette fois-ci aux estimations sur données individuelles, est tracé pour les zones d'emploi sélectionnées. Il peut être également intéressant de noter que cette baisse d'inégalités est cette fois-ci très forte en début de période et un peu moins par la suite.

Le tableau 4 confirme ces constats, non plus sur quelques zones d'emploi particulières, mais sur l'ensemble de celles-ci. Pour ce faire, nous recalculons les indicateurs d'inégalité du graphique 3. La colonne « Brut » donne la moyenne (sur les années observées) des indicateurs d'inégalités

calculés sur la série brute des coûts salariaux. La première colonne « Net » donne ces indicateurs, calculés cette fois-ci sur le coût salarial net des effets de première étape pour les estimations sur données agrégées et la deuxième colonne « Net » la même chose pour les estimations sur données individuelles. Lorsque les effets sectoriels, d'interactions intrasectorielles, de qualification et d'âge sont pris en compte, les inégalités spatiales de salaire sont plus faibles que sur données brutes, et ce quel que soit l'indicateur considéré. Ces inégalités spatiales de salaire sont encore plus faibles lorsque les effets de qualification et d'âge sont remplacés par des effets spécifiques à chaque individu captant par exemple également des différences interindividuelles d'aptitude, à qualification et âge donnés.

5. Coût salarial net des effets de première étape et *détrendé*, données individuelles



Source : Calcul des auteurs.

4. Indicateurs d'inégalité moyens sur l'ensemble de la période étudiée

	Brut	Données agrégées		Données individuelles	
		Net	Résiduel	Net	Résiduel
Max / Min	1,88	1,69	1,47	1,51	1,35
C90 / C10	1,23	1,20	1,09	1,14	1,09
C75 / C25	1,12	1,10	1,04	1,06	1,04
Variation	0,094	0,075	0,036	0,054	0,034

Notes : Max / Min, C90 / C10, C75 / C25, Variation : Moyenne sur l'ensemble des années utilisées (20 points en agrégé, 6 en individuel) ; Brut : Coût journalier net des effets temps ; Net : Coût journalier net des effets temps, secteurs, d'interactions intrasectorielles, de qualification et d'âge (ou d'effet individuel) ; Résiduel : Coût journalier net des effets temps, secteurs, d'interactions intrasectorielles, de qualification et d'âge (ou d'effet individuel) et d'interactions intersectorielles.

Source : Calcul des auteurs.

3.3. Dotations vs interactions locales intersectorielles

Il nous reste à déterminer dans la variabilité spatiale résiduelle (c'est-à-dire le salaire local net des effets sectoriels, d'interactions intrasectorielles, de qualification et d'âge et *détrendé*), la part expliquée par les effets d'interactions locales intersectorielles (liés à la densité de l'emploi, à la superficie et à la diversité sectorielle globale) et la part des effets des dotations. Comme mentionné précédemment, ne disposant pas de données adéquates, cette dernière n'est qu'appréhendée comme, au maximum, la part de la variance non expliquée par les variables d'interactions locales non sectorielles.

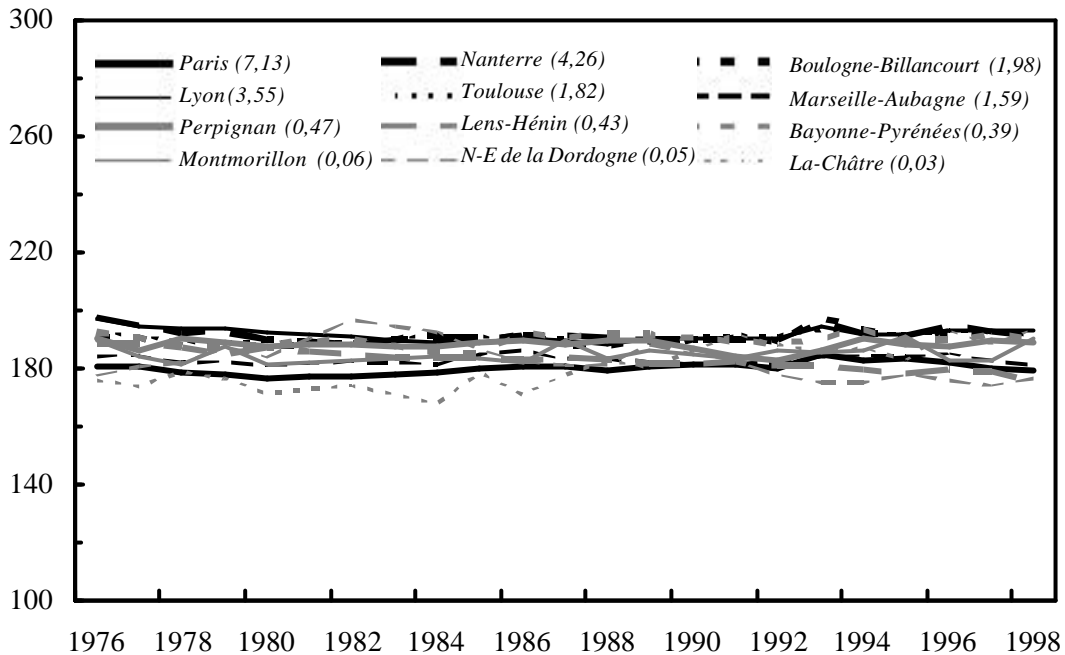
Ainsi, la deuxième étape de l'estimation consiste à régresser les salaires locaux nets des effets de première étape, sur la densité, la superficie de la zone et un indice de diversité sectorielle (inverse d'un indice d'Herfindhal calculé sur les parts de chaque secteur dans l'emploi local). L'effet de densité est fort : doubler la densité induit une augmentation de salaire d'environ 3 %. Rappelons que les zones urbaines sont souvent plus de dix fois plus denses que les zones rurales, sans prendre les cas les plus extrêmes⁽³⁾. À densité donnée, un doublement de la superficie, ce qui revient à doubler l'emploi total, exerce un effet supplémentaire sur le salaire également de l'ordre de 3 %. L'impact de la diversité sectorielle, souvent étudié dans les études empiriques sur la croissance locale, est positif et significatif quand elle est introduite seule. Elle a toutefois un effet beaucoup plus faible et parfois non significatif pour certaines spécifications du modèle lorsque densité et superficie sont simultanément introduites. De plus, cette variable n'explique qu'une part assez faible de la variance spatiale de l'indice de salaire, 10 % lorsqu'elle est introduite seule. Cela n'est en revanche pas le cas de la densité et de la superficie qui en expliquent la grande majorité, 77 %, l'ajout de la diversité ne faisant alors quasiment pas varier ce chiffre.

Toutes ensemble, les variables d'interactions locales non sectorielles expliquent donc 77 % de la variance spatiale des salaires nets des effets de première étape et *détrendés*. Ainsi, les dotations locales ne seraient responsables, au plus, que de 23 % de celle-ci, tout en n'oubliant pas que les effets purement locaux ne font varier le R^2 de première étape que de 0,18 au plus.

Là encore, l'importance des effets d'interactions intersectorielles est clairement illustrée par la représentation du coût salarial résiduel, le terme d'erreur de l'équation de deuxième étape, pour les douze zones d'emploi que nous avons considérées. Le graphique 6 représente ainsi le coût salarial net des effets temporels, sectoriels, d'interactions intra et intersectorielles, ainsi que des effets de qualification et d'expérience, dont les écarts entre zones d'emploi ne peuvent être attribués qu'aux effets de dotations ou à des éléments purement aléatoires.

(3) Le rapport P90/P10 pour la densité d'emploi est égal à 13,2.

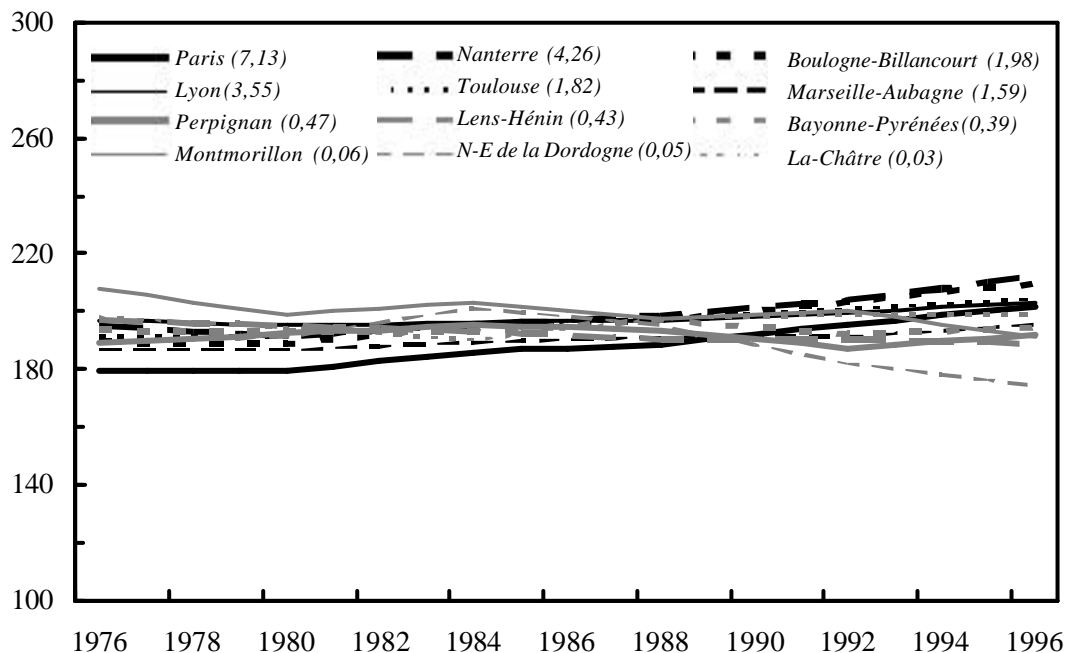
6. Coût salarial résiduel *détrendé*, données agrégées



Source : Calcul des auteurs.

La faiblesse des inégalités de coût salarial résiduel (sur données agrégées) entre zones d'emploi est alors frappante. Le salaire à Nanterre par exemple n'est alors qu'à peine 10 % plus élevé que dans les zones les plus rurales. Les estimations sur données individuelles présentées sur le graphique 7 conduisent aux mêmes ordres de grandeur.

7. Coût salarial résiduel *détrendé*, données individuelles



Source : Calcul des auteurs.

Finalement, les colonnes « résiduel » du tableau 4 donnent la valeur moyenne des indicateurs d'inégalité résiduelle (pouvant être attribuée aux seuls effets de dotations ou à des effets aléatoires). Les variations par rapport aux colonnes « Net » montrent l'impact important des effets d'interaction intersectorielles et notamment de la densité. Les variations par rapport aux inégalités initiales de coût salarial brut confirment que les effets de dotations ne peuvent expliquer qu'une part mineure de celles-ci. Enfin, il est intéressant de noter que les estimations agrégées comme individuelles conduisent à des degrés d'inégalité résiduelle assez comparables, alors que le pouvoir explicatif du modèle en première étape est bien plus élevé pour les estimations individuelles. Ainsi, une conjecture possible serait que les zones d'emploi les plus importantes sélectionnent les individus les plus aptes, même à qualification et âge donnés.

Conclusions

Ce complément a tout d'abord pour objet de clarifier le débat théorique quant aux origines des inégalités spatiales de productivité et de salaire. Nous rappelons que trois grandes familles d'explications peuvent être à la source de ces inégalités : les effets de dotations locales (géographie physique, infrastructures publiques, différences technologiques), les effets des interactions locales (non sectorielles : densité des activités économiques et diversité sectorielle globale ; sectorielles : spécialisation et nombre d'établissements dans un secteur donné) et enfin les effets de différences locales de qualification des employés.

Nous montrons que les inégalités salariales entre les 341 zones d'emploi françaises sont fortes, les salaires étant, par exemple, jusqu'à 60 % plus élevés dans les zones urbaines que dans les zones rurales. En termes d'explication de ces inégalités, les effets de dotations ont un impact deux à trois fois moins fort que celui des interactions intersectorielles. Les interactions locales intrasectorielles, centrales dans certains modèles d'économie géographique sont très significatives, mais n'expliquent qu'une faible part des différences spatiales de productivité. Les différences de qualifications intrasectorielles locales peuvent à l'inverse en expliquer une part assez élevée, supérieure à celle due aux effets simultanés des dotations et des interactions intersectorielles. La comparaison des indicateurs d'inégalités de coût salarial entre zones d'emploi montre une baisse drastique de celles-ci quand les effets sectoriels, d'interactions intrasectorielles, de qualification et d'âge sont éliminés. La baisse est encore plus nettement marquée quand les effets d'interactions intersectorielles sont aussi purgés. Par exemple, le salaire à Nanterre n'est alors plus élevé que de 10 % par rapport à celui des zones d'emploi rurales, seuls les effets de dotations n'étant alors pas contrôlés.

Il apparaît ainsi que jouer sur les dotations des zones d'emploi ne devrait pas avoir d'effet spectaculaire sur les différences de productivité entre

celles-ci. En revanche, l'agglomération spatiale, de façon générale et non nécessairement secteur par secteur, affecte de façon sensible ces inégalités. Par ailleurs, les différences de qualification entre zones d'emploi, même à l'intérieur de chaque secteur, sont responsables d'une large part des inégalités salariales. C'est donc bien aux déterminants des migrations humaines et aux facteurs de localisation des entreprises qu'il faut s'intéresser si l'on veut encore mieux comprendre ces inégalités spatiales de salaire et productivité, et si l'on veut par la suite éventuellement peser dessus. En effet, des phénomènes plus complexes pourraient être à l'œuvre et les éléments présentés interagir de façon imbriquée. Par exemple, de meilleures dotations publiques, de type aménités par exemple, pourrait attirer des migrants plus qualifiés. Un cadre théorique plus complet accompagné d'estimations structurelles serait alors nécessaire pour appréhender de tels phénomènes et mieux comprendre les déterminants des différences de productivité locale et des inégalités salariales spatiales.

Références bibliographiques

- Abdel-Rahman H. (1988) : « Product Differentiation, Monopolistic Competition and City Size », *Regional Science and Urban Economics*, n° 18, pp. 69-86.
- Abowd J., F. Kramarz et D. Margolis (1999) : « High Wage Workers and High Wage Firms », *Econometrica*, n° 67, pp. 251-333.
- Ciccone A. et R. Hall (1996) : « Productivity, and the Density of Economic Activity », *American Economic Review*, n° 86, pp. 54-70.
- Combes P-P. (2000) : « Economic Structure and Local Growth: France, 1984-1993 », *Journal of Urban Economics*, n° 47, pp. 329-355.
- Combes P-P., G. Duranton et L. Gobillon (2002) : « Wages Differences Across French Local Labour Markets: Endowments, Skills, and Interactions », *Document de Travail du CERAS*.
- Dumais G., G. Ellison et E. Glaeser (1997) : « Geographic Concentration as a Dynamic Process », *NBER Working Paper*, n° 6270.
- Duranton G. et V. Monastiriotis (2002) : « Mind the Gap: The Evolution of Regional Inequalities in the UK, 1982-1997 », *Journal of Regional Science*, n° 42, pp. 219-256.
- Durlauf S. et D. Quah (1999) : « The New Empirics of Economic Growth » in *Handbook of Macroeconomics IA*, Taylor et Woodford (eds), Elsevier, pp. 231-304.
- Fujita M. (1988) : « A Monopolistic Competition Model of Spatial Agglomeration: A Differentiated Product Approach », *Regional Science and Urban Economics*, n° 18, pp. 87-124.

- Gérard-Varet L-A. et M. Mougeot (2000) : « L'État et l'aménagement du territoire » in *Aménagement du territoire*, Rapport du Conseil d'Analyse Économique, n° 31, La Documentation française.
- Glaeser E. et D. Maré (2001) : « Cities and Skills », *Journal of Labor Economics*, n° 19, pp. 316-342.
- Glaeser E., H. Kallal, A. Schleifer et J. Sheinkman (1992) : « Growth in Cities », *Journal of Political Economy*, n° 100, pp. 1126-1152.
- Hall R. et C. Jones (1999) : « Why Do Some Countries Produce Much More Output per Worker than Others? », *Quarterly Journal of Economics*, n° 114, pp. 83-116.
- Helsley R. et W. Strange (1990) : « Matching and Agglomeration Economies in a System of Cities », *Regional Science and Urban Economics*, n° 20, pp. 189-212.
- Henderson J.V. (1988) : *Urban Development: Theory, Fact and Illusion*, Oxford, Oxford University Press.
- Henderson J.V., A. Kuncoro et M. Turner (1995) : « Industrial Development in Cities », *Journal of Political Economy*, n° 103, pp. 1067-1152.
- Jaffe A., M. Trajtenberg et R. Henderson (1993) : « Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations », *Quarterly Journal of Economics*, n° 63, pp. 577-598.
- Krugman P. (1991) : « Increasing Returns and Economic Geography », *Journal of Political Economy*, n° 99, pp. 483-499.
- Krugman P. et A. Venables (1995) : « Globalization and The Inequality of Nations », *Quarterly Journal of Economics*, n° 110, pp. 857-880.
- Landes D. (1999) : *The Wealth and Poverty of Nations: Why Some Are So Rich and Others So Poor?*, New York, Norton.
- Lucas R. (1988) : « On the Mechanics of Economic Development », *Journal of Monetary Economics*, n° 22, pp. 3-42.
- Temple J. (1999) : « The New Growth Evidence », *Journal of Economic Literature*, n° 37, pp. 112-156.