

Spatial wage disparities: Sorting matters!

Pierre-Philippe Combes* §

CERAS

Gilles Duranton* ‡

London School of Economics

Laurent Gobillon* †

University College London and CREST

This draft, 6 December 2003

ABSTRACT: Spatial wage disparities can result from spatial differences in the skill composition of the workforce, in non-human endowments, and in local interactions. To distinguish between these explanations, we estimate a model of wage determination across local labour markets using a very large panel of French workers. We control for worker characteristics, worker fixed-effects, industry fixed-effects, and the characteristics of the local labour market. Our findings suggest that individual skills account for a large fraction of existing spatial wage disparities with strong evidence of spatial sorting by skills. Interaction effects are mostly driven by the local density of employment. Not controlling for worker heterogeneity biases estimates of agglomeration economies by up to 100%. We also find evidence of various omitted variable biases and reverse causality between agglomeration and high wages. Finally, endowments only appear to play a small role.

Key words: Local labour markets, spatial wage disparities, panel data analysis, sorting.

JEL classification: R23, J31, J61.

*We are grateful to Lionel Fontagné, Vernon Henderson, Francis Kramarz, Thierry Magnac, Gianmarco Ottaviano, Barbara Petrongolo, Diego Puga, Jean-Marc Robin, Sébastien Roux, Jon Temple, Dan Trefler, and especially Henry Overman for fruitful discussions, advice and encouragements. We also acknowledge Francis Kramarz and Sébastien Roux's kind and efficient help with the data. Seminar and conference participants in Bristol, Brown, CREST, Kiel, Lille, LSE, Marseille, Paris, Philadelphia, Rome, Royal Holloway, Stockholm, Stoke-Rochford, Toronto, UCL, and Villars also provided us with very useful feed-back.

§CERAS, 48, Bd Jourdan, 75014 Paris, France, combes@enpc.fr, <http://www.enpc.fr/ceras/combes/>. CNRS researcher also affiliated with the Centre for Economic Policy Research. Hospitality from the economics department at Boston University is gratefully acknowledged.

‡Department of Geography and Environment, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom, g.duranton@lse.ac.uk, <http://cep.lse.ac.uk/~duranton>. Also affiliated with the Centre for Economic Policy Research, and the Centre for Economic Performance at the London School of Economics. Funding from the Leverhulme Trust is gratefully acknowledged.

†UCL, Department of Economics, Gower Street, London WC1E 6BT, United Kingdom, gobillon@ensae.fr.

1. Introduction

In many countries, spatial disparities are large and a source of considerable policy concern. In most developed countries, the workers in the richest regions have incomes or wages that are typically double those in the poorest regions. In developing countries, the gaps are often larger.¹ In this paper we propose a new approach to account for spatial wage disparities. We implement it on a large panel of French workers.

To explain large spatial wage disparities, three broad sets of explanations can be proposed. First, differences in wages across areas could *directly* reflect spatial differences in the skill composition of the workforce. There are good reasons to suspect that workers may sort across employment areas so that the measured and un-measured productive abilities of the local labour force may vary. For instance, industries are not evenly distributed across areas and require different labour mixes. Consequently, we expect a higher mean wage in areas specialised in more skill-intensive industries. Differences in local amenities may also imply some sorting by skills across areas. All these skills-based explanations essentially assume that the wage of worker i is given by $w_i = As_i$, where s_i denotes individual skills and A , the productivity of labour, is independent of location. Consequently in area a the average wage is the product of the average skill in the area, \bar{s}_a , by the productivity of labour: $w_a = A\bar{s}_a$.

That sorting could be at the root of systematic wage differences between groups of workers is a long-standing concern of labour economists. They researched this question intensively in the case of wage differences across industries (Krueger and Summers, 1988; Gibbons and Katz, 1992; Abowd, Kramarz, and Margolis, 1999) but they have mostly left aside the geographic dimension. On the other hand, scholars interested in regional issues have paid remarkably little attention to this type of explanation.²

Instead, the study of spatial disparities has mostly focused on two alternative strands of explanations, which both assert that these disparities reflect 'true' productivity differences across places. The first alternative contends that wage differences across areas are caused by differences in local non-human endowments (hereafter endowments). For instance, workers in some areas may have a higher marginal product than in others because of geographical features such as a favourable location (like a port or a bridge on a river), a climate more suited to economic activity, or some natural resources. Arguably, local endowments cannot be restricted to natural features and should also encompass factors of production such as public or private capital, local institutions, and technology. More formally, this type of argument implies that in area a with endowments E_a affecting positively the productivity of labour, the wage is given by $w_a = A(E_a)$.

This (very) broad group of explanations is often at the heart of the work done by growth economists. The literature on this topic is extremely voluminous (see Durlauf and Quah, 1999, and Temple, 1999, for recent surveys).

¹See ? for some broad international comparisons and a discussion of the data sources. See also ? for a recent survey on regional inequalities in Europe.

²They usually perform some regressions at the regional level using aggregate measures of skills. In their attempts to use micro-data, Glaeser and Maré (2001) on the urban wage premium in US cities and Duranton and Monastiriotis (2002) on UK regional inequalities stand out as exceptions.

The second alternative argues that some interactions between workers or between firms may take place locally and lead to productivity gains. Interactions-based explanations have a wealth of theoretical justifications. Since Smith (1776), there is a long tradition arguing that the division of labour, which leads to productivity gains, is limited by the extent of the (local) market. Then unsurprisingly, workers in larger markets may enjoy higher wages because of a greater division of labour. It could also be the case that the matching of workers' skills with firms' needs is easier in thicker labour markets (Helsley and Strange, 1990). Input-output linkages between buyers and suppliers can also generate such agglomeration economies (Abdel-Rahman, 1988; Fujita, 1988; Fujita, Krugman, and Venables, 1999).³ Finally non-market interactions, i.e., technological externalities, and in particular those originating from human capital, may also play an important role in explaining wage differences across areas (Lucas, 1988; Rauch, 1993). A key issue in the literature is whether these benefits stem from the size of the overall market (*urbanisation economies*) or from geographic concentration at the industry level (*localisation economies*).⁴ Stated formally, these arguments imply that the mean wage in area a and industry k is given by $w_{a,k} = A(I_a, I_{a,k})$, where I_a and $I_{a,k}$ are two vectors of interaction variables to capture urbanisation and localisation economies, respectively (which we also refer to as between- and within-industry interactions).

Interaction-based explanations have received a lot of attention from urban and regional economists. Work on agglomeration economies is usually done at the aggregate level by regressing a measure of local productivity on a set of variables relating to the extent and local composition of economic activity. Results are generally supportive of the existence of both localisation and urbanisation economies.⁵

In summary, although there is a very large literature dealing with the last two of these three broad explanations (or more specific theories therein), we are not aware of any work using individual data considering all three of them in a unified framework. This is the main purpose of this paper. In our framework, we allow skills, endowments, and interactions to determine local wages. More formally, our model implies that in equilibrium the wage of worker i in area $a(i)$ and industry $k(i)$ is given by $w_i = A(E_{a(i)}, I_{a(i)}, I_{a(i),k(i)})s_i$.⁶

Such a unified framework encompassing skills-, endowments-, and interactions-based explanations is important because it should provide us with a sense of magnitudes about the importance of these three types of explanations in determining wage disparities across areas. These magnitudes are crucial to inform policy and to guide future theoretical work.

However, a unified framework imposes formidable data requirements. To deal properly with skills-based explanations, we must control for unobserved worker heterogeneity. In turn this

³The theories relying on input-output linkages and more generally on closeness to large markets differ starkly with respect to the spatial scale they consider. The focus of Abdel-Rahman (1988) and Fujita (1988) is the city whereas that of the 'new economic geography' (Fujita *et al.*, 1999) is more regional and even inter-regional. We pay some attention to these issues below.

⁴Marshall (1890) was the first to insist on the benefits of local specialisation. The distinction between localisation and urbanisation economies dates back to Hoover (1936).

⁵The main references in the literature are Sveikauskas (1975), Moomaw (1981), Henderson (1986), Sveikauskas, Gowdy, and Funk (1988), Rauch (1993), Ciccone and Hall (1996), and Henderson (2003). See Rosenthal and Strange (2004) for a survey and further references.

⁶See Rice and Venables (2003) for a related typology. In the UK case, their work focuses on the differences between explanations with respect to their aggregate implications.

requires a panel of workers. In our empirical analysis, we use a large scale panel of French workers comprising almost 9 million observations between 1976 and 1998 to account for wage disparities across French employment areas. To our knowledge, this is the first time such a large panel is used to investigate regional and urban questions.

To consider interactions and endowments, we also need to use a wealth of area level variables. However, regressing micro-data on aggregate variables generates some heteroscedasticity and leads to potentially very biased standard errors (Moulton, 1990). To deal with this issue *with panel data*, we apply the novel two-stage procedure developed by Gobillon (2003).⁷

In the first stage, we regress individual wages on some time-varying worker characteristics, a worker fixed-effect, an area-year fixed-effect, an industry fixed-effect, and a set of variables relating to the local characteristics of the industry of employment (to capture local interactions within industry). In this regression, worker fixed-effects capture all the unobserved skill differences between individuals, which do not depend on location. The area-year fixed-effects can be interpreted as local wage indices after controlling for observed and unobserved worker characteristics, industry effects, and localisation economies. Conceptually the first stage of the regression allows us to assess the importance of explanations based on differences in the productive abilities of the workforce (i.e., skills-based explanations) against those highlighting true productivity differences across areas (i.e., between industry interactions and endowments-based explanations).

Our main results for the first stage of the analysis are the following. First, worker fixed-effects alone explain around 70% of the variance of individual wages in the data. By contrast, the explanatory power of area fixed-effects and industry variables is smaller than that of workers' skills by several orders of magnitude. This suggests that the productive abilities of workers matter much more than the environment in which they are working to determine their wages. To our knowledge, it is the first time this result is established rigourously using a large panel of workers to control for individual unobserved heterogeneity.

The second key result is that differences in the skill composition of the labour force account for 40 to 50% of aggregate spatial wage disparities. This suggest that skills are of fundamental importance to explain not only wage disparities across workers but also wage disparities across areas. The reason behind this result is that workers tend to sort across locations according to their measured and unmeasured characteristics. Following up on this, the most surprising and novel result is that the correlation between the local mean of worker fixed-effects and area fixed-effects (which are computed controlling for worker fixed-effects) is large at 0.29. Hence not only do workers sort across locations according to their skills and abilities but they also do so according to local productivity. This sorting of workers across local labour markets has important implications. Most crucially, differences in the local composition of the workforce magnify true local productivity differences. In this respect, an immediate concern is that previous approaches in the literature, which estimate the importance of local interactions or local endowments but do not pay much attention to sorting, are likely to suffer from an important omitted variable problem.

⁷This two-stage empirical strategy (instead of a single-stage estimation) has additional benefits. They are explained in detail below. Note that this procedure is of independent interest since it could be applied to a wide variety of issues (e.g., panel estimation of the determinants of industry fixed-effects in labour economics or hospital fixed-effects in health economics among other potential applications).

In the second stage of the regression, we use the area fixed-effects estimated in the first stage and regress them on a set of time dummies, several variables capturing local interactions between industries, and some controls for local endowments. Conceptually, this allows us to assess the relative importance of between-industry interactions- and endowments-based explanations.

There are several concerns with this second stage. First, many endowments variables might be missing. Second, area fixed-effects and some explanatory variables capturing local interactions (such as employment density) may be simultaneously determined (Ciccone and Hall, 1996). In our analysis we pay considerable attention to these issues. To deal with them, we use a variety of panel data techniques and instrumental variables approaches.

Our findings for the second stage are the following. First, we find evidence of substantial local interactions despite the importance of sorting. Urbanisation economies (measured by the density of local employment) play the most important role, whereas endowments play at best a secondary role.

Second, controlling for sorting nearly halves standard estimates of the intensity of agglomeration economies. The second-stage OLS estimate for the elasticity of wages with respect to employment density is at 3.7%. When we repeat the same estimation with data aggregated by location and industry as typically done in the literature, our estimate for the same coefficient is at 6.3%. Similarly the average effect of employment specialisation in the same industry estimated in the first-stage is at 2.1% with individual data against 4.3% with aggregate data. Put differently, not accounting for the sorting of workers biases standard estimates of agglomeration economies by about 100%.

Third, we observe some reverse causality between high local productivity and high density. We also find that missing amenity variables play a moderate role. Taking all this into account, our preferred estimates for the elasticity of wages with respect to local employment density is around 2%. This is substantially lower than the standard estimates in the literature, which are between 4 and 8% (Rosenthal and Strange, 2004).

Fourth, after controlling for skills and interactions, residual spatial wage disparities are smaller than disparities in mean wages by a factor of around three. This result is of course consistent with a major role for skills-based explanations, a moderate role for interactions, and a weak role for endowments.

The rest of the paper is structured as follows. We first document wage disparities between French employment areas in the next section. Then, in Section 3 we propose a general model of spatial wage disparities. In Section 4, this model is estimated on individual data to assess the importance of skills-based explanations. In Sections 5 and 6, we discuss the issues relating to endowments- and interactions-based explanations and assess their importance. In Section 7, we reproduce our regressions using aggregate data. Finally some conclusions are given in Section 8.

2. Wage disparities across French employment areas

Data description

The data is extracted from the Déclarations Annuelles des Données Sociales (DADS) or Annual Social Data Declarations database. The DADS are collected by the French Institute for Statistics (INSEE) from all employers and self-employed in France for pension, benefits and tax purposes. A report must be filled by every establishment for each of its employees so that there is a unique record for each employee-establishment-year combination. The extract we use covers all employees in manufacturing and services working in France and born in October of even-numbered years. The data run from 1976 to 1998. Because of lack of sampling by INSEE, 1981, 1983 and 1990 are excluded.

The raw data contains 19,675,740 observations. For each observation, we have some basic personal data (age, gender, occupation at the one-digit level but not education), basic establishment level data (including location and firm industry at the three-digit level), number of days worked, and various measures of earnings. For consistency with the model below, we focused only on total labour costs for full-time employees deflated by the French consumer price index. We refer loosely to the real 1980 total labour cost per full working day as the wage.

Workplace location is identified at the level of employment areas (zones d'emploi). Continental France is fully covered by 341 employment areas, whose boundaries are defined on the basis of daily commuting patterns. Most employment areas correspond to a city and its catchment area or to a metropolitan area.

Although the data is of high quality, we carefully avoided a number of pitfalls. These data issues are detailed in Appendix A. After eliminating public sector workers, part-timers (for whom hours are missing), outliers, incomplete observations, and various industries for which there are some coding problems, we ended up with 8,826,422 observations. For reasons of computational tractability, we keep only six points in time, one every four years (1976, 1980, 1984, 1988, 1992, and 1996) and 2,664,474 observations when estimating the model on individual data.⁸ Appendix A and Abowd *et al.* (1999) provide further details on the data and some background information on wage setting in France.

Wage disparities across French employment areas

In this sub-section we briefly document the extent and persistence of wage disparities between employment areas in France. To do this, we consider the mean wage in each French employment area for each year between 1976 and 1996.

Typically, wages in and around Paris are around 15% higher than those of other large French cities such as Lyon or Marseille, 35% higher than those of mid-sized French cities, and 60% higher than those of predominantly rural employment areas. To be more systematic, we can also compute a series of inequality measures between employment areas. The ratio of the highest average to the lowest across all French employment areas remains between 1.62 and 1.88 during the 1976 – 1996

⁸Note that these dates cover different phases of the business cycle to avoid any bias in the estimates.

Table 1. Some simple correlations
Mean local wage in 1998 ($\log w_{a,1998}$) as a function of:

	(1)	(2)	(3)	(4)	(5)	(6)
	$\log w_{a,1997}$	$\log w_{a,1988}$	$\log Den_{a,1998}$	$\log Emp_{a,1998}$	$\log Div_{a,1998}$	$Skil_{a,1998}$
Intercept	0.128* (0.072)	0.424*** (0.132)	5.720** (0.014)	5.147*** (0.025)	5.329*** (0.037)	5.352*** (0.006)
Coefficient	0.976*** (0.013)	0.943*** (0.025)	0.049*** (0.003)	0.049*** (0.004)	0.047*** (0.012)	1.763*** (0.085)
R ²	0.94	0.81	0.51	0.34	0.04	0.56

$w_{a,t}$ is the wage in employment area a and year t ; $Den_{a,t}$ is the density of employment; $Emp_{a,t}$ is total employment; $Div_{a,t}$ is the diversity of employment as measured by an inverse-Herfindhal index, $Div_{a,t} = Emp_{a,t}^2 / \sum_k Emp_{a,k,t}^2$ where subscript k denotes the industries; and $Skil_{a,t}$ is the employment share of professionals. Standard error between brackets. *: significant at 10%, **: significant at 5%, and ***: significant at 1%.

period. The ratio of the ninth to the first decile is between 1.19 and 1.23. Finally, the coefficient of variation also remains between 0.08 and 0.09. All this points to rather large and persistent wage disparities between French employment areas. Reporting the correlation between local wages over time, columns 1 and 2 of Table 1 provide further evidence of the persistence of spatial wage disparities. Even with a 10-year lag, this correlation is larger than 0.9.

Table 1, columns 3 – 6 reports ordinary least squares (OLS) estimates suggesting that local wages are strongly linked to the structural attributes of their employment area. Column 3 regresses the log of the mean local wage in 1998 on the log of the local density of employment in the same year. The coefficient indicates that a doubling of the density yields a 4.9% rise in wage. The explanatory power of this single variable is very strong since the R² is 51%. Similar results are obtained in column 4 when using total employment instead of density as explanatory variable. In column 5, local wages are regressed on an index of industrial diversity. The effect of this variable is also significant but its explanatory power is much weaker. Finally, regressing local wages in column 6 on the share of workers in professional occupations also yields very good results. For this variable, a one point increase is associated with an increase of 1.8% in the mean local wages. This is a large coefficient and the associated R² is also large at 56%.

In summary, denser, more populous, and more educated employment areas seem to command on average a higher wage. The objective of the rest of the paper is to assess the robustness of these basic results and uncover the determinants of spatial wage disparities.

3. Theory and estimation

The model

Consider a competitive representative firm operating in employment area a and industry k in year t . Its profit is:

$$\pi_{a,k,t} = p_{a,k,t} y_{a,k,t} - \sum_{i \in (a,k,t)} w_{i,t} \ell_{i,t} - r_{a,k,t} z_{a,k,t}, \quad (1)$$

where $p_{a,k,t}$ denotes its price and $y_{a,k,t}$ its output. For any given worker i employed in this firm in year t , $w_{i,t}$ and $\ell_{i,t}$ are the daily wage and the number of working days, respectively. Finally,

$z_{a,k,t}$ represents the other factors of production and $r_{a,k,t}$ their price. Note that this specification allows for markets to be either segmented or integrated (when $p_{a,k,t} = p_{k,t}$ and $r_{a,k,t} = r_{k,t}$) for both output and non-labour inputs. Output is Cobb-Douglas in effective labour and the other factors of production:

$$y_{a,k,t} = A_{a,k,t} \left(\sum_{i \in (a,k,t)} s_{i,t} \ell_{i,t} \right)^b (z_{a,k,t})^{1-b}, \quad (2)$$

where the coefficient b is such that $0 < b \leq 1$, $s_{i,t}$ denotes the skills of worker i in year t , and $A_{a,k,t}$ is the total factor productivity in (a,k,t) . At the competitive equilibrium, the wage received in year t by worker i employed in employment area $a(i,t)$ and industry $k(i,t)$ is equal to her marginal product:

$$w_{i,t} = b p_{a(i,t),k(i,t),t} A_{a(i,t),k(i,t),t} \left(\frac{z_{a(i,t),k(i,t),t}}{\sum_{i \in (a,k,t)} s_{i,t} \ell_{i,t}} \right)^{1-b} s_{i,t}. \quad (3)$$

Using the first-order condition for profit maximisation with respect to the other factors and inserting it in equation (3) yields:

$$\begin{aligned} w_{i,t} &= b(1-b)^{\frac{(1-b)}{b}} \left(p_{a(i,t),k(i,t),t} \frac{A_{a(i,t),k(i,t),t}}{(r_{a(i,t),k(i,t),t})^{1-b}} \right)^{\frac{1}{b}} s_{i,t} \\ &= B_{a(i,t),k(i,t),t} s_{i,t}. \end{aligned} \quad (4)$$

As discussed in the Introduction, wage differences across areas can reflect differences in individual skills or alternatively they can also reflect true productivity differences caused either by differences in non-human endowments or by local interactions. Skills are captured by the last term, $s_{i,t}$, in equation (4) whereas the other two explanations enter the term $B_{a,k,t}$ in equation (4). As made clear by this latter term, 'true productivity differences' can work through total factor productivity, $A_{a,k,t}$, or through the price of outputs, $p_{a,k,t}$, or even through the price of non-labour inputs, $r_{a,k,t}$. This implies that we cannot identify price and technology effects separately.⁹

To understand this point better, consider for instance an employment area a , which is located in a mountainous region, and an industry k . Mountains may have a negative effect on wages in (a,k) because shipping the final output of the industry to the main consumer markets is more expensive, which lowers *f.o.b.* prices. Mountains may have another direct negative effect on wages in (a,k) because operating a plant is more difficult when land is not flat. Finally mountains may have a positive effect on wages because some raw materials such as wood may be more readily available. In this toy example, the first effect works through $p_{a,k,t}$, the second through $A_{a,k,t}$, whereas the third goes through $r_{a,k,t}$. With our approach, we can only estimate the overall effect of local characteristics, the presence of mountains say, in area a and industry k . In other words, we can

⁹Separating direct technology effects from those of local costs and prices would require very detailed firm level information about the prices and quantities of final goods, and the costs and quantities of inputs. This is of course well beyond the scope of this paper. Another issue is that the model also assumes perfect substitutability across workers in the same industry and employment area. This implies that any complementarity between for instance high and low-skill workers in the local production function will be identified together with local technological externalities (Ciccone and Peri, 2002).

identify the determinants of spatial wage disparities (i.e., endowments, interactions, and skills) but not the exact channel through which agglomeration economies percolate.¹⁰

Note further that some local characteristics like employment density may have a positive effect on $B_{a,k,t}$ (e.g., agglomeration economies) as well as a negative effect (e.g., congestion). However we are not able to identify these effects separately. We can only estimate the overall effect of a variable. In other words, the estimated effects of interaction variables include both agglomeration economies and any decreasing returns due to crowding.

A micro-econometric specification

To take equation (4) to the data, we need a specification for both the skill term, $s_{i,t}$, and the 'local industry productivity' term, $B_{a,k,t}$. Assume first that the skills of worker i are given by:

$$\log s_{i,t} = X_{i,t}\varphi + \delta_i + \epsilon_{i,t}, \quad (5)$$

where $X_{i,t}$ is a vector of time-varying characteristics, δ_i is a worker fixed-effect, and $\epsilon_{i,t}$ is a measurement error. The error terms are assumed to be i.i.d. across periods and workers.

Turning to $B_{a,k,t}$, which reflects true productivity differences in equation (4), we assume that it is given by:

$$\log B_{a,k,t} = \beta_{a,t} + \mu_{k,t} + I_{a,k,t}\gamma_k, \quad (6)$$

where $\beta_{a,t}$ is an area-year fixed-effect, $\mu_{k,t}$ is an industry-year fixed-effect, and γ_k is the vector of coefficients associated with $I_{a,k,t}$, the vector of interactions variables for each area-industry-year. For within-industry interactions we follow the literature and use the log of the share of the industry in local employment, the log number of industry establishments and the log share of workers in professional occupations in the industry as explanatory variables.¹¹

Note that $B_{a,k,t}$ could instead, and more simply, be estimated using a fixed-effect for each area, industry, and year. There would be three problems with doing this. First, it would force us to include more than 200,000 fixed-effects in the model (341 employment areas \times 99 industries \times 6 years). These would come in addition to the worker fixed-effects introduced in equation (5). Estimating such a large number of worker and area-industry fixed-effects is computationally too demanding.¹² Furthermore, many of these fixed-effects would be estimated with a very small number of workers (if at all). This would raise some problems of both identification and statistical significance.

¹⁰The theoretical micro-foundations of increasing returns generated by local interactions are reviewed in Duranton and Puga (2004). A crucial feature in this literature is that many different microeconomic mechanisms yield similar reduced forms. Following this, Rosenthal and Strange (2004) divide their review of the empirical literature into two main parts: the measurement of agglomeration economies on the one hand and the identification of the mechanisms at stake on the other. The two exercises are sufficiently different (and difficult) to justify separate approaches.

¹¹The local share in employment and the number of establishments are standard variables appearing in most models of localisation economies (Duranton and Puga, 2004). The share of professionals in the industry is a proxy for the average education locally in the industry. This should capture the external effects of human capital in the local industry in the spirit of Lucas (1988) and Rauch (1993).

¹²Non-standard techniques, such as those developed by Abowd *et al.* (1999) would allow us to estimate both worker and area-industry-year fixed-effects. For reasons made clear below, these techniques are not appropriate here.

Combining equations (4), (5), and (6) yields:

$$\log w_{i,t} = \beta_{a(i,t),t} + \mu_{k(i,t),t} + I_{a(i,t),k(i,t),t} \gamma_{k(i,t)} + X_{i,t} \varphi + \delta_i + \epsilon_{i,t}. \quad (7)$$

In equation (7) the interpretation of $I_{a,k,t} \gamma_k$ and $X_{i,t} \varphi$ is problematic. For instance, an industry may employ younger workers. If wages increase with age, this industry will pay lower wages all else equal. We want to think of such systematic industry component as being part of the ‘industry effect’. To do this we centre $I_{a(i,t),k(i,t),t}$ and $X_{i,t}$ around their industry mean. The systematic industry components in $I_{a,k,t} \gamma_k$ and $X_{i,t} \varphi$ are added to the industry fixed-effect to form a total industry effect. For tractability, we also need to limit the number of coefficients in the model and assume that the time trend is the same for all industries so that this total industry effect can be decomposed into an industry fixed-effect and a year effect (which can be normalised to zero for all years since the temporal evolution is also captured by the area-year fixed-effect).¹³ The final specification for the first stage of the analysis is thus:

$$\log w_{i,t} = \beta_{a(i,t),t} + \mu_{k(i,t)} + \tilde{I}_{a(i,t),k(i,t),t} \gamma_{k(i,t)} + \tilde{X}_{i,t} \varphi + \delta_i + \epsilon_{i,t}. \quad (8)$$

where $\tilde{I}_{a(i,t),k(i,t),t}$ is the centred vector of within-industry interactions variables and $\tilde{X}_{i,t}$ is the centred mean of individual time-varying characteristics.

Equation (8) corresponds to an inverse labour demand equation.¹⁴ We estimate the wages of workers (expressed in constant 1980 francs) as a function of their observed and unobserved characteristics (age and its square plus a worker fixed-effect) and of the area and industry in which they are employed. Rather than a full set of area-industry-year fixed-effects, the working environment is characterised by an area-year fixed-effect, an industry fixed-effect, and some observable local characteristics of the industry (log share of the industry in local employment, log number of establishments and the log share of professionals in the industry).

This estimation allows us to identify separately the effects of ‘people’ (skills-based explanations) versus those of ‘places’ (endowments- and interactions-based explanations). It also allows us to give the respective explanatory power of the effects of skills ($\tilde{X}_{i,t} \varphi + \delta_i$), of within-industry interactions ($\tilde{I}_{a,k,t} \gamma_k$), and the joint explanatory power of endowments and between-industry interactions ($\beta_{a,t}$). Next, the second stage of the estimation, which uses $\beta_{a,t}$ as dependent variable, allows us to assess separately the explanatory power of between-industry interactions and endowments. It is presented in detail in Section 5.

¹³Formally, the effects of within-industry interactions, $I_{a,k,t} \gamma_k$, can be decomposed into an industry specific component independent of location, $I_{\cdot,k,t} \gamma_k$, and a component net of national industry effects, $\tilde{I}_{a,k,t} \gamma_k \equiv (I_{a,k,t} - I_{\cdot,k,t}) \gamma_k$ where $I_{\cdot,k,t}$ is the mean of the $I_{a,k,t}$ weighted by local employment in the industry ($I_{\cdot,k,t} = \frac{1}{N_{k,t}} \sum_{a \in (k,t)} N_{a,k,t} I_{a,k,t}$ where $N_{a,k,t}$ is employment in area a , industry k and year t and $N_{k,t}$ is total employment in industry k in year t). Similarly the effect of age can be decomposed into an industry specific component $X_{\cdot,k(i,t),t} \varphi$ and a component net of national industry effect $\tilde{X}_{i,t} \varphi \equiv (X_{i,t} - X_{\cdot,k(i,t),t}) \varphi$. The total industry effect is thus $\mu_{k,t} + I_{\cdot,k,t} \gamma_k + X_{\cdot,k,t} \varphi$. This consists of the industry effect as defined above, plus a national average industry interaction effect and a national average composition effect (in terms of workers’ observable characteristics). Then we assume: $\mu_{k,t} + I_{\cdot,k,t} \gamma_k + X_{\cdot,k,t} \varphi = \mu_k + \rho_t$. Finally, since it is not possible to identify ρ_t and $\beta_{a,t}$ separately, we normalise ρ_t to zero for all years.

¹⁴A competitive wage-setting mechanism is assumed. Any imperfect competition framework where the wage is a mark-up on marginal productivity would lead to similar results since in a log specification this mark-up would enter the constant or the industry fixed-effects if such mark-ups vary between industries. In France, there is some empirical support for the competitive/fixed-mark-up assumption (see Appendix A and Abowd *et al.*, 1999).

Identification of industry and area fixed-effects

Regarding industry fixed-effects, μ_k , note that a worker remaining in the same industry is such that $\mu_{k(i,t)} = \mu_{k(i)}$. In this case, the industry effect cannot be distinguished from the worker effect δ_i . Hence, industry fixed-effects can only be identified with workers changing industries. When taking the first-difference to eliminate the worker effect, any worker moving from, say, industry 1 to industry 2 allows us to identify $\mu_1 - \mu_2$. More generally, the identification of all industry fixed-effects requires all industries to be 'connected' with each other (at least indirectly) through worker flows (see also Abowd *et al.*, 1999). Given the amount of data we have, this condition is easily met.

Turning to area fixed-effects, recall first that workers staying in the same area (and industry) are informative with respect to the variation of the area fixed-effect over time. For instance, workers staying in employment area 1 and 2 between t and $t + 1$ allow us to identify $\beta_{1,t+1} - \beta_{1,t}$ and $\beta_{2,t+1} - \beta_{2,t}$, respectively. Then, workers moving from one area to the other between t and $t + 1$ enable us to identify $\beta_{2,t+1} - \beta_{1,t}$. Using the decomposition $\beta_{2,t+1} - \beta_{1,t} = (\beta_{2,t+1} - \beta_{2,t}) + (\beta_{2,t} - \beta_{1,t})$ and since the first term is already identified thanks to the stayers, movers thus make it possible to identify $\beta_{2,t} - \beta_{1,t}$. By the same token, the same migrants also allow the identification of $\beta_{2,t+1} - \beta_{1,t+1}$ thanks to the decomposition $\beta_{2,t+1} - \beta_{1,t} = (\beta_{2,t+1} - \beta_{1,t+1}) + (\beta_{1,t+1} - \beta_{1,t})$. More generally, to identify area fixed-effects, it is enough to have (i) some workers staying in each of the employment areas between any two consecutive dates and (ii) no area or group of areas with no worker flow to the rest of the country. Again, we have enough observations for these two conditions to be met.

Since the area fixed-effects, just like the industry fixed-effects, are identified only relative to each other, some identification constraints are necessary. With respect to area fixed-effects, we set the coefficient for Central Paris in 1980 to zero. Turning to industries, we take the first industry (meat processing) as the reference: $\mu_1 = 0$.

Estimation method and estimation issues

Our very large number of observations (with a very large number of worker fixed-effects) restricts us to a simple estimation procedure for this first stage. We estimate equation (8) using the *within* estimator. This allows us to compute the coefficients on all time-varying variables and thus recover all parameters except the worker fixed-effects. Only workers appearing at least twice in the panel contribute to the estimation. This leaves us with 653,169 workers representing 2,221,156 observations. Next, we can recover an estimator of each worker fixed-effect by computing each mean of the prediction error. By the Frish-Waugh theorem, this estimator is the OLS estimator.

We have to impose restrictive assumptions when using these estimation methods to obtain unbiased estimators of the parameters. First, the choice of area and industry is assumed to be strictly exogenous. This assumption holds when workers choose their industry and employment area on the basis of their expected wages.¹⁵ This exogeneity is violated when workers make their employment choice on the basis of the actual wages offered by the firms at date t . In this case,

¹⁵For example, this corresponds to the case of a worker migrating to a given area and then deciding to search for a job in a given industry. Alternatively, this is also the case when part of the wage paid by firms is a bonus that is unknown ex-ante.

individual shocks ($\epsilon_{i,t}$) are correlated with the chosen location and industry and estimates are biased. However, as shown in Appendix B (see also Topel, 1986), the bias is much reduced in a dynamic context when workers make their employment decision on the basis of both current and future (expected) wages. Furthermore this concern becomes even less important when workplace and industry decisions are also driven by other factors such as idiosyncratic preferences. Given the difficulty of this problem with such a large number of areas, we leave it aside here and hope to make progress in further work.

A similar type of concern also arises with the characteristics of the local industries in $I_{a,k,t}$. As discussed by Ciccone and Hall (1996) and Ciccone (2002), some local characteristics like a high level of specialisation in an industry could be endogenous to high wages in this industry. We leave these concerns aside here on the ground that these variables only have a small explanatory power (see below).

According to Abowd *et al.* (1999) a wage equation with industry fixed-effects should also contain establishment fixed-effects. This is because these fixed-effects may be correlated with industry fixed-effects. This also applies to area fixed-effects. Such a correlation would bias the estimates when establishment fixed-effects are omitted. However the method developed by Abowd *et al.* (1999) to deal with large scale matched employer-employee data (using both worker and plant fixed-effects) would not allow us to compute the standard deviations for the estimated area fixed-effects that are necessary to perform the second stage of the estimation correctly. This approach would also lack theoretical foundations since area fixed-effects would then have to be computed by calculating a weighted average of establishment fixed-effects by location. A final problem with this alternative approach is that establishment fixed-effects are constrained by the estimation to be constant over time. The resulting area fixed-effects constructed by aggregating time-invariant establishment fixed-effects can then evolve only through the entry and exit of establishments and internal changes in employment and not by changes in the strength of agglomeration economies.

4. Skills and sorting across employment areas using individual data

This section presents the results for the *within* estimation of equation (8).¹⁶ We first present a variance analysis and the results about sorting before commenting on the coefficients.

The importance of workers' skills

To get a first pass at assessing the relative importance of workers' skills against true productivity differences, we can estimate wages as a function of worker fixed-effects *only*. This yields an R^2 of about 70%. When using only area-year fixed-effects, the R^2 is about 15%. A similar R^2 is achieved with only industry and year effects. When we allow for area-year fixed-effects, industry fixed-effects and local industry effects, the R^2 is 31%. When adding worker fixed-effects in the complete baseline specification, it jumps to 80%. These results suggest that skills are of funda-

¹⁶Recall that the explanatory variables are the area-year fixed-effects, the industry fixed-effects, the worker fixed-effects, the worker's age and its square, the log of the share of the industry in local employment (i.e., specialisation), the log of the number of establishments and the log share of professionals in the industry.

Table 2. Summary statistics for the variance decomposition — estimation of equation (8)

Effect of	Std dev	Simple correlation with:		
		$\log w$	δ	$\beta - \theta$
log real wage ($\log w$)	0.367	1.00	0.78	0.26
residuals (ϵ)	0.166	0.45	0.00	0.00
worker effects ($\delta + X\varphi$)	0.294	0.80	0.98	0.09
worker fixed-effects (δ)	0.284	0.78	1.00	0.10
age ($X\varphi$)	0.058	0.23	0.08	0.00
industry fixed-effects (μ)	0.043	0.25	0.16	0.05
within-industry interactions ($\tilde{I}_k\gamma_k$)	0.024	-0.01	0.00	-0.45
within-industry share of professionals	0.011	0.16	0.12	0.29
within-industry establishments	0.019	-0.13	-0.08	-0.62
specialisation	0.017	0.03	0.02	-0.13
area fixed-effects (β)	0.140	0.34	-0.05	0.55
de-trended area fixed-effects ($\beta - \theta$)	0.065	0.26	0.10	1.00
time (θ)	0.118	0.26	-0.11	0.10

2,221,156 observations. All correlations between the effects that are not orthogonal by definition are significant at 1%. The effect of within-industry share of professionals is that of the log of the share of professional times its coefficient (in vector γ_k). The effect of within-industry establishments is that of the log of the number of establishments times its coefficient. The effect of specialisation is that of the log of the industry share in employment times its coefficient. Area fixed-effects are de-trended using the time fixed-effects (θ) estimated in the second stage.

mental importance and play a much greater role than the local environment in the determination of individual wages.

These results are confirmed when we perform a more complete variance analysis as in Abowd *et al.* (1999). Table 2 shows the explanatory power of the different variables for the baseline regression. For each variable or group of variables, the Table reports the standard deviation of *their effect* and their correlation with wages, the worker fixed-effects and the de-trended area fixed-effects.

To construct this Table, we computed for each observation *the effect of each variable* by multiplying its coefficient by its value for the observation. For instance, consider worker i in (a,k,t) . The effect of specialisation is equal to the estimated coefficient on this variable for industry k times the specialisation of area a in this industry. For a group of variables, the sum of the effects is computed. Then, the variability of the effect of each variable across workers can be calculated. When the standard deviation of the effect of a variable is large and when it is highly correlated with wages, this variable has a large explanatory power. When on the contrary the standard deviation of the effect and its correlation with wages are small, this variable explains only a small fraction of the variations of wages.

Worker fixed-effects have by far the largest explanatory power. Their standard deviation, at 0.284, is close to that of wages at 0.364 and the correlation between worker fixed-effects and wages is very high at 0.78. For no other variable, or group of variables, are the standard deviation and the correlation with log wages as high. When looking at the effects of observable worker characteristics, it is worth noting that age and its square also have a moderate explanatory power with a standard deviation at 0.058 and a correlation with log wages at 0.23. Altogether, with a standard deviation of 0.294 and a correlation of 0.80 with wages, the combined effect of individual

Table 3. Spatial wage disparities, 1976 – 1996

	Mean wage	Net wage
Max/Min	1.88	1.51
P90/ P10	1.23	1.14
P75/ P25	1.12	1.06
Coefficient of variation	0.09	0.05

Mean wage refers to the de-trended mean wage by employment area. Net wages are calculated as in equation (9). Figures are the average over 1976-1998 of each inequality measure. Max/Min is the ratio between the two extremes of the distribution. P90/ P10 is the ratio of the ninth to the first decile. P75/ P25 is the interquartile ratio.

observed and unobserved characteristics is of overwhelming importance.¹⁷

Turning to within-industry interactions, their explanatory power is very small. The standard deviation of the effect of all within-industry interaction variables together is less than a tenth of that of worker fixed-effects. Furthermore, the correlation between log wages and the effect of within-industry interactions is close to zero. Within this group of variables, neither the share of professionals, the number of establishments nor specialisation particularly stands out.¹⁸

Finally, the explanatory power of area-year fixed-effects is substantial, albeit much less so than that of worker fixed-effects. Because wages increased everywhere in real terms between 1976 and 1996, a good fraction of the area fixed-effects is explained by the time trend over the period. After taking away this trend however, area fixed-effects still have an explanatory power more important than that of industry, age, or within-industry interactions. Although this result was to be expected, this is rather interesting in light of the small amount of attention location factors have received so far in the labour literature relative to industry and age.

Spatial wage disparities and sorting

To evaluate the importance of workers' skills on spatial wage disparities, we can also study the variations of a wage index net of worker fixed-effects and industry effects. This 'net wage' is computed from the results of the first-stage regression. It corresponds to the local wage obtained by an 'average' worker in an 'average' industry. We can define such an index $w_{\text{net},a,t}$, which we refer to as the net wage, in the following way:

$$\log w_{\text{net},a,t} \equiv W_t + \hat{\beta}_{a,t}, \quad (9)$$

where W_t is a normalising trend such that $w_{\text{net},a,t}$ can be interpreted as a wage (net of the effects of individual skills).¹⁹

These net wages can then be compared with the real mean wages per area computed in Section 2. Table 3 compares systematically disparities in mean and net wages. Depending on the inequality

¹⁷By contrast, industry fixed-effects are rather small. The relative magnitudes of industry and worker fixed-effects are close to those found by Abowd *et al.* (1999) despite the differences in the specification of the econometric model.

¹⁸Note however that the very small effect of specialisation on wages does not mean that this variable is not a major driver behind the composition of economic activity in cities.

¹⁹Formally, we have $W_t \equiv \frac{1}{N} \sum_{j=1}^K N_j \hat{\mu}_j + \frac{1}{N_{t_0}} \sum_{i \in t_0} \hat{\delta}_i + \frac{1}{N_{t_0}} \sum_{m=1}^Z N_{m,t_0} \hat{\beta}_{m,t_0} - \frac{1}{Z} \sum_{m=1}^Z \hat{\beta}_{m,t}$ where $t_0 = 1980$ and Z is the number of areas.

Table 4. Summary statistics for the coefficients estimated in equation (8)

Variable	Number of coefficients	Percentage > 0 at 5%	Percentage < 0 at 5%	P90/P10
area fixed-effects (de-trended)	2046	10%	78%	0.16
industry fixed-effects	99	58%	33%	0.11
age	1	100%	0%	—
squared age	1	0%	100%	—
specialisation	99	95%	0%	0.02
share of professionals	99	81%	3%	0.20
industry establishments	99	1%	85%	0.02

For area fixed-effects, significance is calculated relative to the weighted national mean for the period. For industry fixed-effects, significance is calculated relative to the weighted national mean. Max/Min is the ratio between the two extremes of the distribution. P90/ P10 is the ratio of the ninth to the first decile.

measure taken, disparities in net wages may be as low as half of those in mean wages. Put differently, workers' skills explain 40 to 50% of spatial wage disparities.²⁰

This result is caused by a strong sorting pattern whereby workers with high fixed-effects tend to live in areas where other workers with high fixed-effects also live (and conversely for low fixed-effect workers). To go further on this issue, it is interesting to correlate the average worker fixed-effects within each areas with area fixed-effects. The correlation between the two is large at 0.29. In other words, areas where workers with high individual fixed-effects work are also areas where the productivity of labour (after controlling for skills) is high. Hence workers sort not only along skill lines but also according to area characteristics. An immediate implication is that large spatial wage disparities reflect true productivity differences across areas that are magnified by the sorting of workers by skills.

Analysis of the coefficients

Table 4 reports some summary statistics regarding the coefficients computed in the first stage of the estimation.²¹ Note first that 88% of the area fixed-effects differ significantly from the national mean (weighted for the period). Moreover, the distribution is skewed since only 10% of these area fixed-effects are significantly higher than the mean whereas 78% are significantly lower. This is because a few populous employment areas (Paris, its suburbs, and other large French cities) offer significantly higher wages than the national mean. Industry fixed-effects are also significant since 90% of them differ significantly from their weighted national mean.

²⁰Our computations do not account for the differences in local industry characteristics. However since the effects of the latter are very small, their contribution to explaining wage disparities is marginal.

²¹Our identification constraints ($\mu_1 = 0$ and $\beta_{\text{Paris},1980} = 0$) imply that standard Student's tests about the significance of the industry and area effects with respect to 0 are not very informative because they depend on the choice of references. We instead test the significance of the coefficients with respect to their weighted industry mean or their weighted area mean for a given year. That is, we test the constraints: $\mu_k = \frac{1}{N} \sum_{j=1}^K N_j \mu_j$ and $\beta_{a,t} = \frac{1}{N_t} \sum_{j=1}^Z N_{j,t} \beta_{j,t}$, where $N_{j,t}$ is the number of workers in employment area j in period t , N_t denotes the total number of workers in year t , N_j is the total number of workers in industry j across all years, K is the number of industries, and Z is the number of employment areas. These tests can easily be implemented once the coefficients and their standard errors are estimated. Directly constraining the mean of all area or industry fixed-effects to zero in the estimation would have been computationally too demanding.

In line with previous findings in the literature, we find that most specialisation coefficients are positive and significant. The average for all industries is at 2.1%, which is at the lower bound of the estimates found in the literature (Henderson, 1986; Rosenthal and Strange, 2004). For a typical industry, a doubling of its share in local employment yields a 2.1% increase in its local wages. The largest specialisation coefficients are found for business services (3.6%) and for two high-tech industries, namely medical instruments (3.9%) and artificial fibres (4.3%). At the other end of the spectrum, the five industries with a coefficient not significantly different from zero are oil refinery, air transport, tobacco, production of weapons and bullets, and production of steel. Given the reliance of most of these industries on localised natural advantage (or some localised infrastructure), these results are not very surprising. The average coefficient on the share of professionals across industries is quite large at 11.8% in light with the usual findings in the literature on human capital externalities (see Rauch, 1993, and his followers). Finally, the coefficient on the log number of industry establishments is on average at -1.4% . This coefficient is highest in industries such as machine tools and various instrument industries that produce very differentiated goods. The smallest coefficients are obtained in industries where efficient plant size is expected to be very large like various extractive industries, naval construction, and energy or water utilities.

5. The determinants of area fixed-effects: estimation

So far we have assessed the relative importance of ‘people’ versus ‘places’ to explain spatial wage disparities. Although workers’ skills play a fundamental role, the explanatory power of true productivity differences is not negligible. The objective of the second stage of the estimation is to assess the relative importance of endowments and between-industry interactions in explaining the area-year fixed-effects.

Specification

The area fixed-effects estimated in equation (8) are assumed to be a function of a year fixed-effect, of local between-industry variables (employment density, land area, and diversity) and of some endowments variables. The econometric specification is:

$$\beta_{a,t} = \omega_0 + \theta_t + I_{a,t}\gamma + E_{a,t}\alpha + v_{a,t}. \quad (10)$$

where the θ_t are time dummies and α is a vector of coefficients associated with the endowments variables, $E_{a,t}$. γ is the vector of coefficients associated with local *between-industry* interactions, $I_{a,t}$. The error terms $v_{a,t}$ that reflect local technology shocks are assumed to be i.i.d across areas and periods.²² Finally, we take 1980 as reference year so that the coefficient for this year is set to zero.

²²Note that endowments, between-industry interactions, and productivity shocks are assumed to affect all industries symmetrically. As already stated above, a full (a,k,t) analysis would bring the number of estimated effects well above our computational limits.

Estimation method

Note that equations (8) and (10) constitute the full econometric specification. We speak of a two-stage estimation because in equation (10), the second stage, we use as dependent variable the area fixed-effects estimated in equation (8), the first stage. An alternative would be to perform only a single-stage estimation and use all the explanatory variables at once.

Such single-stage estimation would be problematic for two reasons. First, as shown by Moulton (1990), using aggregate variables in individual level estimations can create very large biases in the standard error for the coefficients on aggregate explanatory variables. This comes from the existence of unobserved local effects $v_{a,t}$. Our approach mostly avoids this pitfall.²³ Second, with a single-stage estimation, we cannot assess the importance of the local shocks, $v_{a,t}$, in the error term.²⁴ Distinguishing local shocks from purely idiosyncratic shocks at the worker level is important since we lack many endowment variables. Their effect will consequently enter the local shock. As robustness check, we nonetheless ran a single-stage estimation and found qualitatively similar results (see Section 6). We can now turn to the different estimation issues.

Heteroscedasticity

In the estimation of equation (10), note first that the true value of the dependent variable, $\beta_{a,t}$, is unknown. We use instead the consistent and unbiased estimators $\hat{\beta}_{a,t}$ provided by the first-stage results. However, the fixed-effects for areas with few workers are less precisely estimated than those for areas with many workers. Thus, the use of $\hat{\beta}_{a,t}$ as dependent variable introduces some heteroscedasticity through measurement errors. This can be dealt with by computing a feasible generalised least-square (FGLS) estimator. The procedure is detailed in Appendix C.

As shown below, the second stage results using the FGLS correction are very close to those obtained with simpler estimation techniques without any correction. This shows that the effects of the measurement errors on the coefficients estimated at the second stage are negligible.²⁵ Consequently, when dealing with endogeneity and omitted variables problems, we will ignore them to keep the econometrics reasonably simple.

Endogeneity

Some local characteristics are likely to be endogenous to local wages. For instance, employment areas receiving a positive technology shock may attract migrants and thus lead to a positive correlation between the second-stage residual and the density of employment. In this particular case, reverse-causality is going to bias the estimates upwards. Hence, reverse-causality is potentially a serious concern for the second stage of the estimation.

²³Standard robust clustering methods cannot be applied here because (i) the model is projected in the *within* dimension and (ii) workers can move between areas. However in the two-stage estimation there may be still some heteroscedasticity generated by unobservable local industry effects in the first stage. However, these variables play a minor role in the estimation while both area and industry fixed-effects capture most of the variation.

²⁴This is because when we project the model into the *within* dimension, the structure of the covariance matrix of error terms becomes very complex. Thus it is impossible to retrieve an estimator for the variance of $v_{a,t}$.

²⁵This is because we have a very large number of observations with many stayers and large flows of movers between areas. This allows us to estimate the area-year fixed-effects very precisely.

This reverse-causality may be caused by 'contemporaneous' local shocks. Considering that these shocks did not have any effect on the distribution of the population decades ago, we can instrument employment density between 1976 and 1998 by the density of population in 1936. Our lag is not as long as that of Ciccone and Hall (1996) in their analysis of the US (between 40 and 60 years instead of more than a century in their case). However, it can be argued that World-War II was in France a large exogenous shock that makes up for the shorter time period.²⁶

Missing variables

The last key estimation concern regards missing variables. Unlike the endogeneity of explanatory variables, missing variables have received only scant attention in the literature. Before outlining how we are dealing with them, we would like to analyse this estimation issue in some detail.

We distinguish three types of missing variables: those related to local (productive) endowments, those related to consumption amenities and those related to the spatial scale of interactions (i.e., market access / interactions between employment areas).

Productive endowments can increase wages through one of the three channels highlighted above (lower exporting costs, cheaper supplies, or higher productivity). There are many possible endowments that may work through these channels. One can think about airports, high-speed train lines, a favourable climate, closeness to a navigable river or a deep-sea harbour, etc. Gathering data for a complete list of endowments is a task much beyond the scope of this paper. Moreover, using a complete set of endowments would raise serious endogeneity concerns. For instance train stations, highways, or airports are likely to be endogenous.

In absence of a complete set of endowments variables, we expect much of the effects of endowments to be captured by the error term. We could treat missing endowments as a random effect entering the residual. However, as a source of higher wages, these endowments are also likely to attract more workers in the area. In this case, failure to control for endowments will bias upwards the estimate of the effect of employment density on productivity.

To deal with this omitted variables problem, we consider three different approaches. Some of these endowments such as climate have a structural (or permanent) nature. To control for any of them, we can estimate equation (10) with time-invariant local fixed-effects. Other endowments such as high-speed train lines are of a less permanent nature. They are akin to the contemporaneous shocks described above. To deal with them, we can again instrument employment density by its past values. Finally, for some endowments, we can use some direct controls. Our set of endowment variables is rather small however. It contains the fraction of the population in an employment area living in a municipality with the following location attributes: seashore, lakes, mountains and cultural or architectural heritage.

Turning to amenities, equation (4) shows that the price of non-labour inputs matters in the determination of local wages. Then it is worth noting that the price of some non-labour factor (such

²⁶Ciccone (2002) instruments density by land area. We cannot do this here as land area is already used as a regressor. *Not* using land area as a regressor implies the absence of scale effects (since the coefficient on land area captures the effect of an increase in the size of the employment area keeping density constant). In the data, we do find evidence of such scale effects.

as land) may not be solely determined by variables playing a direct role in the production function. As highlighted first by Roback (1982), better consumption amenities (i.e., amenities unrelated to production) increase the willingness of consumers to pay for land and thus imply higher local land rents. As a result, firms use less land. In turn, this lowers the marginal product of labour when land and labour are imperfect substitutes in the production function. Put differently, wages may capitalise the effect of non-production variables.

This is in itself not an issue for our purpose. This missing variable problem would only imply more noisy estimates for the wage effects since observationally identical employment areas end up paying different wages. It becomes an issue when consumption amenities are positively correlated with a variable of interest like employment density because, as shown by Wheaton and Lewis (2002), this introduces a negative correlation between this variable and the residuals.²⁷ Because of this, the estimated effect of employment density is potentially biased downwards. However, just as with productive endowments, we can use instrumental variables, time-invariant area fixed-effects and further controls to deal with this missing amenity problem.

Finally, it could well be that wage differences across areas are driven by the proximity to markets for intermediate and final goods. These markets may have a spatial scale larger than employment areas as argued by much of the recent literature (Fujita *et al.*, 1999). Although a full treatment of interactions across employment areas is clearly beyond the scope of this paper, they matter to the extent that they may again bias some of the coefficients. In particular, employment areas with good market access may offer higher wages and be attractive to workers. In the second stage, this would imply again that the coefficient on employment density is over-estimated. To the extent that market access does not vary much over time, it will be controlled for by area fixed-effects. We also constructed and experimented with a series of market access variables that we use as extra controls in equation (10).

To summarise, we face both endogeneity and omitted variables problems that may bias the coefficients, particularly that on employment density (which captures between-industry interactions). Introducing some time-invariant area fixed-effects in (10) will take care of permanent unobserved characteristics. Using long-lagged values of employment density enables us to deal with contemporaneous shocks. Finally, we also use a limited set of controls for location characteristics and market access.

Before turning to the results, note that our estimation method allows us to estimate consistently not only the effect of a particular area on wages (Section 4) but also what determines such area fixed-effects, thanks to the second stage. This is in contrast with the labour market literature (e.g., Krueger and Summers, 1988; Gibbons and Katz, 1992; Abowd *et al.*, 1999), which often estimates industry effects, but usually does not attempt to explain them by the industry characteristics.

²⁷Better consumption amenities imply higher land prices. In turn higher land prices have a negative effect on local wages (equation 4). Since land prices are omitted from the wage equation, their negative effect enters the residual. Given that at the same time, better consumption amenities attract more workers, employment density is then negatively correlated with the residual and its coefficient is thus biased downwards. Such unobserved heterogeneity across locations is not specific to our analysis. It can potentially affect any attempt to estimate agglomeration effects (through wages or the estimation of production functions). It is however barely ever mentioned in this literature. Rauch (1993) and Wheaton and Lewis (2002) are two exceptions.

Table 5. Summary statistics for the variance decomposition — estimation of equation (10)

Effect of	Std dev	Simple correlation with:		
		$\log w$	δ	$\beta - \theta$
between-industry interactions ($I\gamma$)	0.068	0.20	0.12	0.85
density	0.078	0.20	0.12	0.84
land area	0.012	-0.15	-0.08	-0.61
diversity	0.001	0.09	0.02	0.32
residuals (η)	0.037	0.08	-0.06	0.20

2,221,156 observations. All correlations between effects that are not orthogonal by definition are significant at 1%. The effect of density is that of the log of density times its coefficient (in vector γ). Note that density is not instrumented here. The effect of land area is that of the log of the land area times its coefficient. The effect of diversity is that of the log of diversity as defined by an inverse Herfindhal index (as in Table 1) times its coefficient.

6. The determinants of area fixed-effects: results

Interactions between industries are captured by the log of the density of employment per square kilometre ($\log Den$), the log of land area ($\log Area$) and the log of industrial diversity ($\log Div$, the inverse of a Herfindhal index calculated as in Table 1).²⁸ In some regressions we also use a market potential variable (the average density of adjacent employment areas) and four amenity variables (the percentage of population in each employment area living in a municipality with a sea shore, mountains, a lake, and outstanding cultural or architectural heritage). Again, we start by a variance analysis before turning to the coefficients.

The importance of employment density

When we first regress the area fixed-effects on employment density alone, we get a *within* R^2 of 59%. Land area alone yields a R^2 of 48% whereas that of diversity is no more than 5%. Together, employment density, diversity and land area explain 60% of the variance. This suggests that most of the explanatory power comes from employment density and land area.

This result is made more precise by a more complete variance decomposition. The results are reported in Table 5 for the OLS regression including only density, area and diversity as explanatory variables. Within this group of variables, employment density clearly stands out. Land area is of secondary importance whereas the explanatory power of the diversity of local industrial composition is close to nil. The second-stage residuals have a small variance. This suggests a small explanatory power for local endowments.

²⁸Note that to be consistent we use the log values of the share of employment by industry (in the first stage) and of density and land area (in the second stage). This allows us to estimate the effect of a change in composition of activity keeping all else constant, a change in population keeping land area and composition constant, and a change in land area keeping density and composition constant (i.e., an increase in population keeping density constant). The effects of other changes can be easily computed by summing the coefficients. Alternative specifications using for instance industry employment, density, and total employment are certainly possible. However one must be careful with respect to the interpretation of the coefficients (Combes, 2000).

Table 6. Estimation results for equation (10)

Regression	(1) OLS 1	(2) FGLS	(3) OLS 2	(4) <i>within area</i>	(5) 2SLS 1	(6) 2SLS 2
$\log Den$	0.0371*** (0.0008)	0.0357*** (0.0010)	0.0377*** (0.0008)	0.0486*** (0.0038)	0.0245*** (0.0012)	0.0196*** (0.0014)
$\log Area$	0.0113*** (0.0014)	0.0106*** (0.0016)	0.0139*** (0.0018)	- -	-0.0055*** (0.0018)	0.0019 (0.0019)
$\log Div$	0.0020 (0.0023)	0.0006 (0.0025)	-0.0001 (0.0023)	-0.0079*** (0.0030)	0.0192*** (0.0030)	0.0157*** (0.0029)
Market potential			0.0003** (0.0001)			0.0018*** (0.0002)
Amenities	No	No	Yes	No	No	Yes
$\hat{\sigma}$	0.0336	0.0316	0.0328	0.0220	-	-

2,046 observations. Standard error between brackets. *: significant at 10%, **: significant at 5%, and ***: significant at 1%. The market potential of an employment area is defined as the mean log density in adjacent employment areas. Amenities are described by four different variables: (1) percentage of the population in the employment area living in a municipality where sea is a major attraction, (2) percentage living in a municipality where mountains are a major attraction, (3) percentage with lake, and (4) percentage of population living in a municipality with outstanding cultural or architectural heritage. In columns 5 and 6, density is instrumented by population density in 1936. The R^2 for the instrumental regression is 0.82. In column 6, market potential is instrumented by its lagged value for 1936. The R^2 for the instrumental regression is 0.48. We do not report time dummies.

Analysis of the coefficients

The coefficients obtained in the estimation of equation (10) are given in Table 6. The first column reports results for the baseline specification where density, land area and diversity are used as explanatory variables.²⁹ At 3.7%, the coefficient on density is at the lower bound of previous estimates in the literature (Ciccone and Hall, 1996; Ciccone, 2002; Rosenthal and Strange, 2004). This suggests that worker heterogeneity, which we control for, is captured in these other studies in part by density (see Section 7 for more on this). The coefficient on land area is smaller than that on density by a factor of three. A 1% increase in population through a higher density has a much larger wage effect than the same population increase obtained by a larger land area keeping density constant.

When using the same variables directly in equation (8) to perform a single-stage estimation (whose results are available upon request), we find very similar values for the effects of industry characteristics. The average coefficient of industry specialisation is 2.2% (against 2.1% in the two-stage estimation). The coefficient on employment density is also very close: 3.2% (against 3.7% in the two-stage estimation). That on land area shows a larger discrepancy at 2.1% (against 1.1%). The insignificant coefficient on industrial diversity changes sign. These differences between the two-stage and single-stage estimations find their sources in the correlations between the explanatory variables and the error terms. These correlations are not taken into account in both estimations and can lead to different biases. However, the explanatory power of both land area and diversity remain negligible in all cases so that these changes in the coefficients do not alter our conclusions.

²⁹It is likely that employment density does not affect all industries with the same intensity (Henderson, 2003). The two-step estimation prevents us from exploring this issue further. We leave it for future work.

Column 2 in Table 6 performs the same regression as the baseline but uses the FGLS correction discussed above. The differences with the baseline are minimal. This reflects the fact that the area fixed-effects are precisely estimated in the first stage.

In column 3, we added some controls for amenities (seaside, lake, mountains and architectural heritage) and market potential to the baseline regression. These extra controls increase the explanatory power of the regression only marginally since the R^2 (*within* time) rises from 60% to 62%. The signs of the amenity coefficients are negative and significant at 1% (bar that on mountains, which is positive and significant). As explained above, the coefficient on employment density in the baseline could be biased downwards because of missing amenities. It could also be biased upwards because of missing productive endowments. Seaside and lake may act as both consumption amenities and productive endowments. The fact that the coefficient on employment density increases with respect to the baseline suggests that the consumption aspect of the variables dominates. Regarding market potential, we tried a large variety of specifications. The best results were obtained using the weighted density of adjacent employment areas. In this case, the coefficient on market potential is significant but it is very small. This suggests that interactions across employment areas are limited in scope.

To control more systematically for the permanent unobserved characteristics of employment areas, we use a *within* area estimator in column 4. With respect to the baseline, we find that the coefficient on density increases to 4.9% whereas that on diversity becomes negative and significant. Again, following Roback (1982)'s argument, this suggests that there is indeed a downward bias in the baseline with respect to some permanent characteristics of employment areas. Put differently, because of unobserved heterogeneity across areas (location, amenities, etc), which impacts mostly on the consumption side, the baseline estimates for the coefficient on employment density are biased downwards. The negative sign of the coefficient on diversity in the same column is apparently more puzzling. Based on the fact that only the intertemporal variability is used in this estimation, one interpretation for this result is the following. Large declines in mean wages are typically observed in employment areas losing a dominant industry. Over the period, a number of French employment areas experienced large losses of employment in one heavy manufacturing industry. The loss of a dominant industry also implies mechanically a more diverse industrial structure. This generates a negative correlation between wages and diversity within areas.

In column 5 we repeat the baseline estimation but we instrument employment density by population density in 1936.³⁰ This enables us to control for contemporaneous shocks (on technology, endowments and amenities) correlated with density. In comparison with the baseline, the coefficient on density is much lower at 2.5%. This suggests that, as expected, positive local shocks lead to an influx of workers. Here the endogeneity bias is shown to be rather large. Using a similar instrumentation strategy (albeit with more instruments and longer lags), Ciccone and Hall (1996) find no evidence of such endogeneity. This is probably because they conduct their analysis at the level of US states whereas we do ours at the level of French employment areas, which are much smaller. Hence, it is not very surprising that the response of workers to local shocks is larger.

³⁰Instrumenting simultaneously by population density in 1954 and natality and mortality between 1954 and 1962 yields similar results.

The coefficient on diversity is now positive and significant whereas that on land area is negative and significant. This significant coefficient on diversity is probably caused by the direct positive correlation between diversity and current employment density. When instrumenting employment density by relatively long lagged values, the coefficient on diversity may pick up some residual effects of current employment density.

Finally in column 6, we still instrument employment density but we also introduce amenity controls and the market potential variable. Given that the latter is subject to the same endogeneity concerns as employment density, we instrument it by its 1936 value. This is our preferred estimation because the instruments take care of contemporaneous shocks whereas the extra controls capture (some of) the effects of the local structural characteristics of employment areas missing in the baseline. The results are fully consistent with those obtained so far. When employment density is instrumented, its coefficient is lower. Adding amenity variables raises it slightly whereas the market potential variable decreases it (more so when this variable is also instrumented). The final coefficient we find for employment density is just below 2%.

In summary, we find evidence of substantial reverse causality between wages and employment density leading to an upward bias for the OLS estimates. We also find that there are some permanent differences between employment areas with respect to consumption amenities that lead to a downward bias for the coefficient on employment density. Our main result for the second stage is that after controlling for worker observed and unobserved heterogeneity, the estimate for the elasticity of wages with respect to density is lower than in previous literature (2% as opposed to 4 to 8%). In addition, employment density appears to be the only robust variable capturing agglomeration economies.

Residual spatial wage disparities

To examine spatial wage disparities, we can now compute a ‘residual wage’, that is a local wage controlling for skills and all interactions, from the results of the baseline regression for the second stage. We can define such index $w_{\text{resid},a,t}$ (or residual wage) as:

$$\log w_{\text{resid},a,t} \equiv W + \hat{\eta}_{a,t}, \quad (11)$$

where W is defined in a similar way as after equation (9). This residual wage corresponds to the local wage obtained by an ‘average’ worker employed in an ‘average’ industry and in an area with ‘average’ interactions.

The ratio of the highest to the lowest residual wage across all employment areas is 1.35 instead of 1.51 for the de-trended net wage (i.e., the wage after controlling for skills and industry) and 1.88 for the de-trended mean wage. The ratio of the first to the last decile is 1.09 instead of 1.14 and 1.23 for net and mean wages, respectively. For the interquartile ratio, we find 1.04, 1.06 and 1.12 for residual, net, and mean wages respectively. Finally, the coefficient of variation for residual wages is 0.03 against 0.05 for net wages and 0.09 for mean wages. The salient result is thus that once skills and interactions are controlled for, between 60 and 70% of the wage disparities between employment areas disappear. Conversely, the remaining room for a possible role of local public endowments on disparities is necessarily small.

7. Aggregate wage differences across employment areas

Research is usually restricted in the data it can use. Existing studies on regional disparities typically use mean wages (or output per worker) by industry and location.³¹ It is of course impossible to directly implement our micro-founded specifications (8) and (10) with aggregate data. In this section, we first show how the simple model introduced above (where wages are determined *at the worker level*) can be aggregated and estimated at the level of each employment area and industry. We then compare the aggregate data results with those obtained above using individual data and with those in the literature.

Aggregation issues

Given the data, it is natural to first proxy worker fixed-effects by their occupation.³² However, occupations may change over time. They thus proxy for worker fixed-effects (δ_i) in a noisy fashion. We assume $\delta_i = \sum_{k,c} d_{i,k,c,t} \delta_{c,k} + \iota_{i,t}$ where $d_{i,k,c,t}$ is an occupation dummy taking value one when worker i is in occupation c and industry k at date t , $\delta_{c,k}$ is the corresponding coefficient, and $\iota_{i,t}$ is a residual term. Averaging (7) over all $N_{a,k,t}$ workers in the same local industry (a,k) in year t yields:

$$\log w_{a,k,t} = \frac{1}{N_{a,k,t}} \sum_{i \in (a,k,t)} \log w_{i,t} = \beta_{a,t} + \mu_{k,t} + I_{a,k,t} \gamma_k + \frac{1}{N_{a,k,t}} \sum_{i \in (a,k,t)} (X_{i,t} \varphi + d_{i,k,c,t} \delta_{c(i,t),k}) + \zeta_{a,k,t} \quad (12)$$

where $\zeta_{a,k,t} = \frac{1}{N_{a,k,t}} \sum_{i \in (a,k,t)} (\epsilon_{i,t} + \iota_{i,t})$.

If there is some sorting across space or industries leading the mean of the residual term $\iota_{i,t}$ to be correlated with some of the explanatory variables at the (a,k,t) level, the estimated coefficients are biased by aggregation. Put differently, if for instance workers with better unobservable characteristics are more likely to be located in denser areas, wage disparities across areas may be wrongly attributed to urbanisation economies. This is potentially a first major limitation when using aggregate data.

Another aggregation problem in equation (12) regards data availability. Typically, one may have access to the mean wage in an industry and area but not to the mean of log wages.³³ Hence the mean of log wages must be proxied by the log of mean wages. A similar problem arises among the explanatory variables when using (as we do) the squared age of workers. Again the mean of squared individual ages requires individual level data. With aggregate data, it can only be proxied

³¹Henderson (2003) and the literature on human capital externalities (following Rauch, 1993) are two exceptions.

³²Little else is present in the data about the skills of workers once we abstract from the longitudinal dimension of the panel. Fortunately, we can use the information about occupations (self-employed, professional, skilled worker, unskilled or semi-skilled white-collar worker, unskilled or semi-skilled blue-collar worker) which proxy well for skills according to Abowd *et al.* (1999).

³³Specifications with log mean wage are typically used in the literature. They are naturally justified with representative worker models of wage determination. As shown by our model, with heterogeneous workers, mean log wage is needed.

by the square of the mean age. This implies some measurement problems for wages and squared age at the area-industry level.³⁴

We can again centre within-industry interactions and worker time-varying characteristics so that all systematic industry components can be brought together with the industry fixed-effect.³⁵ The main difference with the analysis on individual data is that the total industry effect now also reflects the average occupational composition of the industry. Then, using the same reasoning as with individual data, we obtain:

$$\begin{cases} \log w_{a,k,t} = \mu_k + \beta_{a,t} + \tilde{I}_{a,k,t}\gamma_k + \tilde{X}_{a,k,t}\varphi + \sum_c \tilde{q}_{c,a,k,t}\delta_{c,k} + \zeta_{a,k,t}, \\ \beta_{a,t} = w_0 + \theta_t + E_{a,t}\alpha + I_{a,t}\gamma + v_{a,t}. \end{cases} \quad (13)$$

These two equations mirror equations (8) and (10). As argued above, the share of workers in professional occupations in industry and employment areas should be used as one of the regressors in the vector $\tilde{I}_{a,k,t}$ to capture human capital interactions within industries. However this variable also now appears independently in equation (13) following the aggregation of individual skills. Hence the coefficient on the share of professionals captures both skill composition effects *and* local interactions in the industry. The two cannot be separately identified. This constitutes another limitation of aggregate data.

Finally, the first stage equation must be estimated by weighting each observation by the square-root of its number of workers to avoid heteroscedasticity (Coelho and Ghali, 1973).³⁶ Turning to the second stage (and as previously), we do not know the true values of the area fixed-effects, $\beta_{a,t}$. Hence, we use $\hat{\beta}_{a,t}$ rather than $\beta_{a,t}$ keeping a similar estimation method as before (again see Appendix C). We also impose the same identification conditions: $\mu_1 = 0$ and $\theta_{1980} = 0$.

To summarise, aggregation throws up two main problems. First, sorting with respect to unobserved characteristics can lead to strong biases. Second, skill composition effects can no longer be identified separately from human capital interactions/externalities.

Results

The first stage of the regression with all the variables (7514 in total) has a R^2 of 81% compared with 31% for the same regression with individual data without the worker fixed-effects. This difference is obviously explained by the considerable variation in wages at the worker level that is averaged out by aggregation. Running the regression with only year dummies yields a R^2 of 21% whereas with area-year dummies it rises to 46%.

We then perform the same detailed variance analysis as done in the first stage of the estimation. It shows that the share of professionals has a much larger explanatory power than with individual

³⁴However, these measurement problems are very minor. The correlations between mean log wage and log mean wage by industry and location and that between mean squared age and squared mean age by location are both equal to 0.99.

³⁵Define the centred share of occupation c in (a,k,t) : $\tilde{q}_{c,a,k,t} \equiv q_{c,a,k,t} - q_{c,,k,t}$ where $q_{c,a,k,t} \equiv \frac{1}{N_{a,k,t}} \sum_{i \in (a,k,t)} d_{i,k,c,t}$ is the share of occupation c in (a,k,t) and $q_{c,,k,t}$ its weighted mean across all employment areas. To mirror the approach developed in Section 4, assume $\mu_{k,t} + I_{.,k,t}\gamma_k + X_{.,k,t}\varphi + \sum_c q_{c,,k,t}\delta_{c,k} = \mu_k + \rho_t$, that is the sum of all the industry effects can be decomposed into a time-invariant industry effect and a time effect (which is again normalised to zero).

³⁶Existing work on aggregate data tends not to weight the mean wage for local industries. This would be odd in our context because local industries are not autonomous agents but rather a collection of different workers.

Table 7. Estimation results for the second stage of equation (13)

Regression	(1) OLS 1	(2) FGLS	(3) OLS 2	(4) <i>within</i> area	(5) 2SLS 1	(6) 2SLS 2
$\log Den$	0.0625*** (0.0005)	0.0618*** (0.0005)	0.0628*** (0.0005)	0.0260*** (0.0022)	0.0455*** (0.0008)	0.0416*** (0.0010)
$\log Area$	0.0344*** (0.0008)	0.0359*** (0.0008)	0.0358*** (0.0010)	- -	0.0103*** (0.0013)	0.0165*** (0.0014)
$\log Div$	0.0007 (0.0014)	-0.0008 (0.0014)	0.0001 (0.0013)	-0.0158*** (0.0017)	0.0261*** (0.0021)	0.0236*** (0.0021)
Market potential			0.0002*** (0.0001)			0.0016*** (0.0002)
Amenities	No	No	Yes	No	No	Yes
$\hat{\sigma}$	0.0354	0.0354	0.0327	0.0205	-	-

6,820 observations. Standard error between brackets. *: significant at 10%, **: significant at 5%, and ***: significant at 1%. The market potential of an employment area is defined as the mean log density in adjacent employment areas. Amenities are described by four different variables: (1) percentage of the population in the employment area living in a municipality where sea is a major attraction, (2) percentage living in a municipality where mountains are a major attraction, (3) percentage with lake, and (4) percentage of population living in a municipality with outstanding cultural or architectural heritage. In columns 5 and 6, density is instrumented by population density in 1936. In column 6, market potential is instrumented by its lagged value for 1936. The R^2 for the instrumental regression is 0.48.

data. This is true more generally for the effect of all the explanatory variables we consider. The standard deviation for the wages is at 0.258 against 0.367 with individual data. The standard deviation for the area fixed-effect (after subtracting time effects) is at 0.074 (against 0.065 previously). That for the effect of age and its square is unchanged at 0.058, that for industry fixed-effects is at 0.097 (against 0.043 previously), that for specialisation is at 0.047 (against 0.017), and that for the number of establishments is at 0.035 (against 0.019). Finally with aggregate data the standard deviation for the share of professionals is four times as large at 0.046 (against 0.011 with micro data). The effect of all the occupations (including those that do not enter the micro specification) has a standard deviation equal to 0.110.

With respect to the share of the various occupations the high variance of the effects was to be expected given that these variables now capture both the skill composition of the local industry and some interactions therein. For the other variables (and in particular the nearly tenfold change in the variance of the effect of specialisation), this indicates that some correlation with individual unobserved heterogeneity is present. The same conclusion arises with the second stage of the regression. The R^2 (*within* time) of the second stage of the baseline regression is above what we obtained with individual data at 77% (against 60%). Aggregate variables capture some of the unobserved heterogeneity among workers, which is not controlled for.

The first-stage coefficients are nearly all significant. For instance only 0.3% of the area fixed-effects fail to be significant at 5% (instead of 12% with individual data). Because they capture within-industry interactions together with compositional effects, the coefficients on the share of professionals are much higher than with individual data. More interestingly the specialisation coefficients are also much higher: on average 4.3% against 2.1%. Similar discrepancies occur with regard to the second stage coefficients (see Table 7). In the baseline specification, the coefficient on density is at 6.3% rather than 3.7% with individual data. That on land area is at 3.4% against

Table 8. Correlation between the effects of the variables after aggregation by area and year

	area f.-e.	density	area	diversity	residuals
mean worker f.-e.	0.29	0.44	-0.22	-0.08	-0.08
area f.-e.	1	0.77	-0.34	0.23	0.63
density		1	-0.58	0.20	0
land area			1	0.24	0
diversity				1	0

2,046 observations computed from the estimations at the individual level. All correlations between variables not orthogonal by construction are significant at 1%.

Area fixed-effects are estimated from (8) and we subtracted time fixed-effects estimated from (10). Worker fixed-effects are estimated from (8) and then averaged by employment area. The effects of density, land area and diversity are computed using their coefficients as estimated in (10) times the value of the variable.

1.1% with individual data. In the aggregate data equivalent of our preferred specification (column 6), we find that the coefficient on employment density is still at 4.2% against less than 2% with individual data.

As can be seen from Table 8, these discrepancies are easily explained by the sorting of workers by skills. We have already underlined in Section 4 that the correlation between the average worker fixed-effect by area and the area-year fixed-effect at 0.29 is high in individual regressions. It is even higher (0.53) when the area-year fixed-effects are computed on aggregate data. This implies that in the aggregate analysis area-year fixed-effects capture a large fraction of the worker fixed-effects that is not captured by occupation shares and within-industry interactions. Moreover, average worker fixed-effects also have a high correlation (0.44) with employment density.

In conclusion, when sorting is not taken into account the coefficient on density is over-estimated by 70 to 100%, that on land area is over-estimated by about 200% whereas those on specialisation are over-estimated by 100%. These are clearly large biases.

8. Concluding comments

This paper proposes a general framework to investigate the sources of wage disparities across local labour markets: skills, endowments and within- and between-industry interactions. This framework unites different strands of literature that were so far mostly disjoint. It shows that the research about the 'estimation of agglomeration economies' is closely intertwined with those dealing with 'regional disparities', 'local labour markets' and 'migration'.

Empirically, the main novelty of the paper is to use a very large panel of workers and a consistent approach to exploit it. This allows us to assess precisely the effects of unobserved worker heterogeneity. We find that the effect of individual skills is quantitatively very important in the data. Up to half of the spatial wage disparities can be traced back to differences in the skill composition of the workforce. Workers with better unobserved labour market characteristics tend to agglomerate in the larger, denser and more skilled local labour market. We believe more work

is now needed to understand the nature of this sorting.³⁷

We also pay considerable attention to the issues of reverse causality and missing local variables. As suggested by the literature on amenities, we find indirect evidence of differences in the permanent characteristics of local labour markets leading to a downward bias when measuring agglomeration economies. However, we also find that positive local technological shocks have a positive impact on local employment leading to an upward bias when measuring agglomeration economies. When correcting for these biases, our estimates for economies of density, at around 2%, are much lower than in previous literature. Nonetheless they still play an important role in explaining differences in local wages. The evidence on other types of local interactions such as those taking place within particular industries is more mixed. They are significant but do not matter much quantitatively in explaining local wages disparities. Our approach also suggests at best a modest direct role for local non-human endowments in the determination of local wages. Local public goods in particular may not have a large direct impact on productivity and wage disparities. However, what remain to be tested, and should be the purpose of future research, is whether they might be responsible for the observed spatial sorting of skilled workers, and thus exert an indirect influence on productivity. In any case, the final picture we reach is one where high local wages are primarily the outcome of high-skill workers gathering in dense local labour markets.

³⁷One explanation could be based on a self-selection effect in internal migrations. As suggested long ago by Alfred Marshall, it may be that "the most enterprising, the most highly gifted, those with the highest physique and strongest character go [to the large towns] to find scope for their abilities" (Marshall, 1890). Nocke (2003) proposes a formalisation of this argument. Alternatively, the largest cities may offer some particular amenities that appeal more to the workers commanding the highest wages. A third hypothesis (Glaeser and Maré, 2001) is that workers may learn more in larger cities.

References

- Abdel-Rahman, Hesham M. 1988. Product differentiation, monopolistic competition and city size. *Regional Science and Urban Economics* 18(1):69–86.
- Abowd, John M., Francis Kramarz, Thomas Lemieux, and David N. Margolis. 2000. Minimum wages and youth employment in France and the United States. In David G. Blanchflower and Richard B. Freeman (eds.) *Youth Employment and Joblessness in Advanced Countries*. Cambridge (Mass.): NBER.
- Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. High wage workers and high wage firms. *Econometrica* 67(2):251–333.
- Ciccone, Antonio. 2002. Agglomeration effects in Europe. *European Economic Review* 46(2):213–227.
- Ciccone, Antonio and Robert E. Hall. 1996. Productivity and the density of economic activity. *American Economic Review* 86(1):54–70.
- Ciccone, Antonio and Giovanni Peri. 2002. Identifying human capital externalities: Theory with an application to US cities. Discussion Paper 3350, Centre for Economic Policy Research. URL <http://www.cepr.org/>.
- Coelho, Philip R. P. and Moheb A. Ghali. 1973. The end of the North-South wage differential: Reply. *American Economic Review* 63(4):757–762.
- Cohen, Daniel, Arnaud Lefranc, and Gilles Saint-Paul. 1997. French unemployment: A transatlantic perspective. *Economic Policy: A European Forum* 0(25):265–285.
- Combes, Pierre-Philippe. 2000. Economic structure and local growth: France, 1984–1993. *Journal of Urban Economics* 47(3):329–355.
- Duranton, Gilles and Vassilis Monastiriotis. 2002. Mind the gaps: The evolution of regional earnings inequalities in the UK 1982–1997. *Journal of Regional Science* 42(2):219–256.
- Duranton, Gilles and Diego Puga. 2004. Micro-foundations of urban agglomeration economies. In Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North-Holland (forthcoming).
- Durlauf, Steven N. and Dany T. Quah. 1999. The new empirics of economic growth. In John B. Taylor and Michael Woodford (eds.) *Handbook of Macroeconomics vol. 1A*. North-Holland: Elsevier Science, 231–304.
- Fujita, Masahisa. 1988. A monopolistic competition model of spatial agglomeration: A differentiated product approach. *Regional Science and Urban Economics* 18(1):87–124.
- Fujita, Masahisa, Paul R. Krugman, and Anthony J. Venables. 1999. *The spatial economy: Cities, regions, and international trade*. Cambridge, MA: MIT Press.
- Gibbons, Robert and Lawrence Katz. 1992. Does unmeasured ability explain inter-industry wage differentials? *Review of Economic Studies* 59(3):515–535.
- Glaeser, Edward L. and David C. Maré. 2001. Cities and skills. *Journal of Labor Economics* 19(2):316–342.
- Gobillon, Laurent. 2003. The estimation of aggregate effects on individual outcome using panel data: A simple two-stage method. Processed, University College London.

- Helsley, Robert W. and William C. Strange. 1990. Matching and agglomeration economies in a system of cities. *Regional Science and Urban Economics* 20(2):189–212.
- Henderson, J. Vernon. 1986. Efficiency of resource usage and city size. *Journal of Urban Economics* 19(1):47–70.
- Henderson, J. Vernon. 2003. Marshall's economies. *Journal of Urban Economics* 53(1):1–28.
- Hoover, Edgar M. 1936. *Location Theory and the Shoe and Leather Industries*. Cambridge, MA: Harvard University Press.
- Krueger, Alan B. and Lawrence H. Summers. 1988. Efficiency wages and the inter-industry wage structure. *Econometrica* 56(2):259–293.
- Lucas, Robert E. Jr. 1988. On the mechanics of economic development. *Journal of Monetary Economics* 22(1):3–42.
- Marshall, Alfred. 1890. *Principles of Economics*. London: Macmillan.
- Moomaw, Ronald L. 1981. Productivity and city size? A critique of the evidence. *Quarterly Journal of Economics* 96(4):675–688.
- Moulton, Brent R. 1990. Illustration of a pitfall in estimating the effects of aggregate variables on micro unit. *Review of Economics and Statistics* 72(2):219–256.
- Nocke, Volker. 2003. A gap for me: Entrepreneurs and entry. Processed, University of Pennsylvania.
- Rauch, James E. 1993. Productivity gains from geographic concentration of human capital: Evidence from the cities. *Journal of Urban Economics* 34(3):380–400.
- Rice, Patricia and Anthony J. Venables. 2003. Equilibrium regional disparities: Theory and British evidence. Processed, London School of Economics.
- Roback, Jennifer. 1982. Wages, rents and the quality of life. *Journal of Political Economy* 90(6):1257–1278.
- Rosenthal, Stuart S. and William C. Strange. 2004. Evidence on the nature and sources of agglomeration economies. In Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North-Holland (forthcoming).
- Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: Printed for W. Strahan, and T. Cadell.
- Sveikauskas, Leo A. 1975. The productivity of cities. *Quarterly Journal of Economics* 89(3):393–413.
- Sveikauskas, Leo A., John Gowdy, and Michael Funk. 1988. Urban productivity: City size or industry size. *Journal of Regional Science* 28(2):185–202.
- Temple, Jonathan. 1999. The new growth evidence. *Journal of Economic Literature* 37(1):112–156.
- Topel, Robert H. 1986. Local labor markets. *Journal of Political Economy* 94(3(2)):S111–S143.
- Wheaton, William C. and Mark J. Lewis. 2002. Urban wages and labor market agglomeration. *Journal of Urban Economics* 51(3):542–562.

Appendix A. Data description and background

In this appendix, we further describe the data and the basic features of the wage setting process in France. These descriptions are by no means comprehensive. A very detailed description of the data can be found in Abowd *et al.* (1999). Cohen, Lefranc, and Saint-Paul (1997) and Abowd, Kramarz, Lemieux, and Margolis (2000) propose broad syntheses of recent research on the French labour market.

The DADS data

The DADS extract we use covers around 5% of the employees and self-employed in France. It contained initially 19,675,740 observations from 1976 to 1998. The selection method (persons born in October of even years) and the mandatory aspect of the underlying data guarantee that this sample is representative and not plagued by non-response problems. A few issues must however be discussed.

- *Missing years.* The three missing years (1981, 1983 and 1990) are due to lack of sampling by INSEE during census periods. Since in the individual estimations we keep only one year every four starting in 1976 this does not create any particular problem in our analysis (beyond the obvious missing data issue in the aggregate estimations).
- *Unemployment.* Unemployment spells can be detected only indirectly through interruptions between different episodes of employment for which only the year is known. This impossibility to track precisely unemployment is an issue when estimating a cross-section as there may be some spatial selection with respect to the unobserved abilities of individuals at work.
- *Wages, earnings and labour costs.* For each observation, total net nominal earnings, number of days worked and the work status (full-time or part-time) are known. For full-time workers it is thus possible to calculate their annualised nominal wage. Then mandatory payroll taxes for both employees and employer can be calculated and added to net nominal earnings to obtain total annualised labour costs. These payroll taxes are large (for low wages they are of the same order of magnitude as net earnings) and have increased over time. They also differ across wage levels and work status — for instance the fiscal and pension regimes of executives (*cadres* in French) differs from that of non-executives. These payroll taxes however are uniform across industries (with textile being an exception which is taken into account) and across mainland France.
- *Imputed wages.* For workers in the same establishment in year t and $t + 2$ but no observation for year $t + 1$, the original data contains an imputed value for this missing year. Since our estimation methods apply to unbalanced panels, we deleted these imputed values and ended up with 18,581,470 observations.
- *Missing values and coding errors.* We deleted all the observations for which one or more variables of interest (such as age or location) was missing. We also deleted all workers with

a duration of employment equal to zero, workers with negative wages and workers not born in October of even years who should not be in the data. After these deletions, we were left with 17,495,335 observations. We also deleted all the observations for which we could not determine the industry of employment or the employment area. This left us with 16,458,989 observations.

- *Mainland private sector employees of working age.* We excluded all apprentices and workers not employed in the private sector. We also restricted the sample to workers aged 15 to 65 employed in mainland France. Workers employed in Corsica and overseas territories were deleted to end up with 14,067,326 observations.
- *Part-timers.* Because the number of hours is unknown before 1993, we excluded all part-time workers. In case of multiple observations for a worker over a given year (corresponding to more than one job), we kept only one observation (the one with the most working days). This left us with 10,551,810 observations.
- *Agriculture, fishing and small industries.* The French sectoral classification comprised around 600 finely defined industries before 1993 and 700 after this date following an overhaul of the industrial classification system. We aggregated these finely-defined industries into 114 industries whose definition has remained constant over time. Agriculture and fishing industries are not normally covered by the extract. Remaining workers in these sectors were excluded. We also excluded various other industries:
 - All industries with less than 500 observations over the period were excluded. This includes: Spatial transport (70 observations), Extraction of uranium (129) and Extraction of metals (311).
 - In a few industries, firms with a large number of establishments can aggregate their reporting at the regional level. In the data, one particular branch (which may change from year to year) appears to host all the employees for a large region. This of course biases the estimates regarding the effects of local specialisation. Consequently we excluded financial intermediation, insurance, financial auxiliaries, telecommunications, and postal services.
 - Finally we also excluded a few non-competitive industries: public administration, extra-territorial activities, and associations.

In total 15 industries in 114 were excluded and we ended up with 9,389,838 observations.

- *Outliers.* The initial data had a significant number of outliers with wages either unrealistically high or well below the minimum wage. These seem to be caused by reporting mistakes in the net nominal earnings or in the number of working days. We decided to get rid of the 3% lowest and highest wages for every year.

The final sample contains 8,826,422 observations. When working with the 6 years we selected (1976, 1980, 1984, 1988, 1992, and 1996), the sample contains a total of 2,664,474 observations.

When we aggregate the data by area, industry, and year we end up with 378,022 observations for the 1976-1998 period.

French employment areas and area based data

The units of analysis are the 341 Employment areas (or 'zones d'emploi') which entirely and continuously cover mainland France. Each employment area has an average area of 1570km² which is fairly small (equivalent to a square of 40 by 40 kilometres. Counties in mainland US are about 60% larger). The standard deviation, at 987km², is quite large. French employment areas have been defined primarily by worker's commuting patterns. Other criteria such as access to a major infrastructure have also been used to make French employment areas more homogenous than administrative units and minimise border effects. We think French employment areas closely match the idea of local labour markets. This makes French employment areas ideally suited for our analysis.

Each municipality (or 'commune') is fully included in an employment area. This makes it possible to use the 1988 census of facilities ('inventaire communal') conducted in all municipalities to compute the amenity variables.

Wage setting in France

Before 1982, wages were set through centralised collective bargaining at the industry level. The agreements reached by workers' and employers' unions were binding for all parties involved in the negotiation. These agreements were then usually extended by the French government to all firms and workers in the industry. Despite a low level of unionisation (circa 15%), these agreements would cover up to around 95% of the workforce. In 1982, a new legislation also required firms with at least 50 employees to conduct firm level negotiations. Such agreements could only improve on the industry level agreements. Around 30% of the workforce was covered by these agreements in the mid 1980s (see Abowd *et al.*, 1999, and the references therein for more on this issue).

A high minimum wage is another key feature of the French labour market. The minimum wage was introduced in France in the early 1950s. It was originally indexed on the average blue-collar worker's wage and the consumer price index. In addition to this indexation, the government often increased the minimum wage between 1976 and 1998. According to Abowd *et al.* (2000) the real hourly minimum wage increased by around 40% between 1976 and 1994. Furthermore, over the 1980s payroll taxes also increased as a proportion of the earnings making the increase in minimum wage even larger for employers. The minimum wage is also widely used as an index for other low wages. In 1990, 28% of the workforce employed in France were paid below 120% of the minimum wage (Abowd *et al.*, 2000). Hence, it is at least a third of the French wages which are affected directly by changes in the minimum wage.

These institutional settings are not very different from those in other countries from Continental Europe but they seem to differ widely from those prevailing in the US and UK. Recent comparisons between the French and US labour markets (Cohen *et al.*, 1997; Abowd *et al.*, 1999, 2000) nonetheless suggest that wage setting outcomes in the two countries have many features in common. For

instance, standard wage regressions on individual workers yield coefficients that are remarkably close in the two countries (Abowd *et al.*, 1999). It was also found that the upper half of the wage distribution looks very similar in both countries (Cohen *et al.*, 1997), etc.

Appendix B. Endogeneity of location and industry choices

We examine here the necessary assumptions about migrations and workers flows between industries for the strict exogeneity of the industry and location of employment to be warranted.

Consider a worker i having to choose an employment area and an industry in a static framework. We assume that this worker's utility depends only on her level of consumption of a composite good whose price is the same everywhere. Indirect utility can then be written as a function of the wage: $v = v(w)$. Worker i chooses her employment area and industry so as to maximise her wages net of the (monetary) costs of migration. This choice can be decomposed in three steps.

1) At the beginning of period t , any industry k in an employment area a can be characterised by a wage $w_{i,a,k,t}$. This wage depends not only on individual attributes and local characteristics of the industry, but also on a shock noted $\psi_{i,a,k,t}$. Using (4) and (5), the wage satisfies:

$$\log w_{i,a,k,t} = \log B_{a,k,t} + X_{i,t}\varphi + \delta_i + \psi_{i,a,k,t}. \quad (\text{B } 1)$$

We assume that all the explanatory variables in $B_{a,k,t}$ and $X_{i,t}$ are strictly exogenous.

2) The worker then chooses an employment area $a(i,t)$ and an industry $k(i,t)$ so as to maximise her utility. Assume first that the worker knows the distribution of the shocks $\psi_{i,a,k,t}$ without knowing their exact values. The maximisation programme of the worker is then:

$$\max_{(a,k) \in t} E_{\psi_{i,a,k,t}} [v(w_{i,a,k,t} - c_{a,k})], \quad (\text{B } 2)$$

where $E_{\psi_{i,a,k,t}}$ is the expectation operator on the distribution of $\psi_{i,a,k,t}$, and $c_{a,k}$ is a mobility cost equal to zero when $a = a(i,t-1)$ and $k = k(i,t-1)$. In this case, the choice of $a(i,t)$ and $k(i,t)$ is independent from the realisation of $\epsilon_{i,t} = \psi_{i,a(i,t),k(i,t),t}$. The location and industry of employment are thus determined solely on the basis of exogenous variables entering the wage equation and the mobility costs. Hence, when the worker knows only the distribution of the shocks, the assumption of strict exogeneity is satisfied.

Turning now to the case where the worker can observe all the $\psi_{i,a,k,t}$, the maximisation programme is:

$$\max_{(a,k) \in t} [v(w_{i,a,k,t} - c_{a,k})]. \quad (\text{B } 3)$$

In this case, the choice of $a(i,t)$ and $k(i,t)$ is correlated with the realisation of all shocks $\psi_{i,a,k,t}$, and in particular $\epsilon_{i,t} = \psi_{i,a(i,t),k(i,t),t}$. Under the assumption of strict exogeneity of the explanatory variables, the model is misspecified.

There are finally intermediate cases for which only some $\psi_{i,a,k,t}$ are observed by the worker. If these observed shocks are not correlated with $\epsilon_{i,t}$, the exogeneity assumption is satisfied. If they are, the model is misspecified again.

3) After choosing an employment area and industry, the individual shock, $\epsilon_{i,t}$, is known and the worker is paid according to (7). The worker then faces the same decision at period $t + 1$.

In a dynamic framework

Consider for simplicity that the evolution of all explanatory variables other than area-year and industry dummies, noted $Y_{i\tau}$, are exogenous. We also ignore savings. At period t , the worker chooses her location and industry taking into account all available information including the observed shocks $\psi_{i,a,k,t}$ and their evolution. We introduce the following notations: $Y_i^t = \{Y_{i\tau}\}_{\tau \leq t}$ and $\psi_i^t = \{\psi_{i,a,k,\tau} \mid a \leq Z, k \leq K, \tau \leq t, \psi_{i,a,k,t} \text{ known by } i\}$. The vector of state variables at the beginning of period t is $(\psi_i^{t-1}, a(i, t-1), k(i, t-1))$. Past employment area $a(i, t-1)$ and industry $k(i, t-1)$ enter this vector because mobility costs can depend on them. The history of observed shocks ψ_i^{t-1} is included because it can be used to predict the current and future realisations of shocks. The sequences of expected locations and industries are noted $\{a(i, \tau)\}_{t \leq \tau \leq T}$ and $\{k(i, \tau)\}_{t \leq \tau \leq T}$, respectively, with T the last period of work for i . Any worker solves:

$$\max_{(a_t, k_t) \in t, \dots, (a_T, k_T) \in T} E \left[\sum_{\tau=t}^T \rho^\tau v(w_{i,a_\tau, k_\tau, \tau} - c_{a_\tau, k_\tau}) \mid Y_i^T, \psi_i^{t-1}, Z(i, t-1), K(i, t-1) \right]. \quad (\text{B } 4)$$

with ρ the discount rate.

We can reach different conclusions depending on the dynamic process determining the shocks $\psi_{i,a,k,t}$. If we first suppose that shocks are idiosyncratic, the same conclusions as in the static case apply. The location $a(i, t)$ and the industry $k(i, t)$ are correlated with $\epsilon_{i,t}$ if and only if the worker can collect information on $\epsilon_{i,t}$ at period t . If we suppose instead that shocks follow an $AR(1)$ process and that the worker can obtain some information on $\epsilon_{i,t}$ through her history of shocks ψ_i^{t-1} , then three issues arise:

- 1) The location $a(i, t)$ and the industry $k(i, t)$ are correlated with $\epsilon_{i,t}$. This correlation is however much weaker than in the static case because workers take into account future wages in their mobility decisions. Indeed, the information related to current shock present in future wage shocks is decreasing with the time horizon and becomes negligible when it grows arbitrarily large.
- 2) $a(i, t)$ and $k(i, t)$ are correlated with past shocks $\{\epsilon_{i,\tau}\}_{\tau < t}$ as shocks follow an $AR(1)$ process.
- 3) $a(i, t)$ and $k(i, t)$ are correlated with future shocks $\{\epsilon_{i,\tau}\}_{\tau > t}$. However, the predictive power of the information set at t decreases over time. Thus, the worker can form only inaccurate expectations about future shocks. Thus the correlation between $a(i, t)$ and $k(i, t)$ in the one hand, and $\epsilon_{i,\tau}$, for $\tau > t$, in the other hand, decreases when τ increases.

These three remarks suggest that the results may be biased because the explanatory variables can be correlated not only with present shocks, but also with past and future shocks. However, although we may have more sources of bias than in the static case, these correlations are likely to be weak because workers take future wages into account in their mobility decision while having little information about future shocks. Extensions to other dynamic processes for the shocks are straightforward.

Appendix C. Two-stage estimation

What follows is a complete description of our two-stage estimation procedure.

Equation (10) can be re-written compactly:

$$\beta = D\Phi + \eta, \quad (\text{C } 1)$$

where $\beta = (\beta_{1,1}, \dots, \beta_{Z,T})'$, $\Phi = (w_0, \theta_1, \dots, \theta_T, \gamma)'$, D is the matrix of all explanatory variables after vectorisation, and $\eta = (\eta_{1,1}, \dots, \eta_{Z,T})'$.

Because the exact value of the area fixed-effects is unknown, this equation cannot be directly estimated with OLS. It is however possible to compute a consistent and unbiased estimator of β from the first stage results. Note first that (C 1) can be transformed into:

$$\hat{\beta} = D\Phi + \eta + \Psi, \quad (\text{C } 2)$$

where $\hat{\beta} = (\hat{\beta}_{1,1}, \dots, \hat{\beta}_{Z,T})'$ is the estimator of β obtained in the first stage of the regression and $\Psi = \hat{\beta} - \beta$ is a measurement error. Equation (C 2) can then be estimated in the following way:

1. Compute the OLS estimate of Φ from (C 2):

$$\hat{\Phi}_{OLS} = (D'D)^{-1}D'\hat{\beta} = \Phi + (D'D)^{-1}D'(\eta + \Psi) \quad (\text{C } 3)$$

2. It is then possible to define $\hat{\sigma}^2$ such that:

$$\hat{\sigma}^2 = \frac{1}{tr(M_D)} \left\{ \left(\widehat{\eta + \Psi} \right)' \left(\widehat{\eta + \Psi} \right) - tr \left[M_D \widehat{V}(\Psi) \right] \right\}, \quad (\text{C } 4)$$

where $M_D = I - D(D'D)^{-1}D'$, $\widehat{\eta + \Psi} = \hat{\beta} - D\hat{\Phi}_{OLS} = M_D(\eta + \Psi)$, and $\widehat{V}(\Psi)$ is the estimator of the covariance matrix obtained from the first stage estimation. As shown by Gobillon (2003), $\hat{\sigma}^2$ is an unbiased estimator of σ^2 when η is orthogonal to ϵ . It is also consistent under reasonable assumptions.

3. We can now compute an unbiased estimator of the covariance matrix $V(\eta + \Psi)$:

$$\widehat{V} = \hat{\sigma}^2 I + \widehat{V}(\Psi). \quad (\text{C } 5)$$

4. Measurement errors on the dependant variable create some heteroscedasticity. To control for this, the feasible generalised least-square (FGLS) estimator of Φ can be computed. It is given by:

$$\hat{\Phi}_{FGLS} = \left(D' \widehat{V}^{-1} D \right)^{-1} D' \widehat{V}^{-1} \hat{\beta}. \quad (\text{C } 6)$$

5. Finally, it is possible to compute a consistent estimator of the variance of $\hat{\Phi}_{FGLS}$:

$$\widehat{V} \left(\hat{\Phi}_{FGLS} \right) = \left(D' \widehat{V}^{-1} D \right)^{-1}. \quad (\text{C } 7)$$