

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: The Economics of Agglomeration

Volume Author/Editor: Edward L. Glaeser

Volume Publisher: University of Chicago Press

Volume URL: <http://www.nber.org/books/glae08-1>

Conference Date: November 30-December 1, 2007

Publication Date: unknown

Chapter Title: Estimating Agglomeration Economies with History, Geology, and Worker Effects

Chapter Author: Pierre-Philippe Combes, Gilles Duranton, Laurent Gobillon, S bastien Roux

Chapter URL: <http://www.nber.org/chapters/c7978>

Chapter pages in book:

Estimating agglomeration economies with history, geology, and worker effects

Pierre-Philippe Combes

Gilles Duranton

Laurent Gobillon

Sébastien Roux

13 March 2009

ABSTRACT: Does productivity increase with density? We revisit the issue using French wage and TFP data. To deal with the ‘endogenous quantity of labour’ bias (i.e., urban agglomeration is consequence of high local productivity rather than a cause), we take an instrumental variable approach and introduce a new set of geological instruments in addition to standard historical instruments. To deal with the ‘endogenous quality of labour’ bias (i.e., cities attract skilled workers so that the effects of skills and urban agglomeration are confounded), we take a worker fixed-effect approach with wage data. We find modest evidence about the endogenous quantity of labour bias and both sets of instruments give a similar answer. We find that the endogenous quality of labour bias is quantitatively more important.

Key words: agglomeration economies, instrumental variables, wages, TFP

JEL classification: R12, R23

Pierre-Philippe Combes is a CNRS research professor at GREQAM, University of Aix-Marseille, and Paris School of Economics, and a fellow at the CEPR.

Address: 2 Rue de la Charité, 13236 Marseille cedex 02, France (e-mail: ppcombes@univmed.fr; website: <http://www.vcharite.univ-mrs.fr/pp/combes/>).

Gilles Duranton is a professor of Economics at the University of Toronto where he holds the Noranda Chair in International Trade and Development. He is also a fellow at the CEPR.

Address: 150 Saint George Street, Toronto, Ontario M5S 3G7, Canada (e-mail: gilles.duranton@utoronto.ca; website: <http://individual.utoronto.ca/gilles/default.html>).

Laurent Gobillon is a researcher Institut National d'Études Démographiques (French Institute of Demographics), an associate researcher at CREST, Paris, an associate researcher at the Paris School of Economics and a research affiliate at the CEPR.

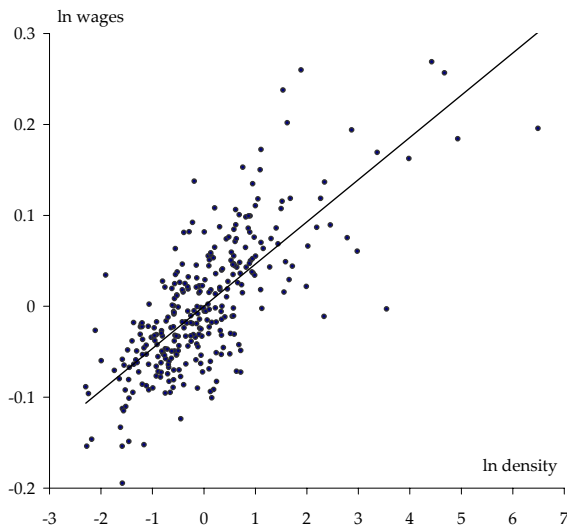
Address: 133 Boulevard Davout, 75980 Paris cedex 20, France (e-mail: laurent.gobillon@ined.fr; website: <http://laurent.gobillon.free.fr/>).

Sébastien Roux is a researcher at Centre de Recherche en Économie et Statistique (CREST-INSEE), Paris.

Address: 15 Boulevard Gabriel Péri, 92245 Malakoff cedex, France (e-mail: roux@ensae.fr).

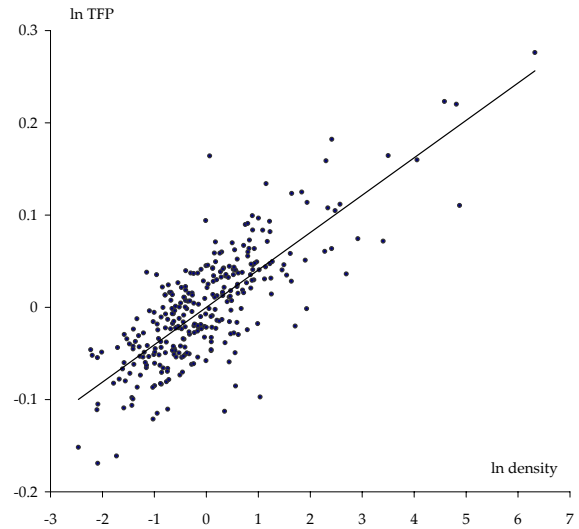
Acknowledgements: We are grateful to Alejandra Castrodad for superb research assistance with the GIS work and to France Guérin-Pace for providing us with historical data for French municipalities. Thanks also to Jason Faberman, Megan MacGarvie, Stuart Rosenthal, Will Strange, AEA, NARSA and NBER conference participants, and more particularly to the editor, Ed Glaeser. Financial support from the Centre National de la Recherche Scientifique (Combes) and from the Canadian Social Science and Humanities Research Council (Duranton) is gratefully acknowledged.

Figure 1. Productivity and employment density in France



(a) Wages and employment density

(306 employment areas, 1976-1996 average)



(b) TFP (Olley-Pakes) and employment density

(306 employment areas, 1994-2002 average)

Source: DADS, BRN, RSI, SIREN and authors' calculations. All variables are centred around their mean. The R-squared is 56% in panel (a) and 61% in panel (b). See the rest of the paper for the details of the calculations.

1. Introduction

Productivity and wages are higher in larger cities and denser areas. This fact was first noted by Adam Smith (1776) and Alfred Marshall (1890) and has been confirmed by the modern empirical literature on this topic (see Rosenthal and Strange, 2004 for a review). The measured elasticity of local productivity with respect to employment density is typically between 0.04 and 0.10. We confirm this on French data. Figure 1 (a) plots mean log wages against employment density over 1976-1996 for 306 French employment areas. The measured density elasticity of wages is 0.05. Figure 1 (b) conducts a similar exercise using log TFP for the same 306 employment areas over 1994-2002. The measured density elasticity of TFP is 0.04.

To draw inference from figure 1, two fundamental identification problems must be dealt with. First, density and measures of productivity (wage or TFP) may be simultaneously

determined. This could happen because more productive places tend to attract more workers and, as a result, become denser. An alternative explanation, albeit equivalent from an econometric perspective, is that there may be a missing local variable that is correlated with both density and productivity. We refer to this issue as the ‘endogenous quantity of labour’ problem. Since Ciccone and Hall (1996), a standard way to tackle this problem is to use instrumental variables (IV).

The second major identification problem is that more productive workers may sort into denser areas. This may occur for a variety of reasons. For instance, skilled workers may have a stronger preference for high density, perhaps because density leads to better cultural amenities. Alternatively, the productivity benefits of high density may be stronger for skilled workers. These explanations suggest that it is not only density that we expect to be simultaneously determined with productivity but also the characteristics of the local workforce. To make matters worse, we expect characteristics that are not usually observed by the statistician such as ambition or work discipline to matter and be spatially unevenly distributed. For instance, French university professors may have similar observable characteristics everywhere but a disproportionate fraction of the better ones are working in or around Paris. We refer to this problem as the ‘endogenous quality of labour’ problem. Since Glaeser and Maré (2001), a standard way to tackle this problem is to use the longitudinal dimension of the data.

One may also be concerned that density affects productivity in a myriad of ways, direct and indirect (see Duranton and Puga, 2004 for a review). Denser markets allow for a more efficient *sharing* of indivisible facilities (e.g., local infrastructure), risks, and the gains from variety and specialisation. Next, denser markets also allow for a better *matching* between employers and employees, buyers and suppliers, partners in joint-projects, or entrepreneurs and financiers. This can occur through both a higher probability of finding a match and a better quality of matches when they occur. Finally, denser markets can facilitate *learning* about new technologies, market evolutions, or new forms of organisation. Some of these mechanisms (e.g., matching) may have instantaneous effects while others (e.g.,

learning) may take time to materialise. ¹

Our paper addresses the issues of endogenous quantity and endogenous quality of labour. We do not attempt to distinguish between the different channels through which density could affect productivity and only aim at estimating a total net effect of density on wages. To deal with the endogenous quantity of labour problem we take an IV approach using both history and geology as sources of exogenous variation for population. To deal with the endogenous quality of labour problem, we proceed as in Combes, Duranton, and Gobillon (2008a) and use the longitudinal dimension of extremely rich wage data. We impose individual fixed effects and local time-varying fixed effects in a wage regression. This allows us to separate local from individual effects and reconstruct some local wages net of individual observed and unobserved effects. Note that both approaches are necessary to identify the effect of density on productivity. Neither approach on its own would be sufficient.

Our main results are the following. The raw elasticity of mean wages to density is slightly below 0.05. Controlling only for the endogenous quantity of labour bias lowers this estimate to around 0.04. Historical and geological instruments lead to roughly the same answer. Controlling only for the endogenous quality of labour bias yields an even lower density elasticity of 0.033. Controlling for both source of biases leads to a coefficient of 0.027. When we also control for the fact that agglomeration economies spill over the spatial units boundaries, our preferred estimate for the elasticity of wages to local density stands at 0.02. These results are broadly confirmed when we use an alternative measure of productivity, TFP, rather than wages.

We draw a number of conclusions from this work. First, even though we control for two major sources of bias we still find evidence of small but significant agglomeration effects. Second, the sorting of workers across places is a quantitatively more important issue than their indiscriminate agglomeration in highly productive locations. Third, the importance of unobserved labour quality implies that wages should be favoured over TFP and other productivity measures since wage data are our main hope to deal with unobserved worker

characteristics.

The rest of this chapter is as follows. Section 2 provides a simple model of productivity and wages in cities and discusses the two main estimation issues. Section 3 presents the wage data and our approach to the endogenous quality of labour bias. Section 4 presents our instruments and discusses the details of our instrumentation strategy. Our results for wages are presented in section 5 while those for productivity follow in section 6. Finally, section 7 concludes.

2. Identification issues when estimating agglomeration effects

We consider a simple theoretical model of the relationship between local characteristics and wages or productivity. Take a competitive firm i operating under constant returns to scale. Its output y_i depends on the amounts of capital k_i and labour l_i it uses and its total factor productivity A_i :

$$y_i = A_i k_i^\alpha l_i^{1-\alpha}. \quad (1)$$

If all firms face the same interest rate r , the first-order conditions for profit maximisation imply that the wage *rate* is given by:

$$w_i = (1 - \alpha) \left(\frac{\alpha}{r} \right)^{\alpha/(1-\alpha)} A_i^{1/(1-\alpha)}. \quad (2)$$

Taking logs directly leads to:

$$\ln w_i = \text{Constant} + \frac{1}{1 - \alpha} \ln A_i. \quad (3)$$

The whole focus of the agglomeration literature is then on how the local characteristics of area a where firm i is located determine productivity. We assume that TFP depends on a vector of local characteristics X_a and (observed and unobserved) firm characteristics μ_i :

$$\ln A_i = X_{a(i)} \varphi + \mu_i. \quad (4)$$

Inserting into (3) implies:

$$\ln w_i = \text{Constant} + \frac{1}{1 - \alpha} \left(X_{a(i)} \varphi + \mu_i \right). \quad (5)$$

This equation can in principle be estimated using wage data and local characteristics. An alternative strategy is to insert (4) into (1), take logs, and estimate:

$$\ln y_i = \alpha \ln k_i + (1 - \alpha) \ln l_i + X_{a(i)} \varphi + \mu_i. \quad (6)$$

Hence both wage and firm level (TFP) data can be used to estimate the coefficients of interest in the vector φ or $\varphi/(1 - \alpha)$.² The first identification problem when estimating (5) or (6) is that the effect of local characteristics, $X_{a(i)}$, on wages and productivity may not be causal (endogenous quantity of labour bias). In other words, unobserved local determinants of firm productivity that are part of the error term μ_i may well be correlated with $X_{a(i)}$. Second, local characteristics of workers that are not observed, and therefore not included in $X_{a(i)}$, may not be comparable across areas (endogenous quality of labour bias). Again this creates some correlation between μ_i and $X_{a(i)}$.³

To provide further justification for equations (5) and (6) and clarify some issues regarding endogenous quantity of labour bias, we note that the literature on the microfoundations of agglomeration (e.g. Duranton and Puga, 2004) typically leads to equilibria where the wage in area a , w_a , depends on a local productivity shifter, B_a , and the local workforce, N_a :

$$w_a = B_a N_a^\delta, \quad (7)$$

This equation is consistent with (5) when agglomeration effects are such that $A_i^{1/(1-\alpha)} = B_a N_a^\delta e^{\mu_i}$. B_a can be viewed as a short-hand for all variables other than N_a in X_a . With $\delta > 0$, individual wages increase with N_a . If N_a is exogenously determined, it can be part of the vector of local characteristics X_a and δ be appropriately estimated with OLS. Following Roback (1982) and subsequent literature, we may assume instead that workers choose their city of residence. This choice is determined, through utility maximisation, by the difference between the wage and the local cost of living:

$$U_a = w_a - C_a N_a^\lambda, \quad (8)$$

In any city, the cost of living increases with the city workforce and depends on other characteristics such as amenities which have utility costs (and benefits). A spatial equilibrium equalises utility across cities.

Assuming $\lambda > \delta$ and normalising equilibrium utility to zero, the above yields:

$$N_a = \left(\frac{B_a}{C_a} \right)^{\frac{1}{\lambda-\delta}} \quad \text{and} \quad w_a = B_a^{\frac{\lambda}{\lambda-\delta}} C_a^{\frac{-\delta}{\lambda-\delta}}. \quad (9)$$

At one extreme if there are no differences in productivity across cities other than due to population difference (i.e., $B_a = B$) and only costs vary, the OLS coefficient on log workforce when regressed against log wage will be (appropriately) δ . In the opposite case where costs are the same everywhere ($C_a = C$) and only productivity vary, the regression will estimate instead λ . In the general case where both costs and productivity vary, the estimated coefficient on log workforce will be between δ and λ .⁴ The intuition for that result is that if the variation in local workforce comes solely from local costs, it is ‘exogenous’ and the proper coefficient is estimated. If instead the workforce is determined by the variation in productivity, wages in equilibrium only reflect the extent to which local costs increase with the size of the workforce. We need to keep this point in mind for our estimation strategy below.

This problem actually goes deeper than that. Our model considers only two factors of production, labour and physical capital, which is mobile and whose price can reasonably be taken to be constant everywhere. Then the term associated with its price r enters the constant and raises no further problem. However, land may also enter as a factor of production. Unlike for capital, the price of land varies across areas. Following again Roback (1982), we expect better consumption amenities to draw in more population and imply higher land prices. Firms will thus use less land. In turn, this lowers the marginal product of labour when land and labour are imperfect substitutes in the production function (as might be expected). Put differently, non-production variables may affect both population patterns and be capitalised into wages. To deal with this problem, we could attempt to control for local variables that directly affect consumer utility and thus land prices.

However, our range of controls is limited and we are reluctant to use a broad range of local amenities since many of them are likely to be simultaneously determined with wages.

Faced with reverse causality and missing variables that potentially affect both wages and the density of employment, our strategy is to rely on instrumental variables.⁵ Hence, we are asking to our instrument to deal with both the reverse-causality problem and the missing variable issue highlighted here.⁶

Turning to the endogenous quality of labour bias, note that the quantity derived in equation (2) and used throughout the model is a wage rate per efficiency unit of labour. Even if we are willing to set aside the issue that different types of labour should be viewed as different factors of production, not all workers supply the same number of efficiency units of labour per day. However, the data for individual workers is about their daily earnings, that is their wage *rate* times the efficiency of their labour. For worker j employed by firm i it is convenient to think of their earnings as being $W_j = w_{i(j)} \times s_j$ where their level of skills s_j is assumed to map directly into the efficiency of their labour. Hence, individual skills must be conditioned out from the regression to estimate (5) properly. Otherwise, any correlation between local characteristics and the skills of the local workforce will lead to biased estimates for agglomeration effects. Put differently, the quality of workforce in an area is likely to be endogenous. Previous work on French data (Combes *et al.*, 2008a) leads us to believe that this is a first-order issue.

To deal with this problem of endogenous labour quality, a number of approaches can be envisioned. The first would be to weigh the workforce by a measure of labour quality at the area level and try to instrument for labour quality just like we instrument for labour quantity. Instruments for labour quality are very scarce. The only reasonable attempt is by Moretti (2004) who uses land-grant colleges in US cities to instrument for the local share of workers with higher education. In any case, this is unlikely to be enough because we also expect unobservables such as ambition or work discipline to matter and be spatially unevenly distributed (Bacolod, Blum, and Strange, 2009).

To tackle sorting heads on, previous literature has attempted to use area characteristics

at a different level of spatial aggregation. For instance, Evans, Oates, and Schwab (1992) use metropolitan characteristics to instrument for school choice while Bayer, Ross, and Topa (2008) use location at the block level and assume an absence of sorting conditional on neighbourhood choice.⁷ In our data, although we know location at the municipal level, we are loathe to make any strong spatial identifying assumption of that sort. A more satisfactory alternative would be to estimate a full system of equations, modelling explicitly location choice. Unfortunately, due to both the difficulty of finding meaningful exclusion restrictions and the complications introduced by the discrete nature of the choice among many locations, this is a difficult exercise. Dahl (2002) proposes a new approach to this problem but this can be applied to cross-section data only.

The last existing approach is to use the longitudinal dimension of the data as in Glaeser and Maré (2001), Moretti (2004) and Combes *et al.* (2008a). This is the approach we follow. The details of our methodology are described in the next section.

3. Sorting and wage data

Choice of spatial zoning, sectoral aggregation, and explanatory variables

The choice of geographical units could in principle be of fundamental importance. With the same data, there is no reason why a partial correlation that is observed for one set of spatial units should also be observed for an alternative zoning. In particular, the shape of the chosen units may matter. However, Briant, Combes, and Lafourcade (2007) compare the results of several standard exercises in spatial economics using both official French units, which were defined for administrative or economic purposes, and arbitrarily defined ones of the same average size (i.e., squares on a map). Their main finding is that to estimate agglomeration effects, the localisation of industries, and the distance decay of trade flows across areas, the *shape* of units makes no difference.

With respect to our choice of units, we opt for French employment areas ('zones d'emploi'). Continental France is fully covered by 341 employment areas, whose bound-

aries are defined on the basis of daily commuting patterns. Employment areas are meant to capture local labour markets and most of them correspond to a city and its catchment area or to a metropolitan area. This choice of relatively small areas (on average 1,500 km²) is consistent with previous findings in the agglomeration literature that agglomeration effects are in part very local (Rosenthal and Strange, 2004). Nevertheless, we are aware that *different spatial scales* may matter with respect to agglomeration effects (see Briant *et al.*, 2007, and previous literature). We need to keep this important issue in mind when deciding on a specification.

Turning to the level of sectoral aggregation, a key question regards whether the benefits from agglomeration stem from the size of the overall local market (*urbanisation economies*) or from geographic concentration at the sector level (*localisation economies*). Although we want to focus on overall scale effects, sector effects cannot be discarded. Previous results for France suggest that they matter although they are economically far less important than overall scale effects (Combes *et al.*, 2008a). In the following, we work at the level of 114 three-digit sectors.⁸

The main explanatory variable we are interested in is employment density. It is our favourite measure of local scale. Since Ciccone and Hall (1996), density-based measures have often been used to assess overall scale effects. Their main advantage compared to alternatives measures of size such as total employment or total population is that density-based measures are more robust to the zoning. In particular, Greater Paris is divided into a number of employment areas. The true economic scale of these Parisian employment areas is much better captured by their density than any absolute measure of employment.

To repeat, French employment areas are relatively small and determined by commuting patterns. On the other hand, input-output linkages may not be limited by commuting distances. Hence we expect some agglomeration effects to take place at a scale larger than employment areas. There is by now a lot of evidence that the market potential of an area matters (Head and Mayer, 2004). In some regressions, we thus also consider the market potential of an area that we define as the sum of the density of the other areas weighted

by the inverse distance to these areas.⁹ Experimenting with other measures leads to very similar results.

Main wage data

We use an extract from the Déclarations Annuelles des Données Sociales (DADS) or Annual Social Data Declarations database from the French statistical institute (INSEE). The DADS are collected for pension, benefits and tax purposes. Establishments must fill a report for each of their employees every year. An observation thus corresponds to an 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.

For each observation, we know the age, gender, and occupation at the two-digit level. Except for a small sub-sample, education is missing. We also know the number of days worked but not hours for all years so that we restrict ourselves to full-time employees for whom hours are set by law. For earnings, we focus on total labour costs deflated by the French consumer price index. We refer to the real 1980 total labour cost per full working day as the wage. The data also contains basic establishment level information such as location and three-digit sector.

The raw data contains 19,675,740 observations between 1976 and 1996 (1981, 1983, and 1990 are missing). The details of the cleaning of the data is described in Combes *et al.* (2008a). After selecting only full-time workers in the private sector, excluding outliers, dumping a number of industries with reporting problems, and deleting observations with coding problems, we end up with 8,826,422 observations. For reasons of computational tractability, we keep only six points in time (every four years: 1976, 1980, 1984, 1988, 1992, and 1996), leaving us with 2,664,474 observations.

Using the above data, we can construct a number of variables for each year. Our main explanatory variable, employment density can be readily calculated from the data.

¹⁰ So can market potential. For each area and sector, we also compute the number of

establishments, the share of workers in professional occupations, and the share of the sector in local employment. As controls we also use three amenities variables. These amenities variables are the share of population located on a sea shore, mountains, and lakes and waterways. These variables come from the French inventory of municipalities. We aggregate them at the level of employment areas, weighting each municipality by its population. ¹¹ Table 1 below reports a number of descriptive statistics for French employment areas.

Three wages

The simplest way to implement equation (5) is to compute the mean wage for each area and year, and take its log:

$$W_{at}^1 \equiv \ln \bar{w}_{at} \equiv \ln \left(\frac{1}{N_{at}} \sum_{j \in (a,t)} w_{jt} \right). \quad (10)$$

where w_{jt} is the wage of worker j and year t and N_{at} the number of workers in area a and year t .

We can then use W_{at}^1 as dependent variable to be explained by local employment density and other local characteristics in equation (14). Using a simple log mean like W_{at}^1 throws a number of problems. First, when using mean wages we do nothing regarding the endogenous quality of labour bias. Second, we do not condition out sector effects. ¹²

To deal with these two problems, a first solution is to use all the available observables about workers and proceed as follows. We first compute a mean wage per employment area, sector, and year:

$$\bar{w}_{ast} \equiv \frac{1}{N_{ast}} \sum_{j \in (a,s,t)} w_{jt}. \quad (11)$$

This wage can then be regressed on a number of (mean) characteristics of the workers and the local sector. More specially we can estimate the following first step regression:

$$\ln \bar{w}_{ast} = W_{at}^2 + \gamma_s + X_{ast} \varphi + \epsilon_{ast}. \quad (12)$$

In this equation, γ_s is a sector dummy, and X_{ast} is a set of characteristics for sector s in area a and year t and the workers employed therein. To capture sector effects we use in X_{ast} the

(log) share of local employment in sector s and the (log) number of local establishments in this sector. Also in X_{ast} , the mean individual characteristics are the age, its square, and the shares of employment in each of 6 skill groups.¹³ In equation (12), the coefficient of interest is W_{at}^2 , a fixed effect for each employment area and year. When estimating (12), all local sector and mean individual characteristics are centred and the observations are weighted by the number of workers in each cell to avoid heteroscedasticity.

The coefficients W_{at}^2 can, in a second step, be regressed on local employment density and other local characteristics as stipulated by equation (5). While further details and justifications about the estimation of (12) are given in Combes *et al.* (2008a), three important issues need to be briefly discussed. First, the approach described here first estimates local fixed effects before using them as dependent variable in a second step. We prefer this two-step approach to its one-step counterpart for reasons made clear below.

Next, estimating (12) with OLS may condition out sectoral effects but it does not take care of the possible simultaneity between mean sector wages and local sector characteristics. A high level of specialisation in a certain sector may induce high wages in this sector. Alternatively high local wages may simply be a reflection of strong local advantage also leading to a high level of specialisation. We acknowledge this concern at the sector level but we do not deal with it. The main reason is that the coefficients for local specialisation and the number of establishments, although significant, only explain a very small part of the variation in (12) (Combes *et al.*, 2008a).

Finally, controlling for observable labour market characteristics including one-digit occupational categories (for lack of control for education) attenuates concerns about the endogenous quality of labour bias. However, they do not eradicate them entirely.

A more powerful way to deal with the endogenous quality of labour bias is to estimate:

$$\ln w_{jt} = W_{a(j)t}^3 + \gamma_{s(j)t} + X_{a(j)s(j)t}^1 \varphi_{s(j)t}^1 + X_{jt}^2 \varphi^2 + \theta_j + \epsilon_{jt}. \quad (13)$$

This equation is estimated at the level of individual workers and contains a worker fixed effect θ_j which controls for all fixed individual characteristics.¹⁴ The use of individual data also allows us to control for individual characteristics X_{jt}^2 (age and its square) sep-

arately from (centred) local industry characteristics X_{ast}^1 . The latter contain the share of local employment of the sector, the local number of firms in the sector, and the local share of professional workers. The coefficient of interest in equation (13) is the wage index W_{at}^3 for each area and year after conditioning out sector effects, observable time-varying individual characteristics, and all fixed individual characteristics. If we ignore again the possible endogeneity of local sector characteristics, the main issue when estimating (13) regards the endogeneity of location or sector choices. However, because we have sector effects and time-varying local effects, W_{at}^3 , problems only arise when we have spatial or sector sorting based on the worker-specific errors. In particular, there is no bias when sorting is based on the explanatory variables, *including individual, area-year, and industry fixed effects*. More concretely, there is a bias when the location decision is driven by the exact wage that the worker can get at locations in a given year but there is no bias when workers base their location decision on the average wage of other workers in an area and their own characteristics, i.e., when they make their location decision on the basis of their expected wages. See Combes *et al.* (2008a) for further discussion.

Note that we prefer this two-step approach, which first estimates (12) or (13) before regressing W_{at}^2 or W_{at}^3 on local characteristics, to its corresponding one-step counterpart for three reasons. First, we can properly take into account correlations between area-sector variables and error terms at the area level. Second, a two-step approach allows us to account for area-specific error terms when computing the standard errors for the coefficients we estimate. Doing so is important because Moulton (1990) shows that standard errors can be seriously biased otherwise. Accounting for area-specific errors with a one-step approach is not possible given that workers can move across areas. Third, we can conduct a variance decomposition for the second stage.¹⁵

Finally, to avoid identifying out of the temporal variation, we average the three wage variables and all the explanatory variables across the six years of data we use.¹⁶ Before turning to our results, it is interesting to note that these three local wage variables are strongly correlated with one another. The correlation between W^1 and W^2 is 0.87, the

Table 1. Summary statistics for our main variables (averages across 306 employment areas).

	Mean	Std. dev.
Mean wage (1976–1996, in 1980 French Francs, per day)	207.9	15.8
W^1	5.3	0.074
W^2	5.2	0.070
W^3	-0.04	0.049
Employment density (workers per sq. km)	64.4	543.0
ln employment density	2.4	1.2
Market potential (workers km per sq. km)	108.1	139.9
ln market potential	4.4	0.7
1831 Urban population density (inh. per sq. km)	38.2	419.8
1881 Urban population density (inh. per sq. km)	106.8	1232.3
Sea (average % municipalities on a coast line)	8.8	21.1
Lake (average % municipalities on a lake)	17.2	12.9
Mountain (average % municipalities on a mountain)	9.8	19.7

Source: DADS for the first eight lines, historical censuses for the next two, and 1988 municipal inventory for the last three. For sea, lake and mountain, we have for each employment area the percentage of municipalities on a coast, with a lake, or on a mountain. We average this quantity across employment areas.

correlation between between W^1 and W^3 is 0.81, while the correlation between W^2 and W^3 is 0.91. Table 1 reports a number of descriptive statistics for French employment area.

4. Instruments

That the estimation of agglomeration economies could be plagued by simultaneity was first articulated by Moomaw (1981). To preview of our IV approach, we note first that using historical variables such as long lags of population density to instrument for the size or density of local population is standard since Ciccone and Hall's (1996) pioneering work.

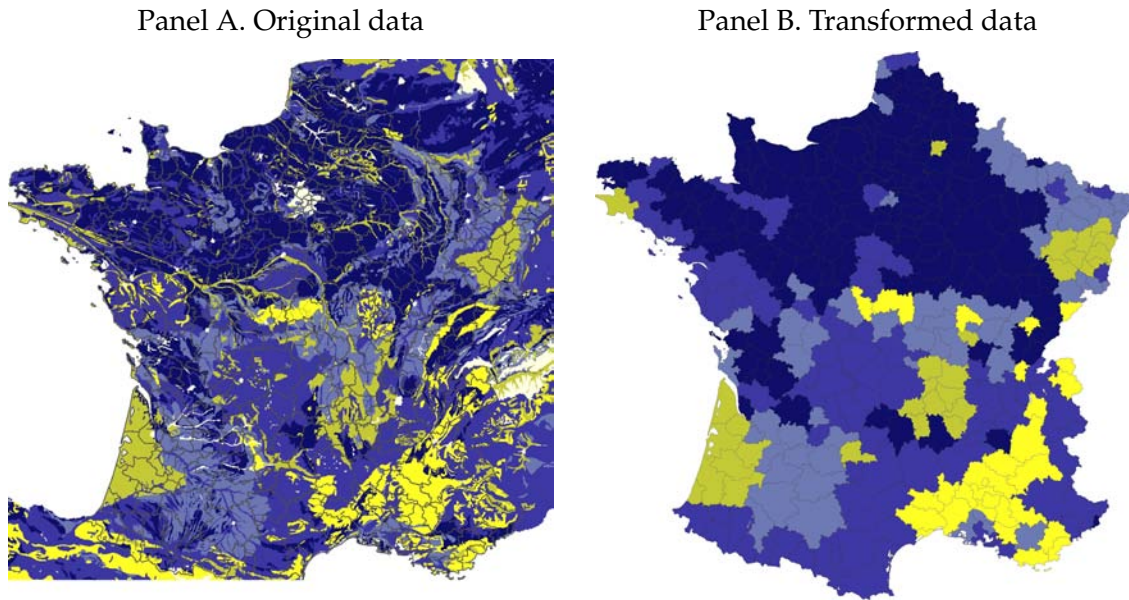
To the extent that (i) there is some persistence in the spatial distribution of population and (ii) the local drivers of high productivity today differ from those of a long gone past, this approach is defensible. An alternative strategy is to use the nature of soils since geology is also expected to be an important determinant of settlement patterns. Some soils are more stable than others and can thus support a greater density of people. More fertile lands may have also attracted people in greater number, etc. To the extent that geology affects the distribution of population (i.e., labour supply) and does not otherwise cause productivity (i.e., labour demand) because fertile lands are no longer a relevant driver of local wealth, it can provide reasonable instruments to explain the distribution of employment. Except by Rosenthal and Strange (2008) in a slightly different context, geology has not been used to instrument for the distribution of population.

Description of the instruments

Our first set of instruments is composed of historical populations from early French censuses. For 26 French censuses prior to our earliest year of data (1976) we know the 'urban' population for each municipality. Among available censuses we choose the earliest one from 1831 and another from 1881, 50 years later.¹⁷ We also experimented with other years. Unfortunately, urban population in historical censuses is only reported above a threshold of 5,000. For 1831, there are 35 employment areas for which no municipality had an urban population above 5,000. A small majority of them are rural areas while the others are densely populated employment areas with strong municipal fragmentation. We think of this as being measurement error. To minimise weak instrument problems, we drop these 35 employment areas.

Our second group of instruments is composed of geological variables from the European Soil Database (ESDB) compiled by the European Soil Data Centre. The data originally come as a raster data file with cells of 1 km per 1 km. We aggregated it at the level of each employment area.¹⁸ Given that soil characteristics are usually discrete, we use the value that appears most often in each area. To take an illustrative example,

Figure 2. Geological characteristics: Water capacity of the subsoil



Source: European Soil Database. Panel A represents the initial raster data. Panel B represents the transformed version of the same data after imputation of the missing values for 7 employment areas in Greater Paris. In both panels, the darkest shade of grey corresponds to 'very high' (i.e., above 190 mm), the second darkest shade corresponds to 'high' (between 140 and 190 mm) followed by 'medium' (100 – 140 mm), 'low' (5 – 100 mm), and 'very low' (0 – 5 mm). Missing values in panel A (around Paris) are in white.

the initial and transformed data for the water capacity of the subsoil are represented in figure 2. For a small number of densely populated employment areas in Greater Paris, the most important category is sometimes "missing". When this is the case, we turn to the second most important category. In the rare instances where the information is missing from all the pixels in an employment area, we impute the value of a neighbouring area (chosen because it takes similar values for other soil characteristics). For instance, the water capacity of the subsoil in Central Paris is missing. We impute the value of that of its close neighbour Boulogne-Billancourt.

In total, we generate 12 variables from the ESDB.¹⁹ The first four describe the nature of the soils according to the mineralogy of their subsoil (3 categories) and topsoil (4

categories) and the nature of the dominant parent material at a broad level of aggregation (6 categories) and at a finer level (with 20 categories). More precisely, the mineralogy variables describe the presence of various minerals in the topsoil (the first layer of soil, usually 5 to 15 cm deep) and the subsoil (the intermediate layer between the topsoil and the bedrock). The dominant parent material of the soil is a description of the underlying geological material (the bedrock). Soils usually get a great deal of structure and minerals from their parent material. The more aggregate dominant parent material variable (in 6 categories) contains entries such as igneous rocks, glacial deposits, or sedimentary rocks. Among the latter, the detailed version of the same variable (with 20 categories) distinguishes between calcareous rocks, limestone, marl, and chalk.

The next seven geological characteristics document various characteristics of the soil including the water capacity of the subsoil (5 categories) and topsoil (3 categories), depth to rock (4 categories), differentiation (3 categories), erodibility (5 categories), carbon content (4 categories), and hydrogeological class (5 categories). Except for the hydrogeological class which describes the circulation and retention of underground water, the meaning of these variables is relatively straightforward. Finally, we create a measure of local terrain ruggedness by taking the mean of maximum altitudes across all pixels in an employment area minus the mean of minimum altitudes. This variable thus captures variations of altitude at a fine geographical scale.

Relevance of the instruments

Following equations (5) and (6), the specifications we want to estimate are:

$$\ln W_a = \text{Constant} + X_a \varphi^W + \mu_a^W \quad (14)$$

and

$$\ln TFP_a = \text{Constant} + X_a \varphi^{TFP} + \mu_a^{TFP}, \quad (15)$$

where μ_a^W and μ_a^{TFP} are the errors terms for the wage and TFP equations. The vector of dependent variables X_a contains the three amenity variables discussed above, (log) em-

ployment density, and sometimes market potential. These last two variables are suspected of being simultaneously determined with wages and TFP.

Estimating the effect of employment density and market potential on local wages and productivity using instrumental variables can yield unbiased estimates provided that the instruments satisfy two conditions, relevance and exogeneity. Formally, these conditions are

$$Cov(Density_a, Z_a | \cdot) \neq 0, \quad Cov(MarketPotential_a, Z_a | \cdot) \neq 0, \quad (16)$$

and

$$Cov(\mu_a^X, Z_a) = 0 \quad \text{for } X = W \text{ and } X = TFP, \text{ respectively,} \quad (17)$$

where Z denotes the set of instruments. We begin by discussing the ability of our instruments to predict contemporaneous employment density and market potential conditionally to the other controls.

The stability of population patterns across cities over time is a well documented fact (see Duranton, 2007, for a recent discussion). This stability is particularly strong in France (Eaton and Eckstein, 1997). The raw data confirm this. Table 2 presents pairwise correlations between our four historical instruments and current employment density and market potential.²⁰ For the sake of comparison with geology variables below, we also report the R-squareds of the corresponding univariate regressions. We can see that the log urban population densities of 1831 and 1881 are good predictors of current employment density. Past market potentials computed from 1831 and 1881 urban populations also predict current market potential extremely well.

Turning to geological characteristics, we expect the nature of soils and their characteristics to be fundamental drivers of population settlements. Soil characteristics arguably determine their fertility. Since each soil characteristic is described by several discrete variables, it is not meaningful to run pairwise correlations as with historical variables. Instead, table 3 reports the R-squared when regressing employment density and market potential against various sets of dummies for soil characteristics. The results show that some geological characteristics like the dominant parent material or the depth to rock

Table 2. R-squareds of univariate regressions and pairwise correlations : historical vs. density and market potential (1976-1996)

	ln(employment density)	ln(market potential)
ln(1831 density)	0.57 (0.75)	0.05 (0.24)
ln(1881 density)	0.78 (0.88)	0.10 (0.33)
ln(1831 market potential)	0.21 (0.46)	0.96 (0.98)
ln(1881 market potential)	0.22 (0.47)	0.99 (0.99)

306 observations.

Adjusted R-squared in plain text and pairwise correlations between parentheses.

have good explanatory power. Other soil characteristics such as their mineralogy or their carbon content are less powerful predictors of current population patterns. Note also that soil characteristics tend to be better at explaining the variations of market potential than employment density. This is not surprising since most soil characteristics vary relatively smoothly over fairly large spatial scales while variations in density are more abrupt and take place at smaller spatial scales.

While the correlations and R-squareds reported in tables 2 and 3 are interesting, equation (16) makes clear that the relevance of an instrument depends on the *partial* correlation of the instrumental variables and the endogenous regressor. To assess these partial correlations, table 4 presents the results of OLS regressions of log density on our instrumental variables and controls. Table 5 reports results for a similar exercise with market potential.

Column 1 of table 4 examines the partial correlation between employment density and 1831 population density while conditioning out amenities (sea, lake, and mountain). Column 2 performs a similar regression using 1881 instead of 1831 population density. In both columns, the coefficient on past density is highly significant and close to unity. In columns 3 to 9, we regress contemporaneous employment density on a series of soil dummies concerning their mineralogy, dominant parent material, water capacity, carbon

Table 3. R-squareds when regressing density and market potential on soil characteristics

	ln(emp. density)	ln(market pot.)
Subsoil mineralogy (2 dummies)	0.02	0.06
Topsoil mineralogy (3 dummies)	0.02	0.06
Dominant parent material (5 dummies)	0.11	0.31
Dominant parent material (19 dummies)	0.13	0.48
Topsoil water capacity (2 dummies)	0.03	0.23
Subsoil water capacity (3 dummies)	0.01	0.32
Depth to rock (3 dummies)	0.10	0.35
Soil differentiation (2 dummies)	0.07	0.19
Erodibility (4 dummies)	0.04	0.19
Carbon content (3 dummies)	0.04	0.04
Hydrogeological class (4 dummies)	0.01	0.04
Hydrogeological class (4 dummies)	0.05	0.10

Adjusted R-squareds. 306 observations.

Table 4. First stage: Density

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
ln(1831 density)	0.906								
	(0.046) ^a								
ln(1881 density)		0.924							
		(0.030) ^a							
Ruggedness			- 0.710						
			(0.224) ^a						
Subsoil mineralogy	N	N	N	Y	N	N	N	N	N
Dominant parent material (6 categories)	N	N	N	N	Y	N	N	N	N
Subsoil water capacity	N	N	N	N	N	Y	N	N	N
Soil carbon content	N	N	N	N	N	N	Y	N	N
Depth to rock	N	N	N	N	N	N	N	Y	N
Soil differentiation	N	N	N	N	N	N	N	N	Y
R-squared	0.58	0.78	0.07	0.07	0.17	0.06	0.10	0.15	0.11
F-test (H_0 – All instruments zero)	395.7	1018.8	10.0	5.5	9.1	1.7	6.5	12.6	12.3
Partial R-squared	0.57	0.77	0.03	0.04	0.13	0.02	0.06	0.11	0.08

Dependent variable: ln(employment density). 306 observations for each regression.
 All regressions include a constant and three amenity variables (sea, lake, and mountain).
 Standard errors in parentheses. *a, b, c*: corresponding coefficient significant at 1, 5, 10%.

Table 5. First stage: Market potential

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
ln(1831 market pot.)	1.026								
	(0.012) ^a								
ln(1881 market pot.)		0.970							
		(0.007) ^a							
Ruggedness			- 0.339						
			(0.111) ^a						
Subsoil mineralogy	N	N	N	Y	N	N	N	N	N
Dominant parent material (6 cat.)	N	N	N	N	Y	N	N	N	N
Subsoil water capacity	N	N	N	N	N	Y	N	N	N
Soil carbon content	N	N	N	N	N	N	Y	N	N
Depth to rock	N	N	N	N	N	N	N	Y	N
Soil differentiation	N	N	N	N	N	N	N	N	Y
R-squared	0.97	0.99	0.23	0.24	0.43	0.41	0.28	0.44	0.31
F-test (H_0 – All instruments zero)	7106.47	21503.0	9.4	7.3	23.1	26.3	11.3	41.2	24.0
Partial R-squared	0.96	0.99	0.03	0.05	0.28	0.26	0.10	0.29	0.14

Dependent variable: ln(market potential). 306 observations for each regression.

All regressions include a constant and three amenity variables (sea, lake, and mountain).

Standard errors in parentheses. *a, b, c*: corresponding coefficient significant at 1, 5, 10%.

content, depth to rock, and soil differentiation. For lack of space, we do not report all the coefficients but it must be noted that at least one dummy is significant at 5% in each regression.

The comparison of R-squareds in columns 1-2 versus 3-9 shows immediately that long lags of population density explain a greater share of the variations in contemporaneous employment density than soil characteristics. To make a more formal assessment of the relevance of our instruments we turn to the weak instrument tests developed by Stock and Yogo (2005).²¹ Table 4 presents the relevant F-statistics. The two lagged density instruments in columns 1 and 2 have F-statistics close to 400 and 1000, respectively. This makes them very strong in light of the critical values reported by Stock and Yogo (2005) in their tables 1-4. The soil instruments are weaker by comparison and fall below the critical values of Stock and Yogo (2005) with TLS. To avoid the pitfalls of weak instruments, a number of possible strategies can be envisioned. First, it would be possible to increase the strength of some soil instruments by considering only the more relevant dummies and dropping insignificant ones. In absence of a well articulated theory of how soils affects economic development, we acknowledge an element of 'data mining' in our use of soil characteristics. We are nonetheless reluctant to push it to such extremes. Second, we experiment below with estimation strategies that are less sensitive to weak instruments such as limited information maximum likelihood (LIML) as advocated by Andrews and Stock (2007). Third, we repeat the same regressions with different sets of soil instruments and see how this affects the coefficient(s) of interest. Obtaining the same answer over and over again would be reassuring.

In table 5, we repeat the same exercise with market potential using lagged values of that variable and the same set of soil instruments as in table 4. Both historical and soil variables are much stronger instruments for market potential than for employment density. For historical variables, the reason is that market potential is computed as a weighted mean of employment density. As a result this washes out much idiosyncratic variation and naturally yields higher R-squareds. Put differently, soil variables are better replicating the

smooth evolution of market potential than the spikes of employment density. The facts that in column 1 the coefficient on 1831 market potential is essentially one and that the partial R-squared is 95% also indicate that we should not expect much difference between OLS and TSLS below.

Because both market potential and soil characteristics vary smoothly over space, one may worry that the good explanatory power of soil characteristics may be spurious. This will be the case if some large areas with particular soil characteristics spuriously overlap with areas of particularly high or low market potential. However, a detailed reading of the coefficients on soil dummies (not reported in table 5) indicates that this is not the case. For instance, areas for which the dominant parent material is conditionally associated with the lowest market potential are eolian sands, molasse (sand stone), and ferruginous residual clay. Sands, which drain very fast, and ferruginous clay, a heavy soil which does not drain at all, do not lead to very fertile soils. On the other hand, the parent materials associated with a high market potential are loess, a notably fertile type of soil, and chalk, a stable and porous soil which can be very fertile provided it is deep enough. Similarly, a high water capacity of the subsoil is associated with a higher market potential as could be expected.

Instrument exogeneity

Equation (17) gives the second condition that must be satisfied by a valid instrument: orthogonality to the error term. Intuitively, the difficulty in inferring the effect of density and market potential on wages and TFP arises because of the possibility that a missing local characteristic or some local shocks might be driving both population location and economic outcomes. To overcome this problem, we require instruments which affect wages and TFP only through the spatial distribution of population. That is, as made clear above, we need our instruments to affect the supply of labour, but not directly productivity. We now discuss the *a priori* arguments why our instruments may (or may not) satisfy this exogeneity condition.

We begin with historical variables dating back to 1831. Long-lagged values of the

same variable obviously remove any simultaneity bias caused by 'contemporaneous' local shocks. For such simultaneity to remain, we would need these shocks to have been expected in 1831 and have determined population location at the time. This is extremely unlikely. However, endogeneity might also arise because of some missing permanent characteristic that drives both past population location and contemporaneous productivity. A number of first-nature geographic characteristics such as a coastal location may indeed explain both past population location and current economic outcomes. In our regressions we directly control for a number of such first-nature characteristics (coast, mountain, lakes and waterways).

Hence, the validity of long population lags rests on the hypothesis that the drivers of population agglomeration in the past are not related to modern determinants of local productivity after controlling for first-nature characteristics of places. The case for this relies on the fact that the French economy in the late 20th century is very different from what it was in 1831. First, the structure of the French economy in the late 20th century differs a lot from that of 1831. In 1831, France was only starting its industrialisation process, whereas it is de-industrialising now. Manufacturing employment was around 3 million in 1830 against more than 8 million at its peak in 1970 and less than 6 today (Marchand and Thélot, 1997). Then, agriculture employed 63% of the French workforce against less than 5% today. Since 1831, the workforce has also doubled. Second, the production techniques in agriculture, manufacturing and much of the service industries are radically different today from what they were more than 150 years ago. With technological change, the location requirements of production have also changed considerably. For instance, the dependence of manufacturing on sources of coal and iron has disappeared. Third, the costs of shipping goods and transporting people from one location to another have declined considerably. Actually, 1831 coincides with the construction of the first French railroads. Subsequently, cars, trucks and airplanes have further revolutionised transport. At a greater level of aggregation, trade has also become much easier because of European integration over the last 50 years. Fourth, other drivers of population location

not directly related to production have changed as well. With much higher standards of living, households are arguably more willing to trade greater efficiency against good amenities (Rappaport, 2007). Some previously inhospitable parts of the French territory such as its Languedocian coast in the South have been made hospitable and are now developed, etc. Finally, since 1831, France has been ruled by, successively, a king, an emperor, and presidents and prime ministers from 5 different republics. The country also experienced a revolution in 1848, a major war with Germany in 1870, and two world wars during the 20th century.

With so much change, a good case can indeed be made that past determinants of population location are not major drivers of current productivity. As a result, historical variables are the instrument of choice for current population patterns since Ciccone and Hall (1996). They have been widely used by the subsequent literature.

Although the *a priori* case for historical instruments is powerful, nothing guarantees that it is entirely fool-proof. The fact that long lags of the population variables usually pass over-identification tests and other *ex post* diagnostics may not constitute such a strong argument in favour of their validity. Population variables are often strongly correlated with one another so that any permanent characteristics that affects both measures of past population location and contemporaneous productivity may go un-noticed due to the weak power of over-identification tests when the instruments are very similar and thus highly correlated.

We now consider geological characteristics. The *a priori* case for thinking that geological characteristics are good instruments hinges first on the fact that they have been decided mostly by nature and do not result from human activity. This argument applies very strongly to a number of soil characteristics we use. For instance, soil mineralogy and their dominant parent material were determined millions of years ago. Other soil characteristics might seem more suspect in this respect. For instance, a soil's depth to rock or its carbon content might be an outcome of human activity. In the very long-run, there is no doubt that human activity plays a role regarding these two characteristics.

Whether recent (in geological terms) economic activity can play an important role is more doubtful (e.g., Guo and Gifford, 2002). A second caveat relates to the measurement of some soil characteristics. In particular, it is hard to distinguish between a soil's intrinsic propensity to erodibility from its actual erosion (see Seybold, Herrick, and Brejda, 1999). In relation to these two worries, our wealth of soil characteristics implies that we can meaningfully compare the answers given by different soil characteristics as instruments in different regressions. We can also use over-identification tests to assess this issue more formally.

Nonetheless, that soils predate patterns of human settlement does not ensure that any soil characteristics will automatically satisfy condition (17) and be valid instruments. More specifically, we expect soil characteristics to have been a major determinant of local labour demand in the past. The main argument for the validity of geological instruments is then that soil quality is no longer expected to be relevant in an economy where agriculture represents less than 5% of employment. We also exclude agricultural activities from our data. Put differently, the case for geological characteristics relies on the fact that this important, though partial, determinant of past population location is now largely irrelevant. Hence, like with historical instruments, the *a priori* case for geological instruments is strong but there is no way to be entirely sure.

It is important to note that the cases for the validity of historical and geological variables as instruments differ. Historical variables are 'broad' determinants of current population location. Soil characteristics are narrower but more 'fundamental' determinants of current population location. Put differently, although we expect soils to have determined history, they were not the sole determinants of population patterns in 1831. Geological characteristics also explain current patterns of employment density over and above past employment density. If one group of instruments fails, it is unlikely that the second will do so in the same way. Finally, it is also important to keep in mind that these two sets of instruments can only hope to control for the endogenous quantity of labour bias. That a higher density can lead to the sorting of better workers in these areas is not taken care of

by these instruments. Put differently, we expect the endogenous quality of labour bias to remain. Moving from crude measures of wage such W^1 to more sophisticated ones (W^2 and most of all W^3) is designed to tackle this second issue.

5. Main wage results

Table 6 presents the results of three simple regressions for our three wages: W^1 , the mean local wage as computed in (10), W^2 , the wage index after conditioning out sector effects and observable individual characteristics as estimated in (12), and W^3 , the wage index from (13) which also conditions out individual fixed effects. In columns 1, 2, and 3, these three wages are regressed on log employment density controlling for three amenity variables (coast, lakes and waterways, mountain) using OLS. The measured density elasticity of mean wages is at 0.048. This is very close to previous results in the literature (Ciccone and Hall, 1996, Ciccone, 2002). Controlling for sector effects in column 2 implies a marginally higher estimate of 0.051 for the density elasticity and significantly improves the explanatory power of employment density. This suggests that, although the local characteristics of the sector of employment matter, conditioning out sector effects does not affect our estimates of the density elasticity. Controlling also for unobserved individual characteristics yields a significantly lower elasticity of 0.033. This suggests that a good share of measured agglomeration effects are in fact attributable to the unobserved characteristics of the workforce. More specifically, workers who command a higher wage on labour market sort in denser areas.

In columns 4, 5, and 6, we perform the same regressions as in columns 1, 2, 3 but we instrument employment density with 1831 urban population density. Compared to their corresponding OLS coefficients, the TSLS coefficients for employment density are marginally lower. The instrument is very strong with a first-stage F (or Cragg-Donald) statistic close to 400. In columns 6, 7, and 8, we add 1881 population density as a second instrument for employment density. The results are virtually undistinguishable from those of columns 4, 5, and 6. With two instruments, it is also possible to run Sargan

Table 6. Local wages as a function of density: OLS and historical instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
	W ¹	W ²	W ³	W ¹	W ²	W ³	W ¹	W ²	W ³
	OLS	OLS	OLS	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs
ln(density)	0.048	0.051	0.033	0.040	0.042	0.026	0.040	0.044	0.027
	(0.002) ^a	(0.002) ^a	(0.001) ^a	(0.003) ^a	(0.002) ^a	(0.002) ^a	(0.003) ^a	(0.002) ^a	(0.002) ^a
Instruments used:									
ln(1831 density)	-	-	-	Y	Y	Y	Y	Y	Y
ln(1881 density)	-	-	-	-	-	-	Y	Y	Y
First stage statistics	-	-	-	395.7	395.7	395.7	518.7	518.7	518.7
Over-id test <i>p</i> -value	-	-	-	-	-	-	0.99	0.19	0.21
R-squared	0.56	0.72	0.65	-	-	-	-	-	-

306 observations for each regression.

All regressions include a constant and three amenity variables (sea, lake, and mountain).

Standard errors in parentheses. *a*, *b*, *c*: corresponding coefficient significant at 1, 5, 10%.

tests of over-identification. They are passed in all three cases with p -values above 10%. However, we can put only a limited weight on this test because the correlation between 1881 and 1831 density is high at 0.75.

If we think of table 6 as our baseline, a number of findings are worth highlighting. The density elasticity of mean wages is 0.048 (column 1). Controlling for the endogenous quality of labour bias through a fixed-effect estimation reduces the size of the coefficient by about a third to 0.033 (column 3). Controlling for the endogenous quantity of labour bias using long historical lags as instruments reduces it by another fifth to 0.027 (column 9). Hence this table provides evidence about both the quality and quantity of labour being simultaneously determined with productivity. It also suggests that the endogenous quality of labour bias is more important than the quantity bias.

Next, table 7 reports results for number of regressions which all use geological characteristics as instruments for employment density. Following the results of table 4, we expect geological instruments to be on the weak side. Furthermore, table 5 also makes clear that geological characteristics appear to explain market potential better than employment density. Hence, we need to keep in mind that our geological instruments are correlated with a variable, market potential, that is (for the time being) missing from the regression and suspected to have an independent effect on wages. As a consequence, IV estimations that rely solely on geological characteristics may not perform very well and should be interpreted with caution.

In each of the regressions in table 7, we use two different soil characteristics. Since, except for ruggedness, each soil characteristic is documented with a series of dummy variables, we could technically run over-identification tests while instrumenting for only one characteristic. However such tests may not be economically meaningful since we would end up testing for over-identification using the particular categorisation of the ESDB. We experimented extensively with soil characteristics. The results we report in the table are representative of what is obtained using any combination of the soil characteristics listed in the table. With them, over-identification tests are usually passed. This is not the case

Table 7. Local wages as a function of density: geological instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	W^1	W^2	W^3	W^3	W^3	W^3	W^3	W^3
	TOLS	TOLS	TOLS	LIML	LIML	LIML	LIML	LIML
ln(density)	0.042	0.047	0.038	0.038	0.048	0.050	0.048	0.047
	(0.010) ^a	(0.008) ^a	(0.006) ^a	(0.006) ^a	(0.005) ^a	(0.005) ^a	(0.005) ^a	(0.005) ^a
Instruments used:								
Subsoil mineralogy	Y	Y	Y	Y	Y	N	N	N
Ruggedness	Y	Y	Y	Y	N	N	N	N
Depth to rock	N	N	N	N	Y	Y	Y	N
Soil carbon content	N	N	N	N	N	Y	N	N
Topsoil water capacity	N	N	N	N	N	N	Y	Y
Dominant parent material (6 cat.)	N	N	N	N	N	N	N	Y
First stage statistics	6.2	6.2	6.2	6.2	8.3	8.2	8.1	6.8
Over-id test <i>p</i> -value	0.99	0.90	0.67	0.67	0.45	0.12	0.34	0.15

306 observations for each regression.

All regressions include a constant and three amenity variables (sea, lake, and mountain).

Standard errors in parentheses. *a*, *b*, *c*: corresponding coefficient significant at 1, 5, 10%.

with the other soil characteristics.

More precisely, in column 1 of table 7, we regress mean wages on density and other controls using subsoil mineralogy and ruggedness as instruments for employment density. We obtain a density elasticity of 0.042, which is consistent with what we find in table 6 when we use historical variables. We repeat the same regression in columns 2 and 3 using W^2 and W^3 as dependent variables. In columns 3, the coefficient is slightly above its OLS counterpart rather than slightly below when using historical instruments. The difference is nonetheless not significant. Before going any further, note that the low first-stage statistics in columns 1-3 raises some questions about the strength of these geological instruments. With weak instruments a number of authors (e.g., Stock and Yogo, 2005) now argue for the superiority of the LIML estimator to the TSLS estimator. Column 4 of table 7 reports the LIML estimate for a specification similar to column 3. The TSLS and LIML results are the same.²²

In columns 5 to 8, we report LIML results regarding our preferred measure of wages, W^3 , for further combinations of instruments. The coefficient on employment density is positive and highly significant in all cases. However, it is above its OLS counterpart rather than below, even more so than in column 4. This discrepancy between the IV results using history in table 6 and those using geology in table 7 is due to the fact that soil variables are not only correlated with the employment density, but also with the market potential, which is missing. As a result, the density elasticities in table 7 may be biased upwards. To see this, note that in column 4 the correlation between the predicted values of employment density obtained from the instrumental regression and actual density is 0.29. The correlation between predicted density and actual market potential (omitted from the regression) is nearly as high at 0.27. In column 5, the problem is even worse since the correlation between predicted and actual density is 0.37 while the correlation between predicted density and market potential is 0.48.²³

To explore this problem further, we now consider historical and geological instruments at the same time. Table 8 reports the results for a number of regressions using both 1831

Table 8. Local wages as a function of density: historical and geological instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	W^1	W^2	W^3	W^3	W^3	W^3	W^3	W^3
	TSLS	TSLS	TSLS	GMM	TSLS	TSLS	TSLS	TSLS
ln(density)	0.040	0.042	0.027	0.027	0.027	0.027	0.027	0.027
	(0.003) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a
Instruments used:								
ln(1831 density)	Y	Y	Y	Y	Y	Y	Y	Y
Subsoil mineralogy	Y	Y	Y	Y	N	Y	N	N
Ruggedness	N	N	N	N	Y	Y	Y	N
Hydrogeological class	N	N	N	N	N	N	Y	N
Topsoil water capacity	N	N	N	N	N	N	N	Y
First stage statistics	138.7	138.7	138.7	116.2	208.7	108.1	69.8	76.2
Over-id test <i>p</i> -value	0.98	0.83	0.15	0.13	0.31	0.21	0.53	0.02

306 observations for each regression.

All regressions include a constant and three amenity variables (sea, lake, and mountain).

Standard errors in parentheses. *a*, *b*, *c*: corresponding coefficient significant at 1, 5, 10%.

density and some soil characteristics. In all cases, the instruments are strong because of the presence of 1831 density. Subsoil mineralogy (along with 1831 density) is used in columns 1 to 3 to instrument for density and explain W^1 , W^2 , and W^3 . The results are the same as those of columns 4-6 of table 6 which use only 1831 density while they differ more with those of columns 1-3 of table 7 which use subsoil mineralogy (together with ruggedness) but not 1831 density. This is unsurprising given that 1831 density is a much stronger instrument. Using a GMM-IV estimation rather than TSLS in column 4 does not change anything. Using ruggedness or hydrogeological class in columns 5-7 also implies a similar coefficient on density. With these three soils characteristics (and 1831 density) the over-identification test is passed. For the other soil characteristics however, this test is failed. An example is given in column 8 with topsoil water capacity. This is in line with the results of the previous table that a majority of soil characteristics do not give the same answer as 1831 density when used as instruments to estimate the density elasticity of wages.

To confirm that this problem is due to the strong correlation between soil characteristics and market potential, table 9 reports results for a number of regressions in which market potential is added as a control. In columns 1-3, our measures of wages W^1 , W^2 , and W^3 are regressed on density and market potential using OLS. The measured elasticity of wages with respect to market potential is between 0.01 and 0.03. It is also interesting to note that the density elasticity is slightly lower than in column 1-3 of table 6 where market potential is omitted. In columns 4 to 9, we instrument employment density with 1831 density and a range of soil characteristics. The density elasticity is very stable at 0.02 while the market potential elasticity is also very stable at 0.034. Importantly, the over-identification tests are passed (whereas they fail without market potential as a control). More generally, the over-identification test is passed for most combinations of geological instruments and 1831 density. The main systematic failure occurs when the dominant parent material dummies are used. It should be noted that 1831 density is a much stronger instrument and as a result it 'does most of the work' in generating the predicted density

Table 9. Local wages as a function of density and (exogenous) market potential: historical and geological instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
	W^1	W^2	W^3	W^3	W^3	W^3	W^3	W^3	W^3
	OLS	OLS	OLS	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs
ln(density)	0.042	0.048	0.026	0.020	0.020	0.020	0.020	0.020	0.020
	(0.003) ^a	(0.002) ^a	(0.001) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a
ln(market pot.)	0.024	0.012	0.027	0.034	0.034	0.034	0.034	0.034	0.034
	(0.006) ^a	(0.004) ^a	(0.003) ^a	(0.003) ^a	(0.003) ^a	(0.003) ^a	(0.003) ^a	(0.003) ^a	(0.003) ^a
Instruments used:									
ln(1831 density)	-	-	-	Y	Y	Y	Y	Y	Y
Subsoil mineralogy	-	-	-	Y	N	N	N	N	N
Ruggedness	-	-	-	N	Y	N	N	N	N
Subsoil water capacity	-	-	-	N	N	Y	N	N	N
Depth to rock	-	-	-	N	N	N	Y	N	N
Erodibility	-	-	-	N	N	N	N	Y	N
Soil differentiation	-	-	-	N	N	N	N	N	Y
First stage statistics	-	-	-	128.5	191.6	80.9	96.9	77.9	130.0
Over-id test <i>p</i> -value	-	-	-	0.72	0.82	0.42	0.54	0.37	0.10
R-squared	0.59	0.73	0.73	-	-	-	-	-	-

306 observations for each regression.

All regressions include a constant and three amenity variables (sea, lake, and mountain).

Standard errors in parentheses. *a, b, c*: corresponding coefficient significant at 1, 5, 10%.

at the first stage. This greater strength of past density may explain the stability of the coefficients. Nonetheless, in each of the IV regressions of table 9, at least one soil dummy (and usually more) is significant (and usually highly so). This implies that we can run meaningful over-identification tests. The fact that their p -values is usually well above 10% is strongly suggestive that 1831 density and a broad range of soil characteristics all support this 0.02 estimate for the density elasticity of wages.

Finally in table 10 we consider that market potential could also be endogenous. In columns 1-3, we use only historical instruments: 1831 and 1881 density as well as 1831 market potential. The results for W^3 in column 3 are similar to the IV results in table 8. In columns 4-9, we use 1831 employment density with, in each regression, two different soils characteristics among erodibility, carbon content, subsoil water capacity, depth to rock, ruggedness, and soil differentiation. The over-identification test is always passed in the table. Although the results are not reported here, this test is also passed for all the other pairwise combinations of these characteristics (except the combination soil differentiation and carbon content for which the test marginally fails). For our preferred concept of wage, W^3 , the coefficients on density and market potential are very stable and confirm the estimates of column 3 with historical instruments and those of the previous table where market potential is taken to be exogenous. This stability across columns 3-9 is interesting because geological variables and past density in columns 4-9 are not as strong sets of instrument as the combination of past density and past market potential. Our preferred estimate for the elasticity of wages with respect to employment density is 0.02. With respect to market potential, our preferred estimate is at 0.034.

While regressing mean wages on employment density leads to a measured elasticity of 0.05, adding further controls and correcting for the endogenous quality and quantity of labour biases bring this number down to about 0.02.

Table 10. Local wages as a function of density and (endogenous) market potential: historical and geological instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
	W^1	W^2	W^3	W^3	W^3	W^3	W^3	W^3	W^3
	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs
ln(density)	0.033	0.040	0.020	0.018	0.019	0.020	0.020	0.020	0.020
	(0.003) ^a	(0.003) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.003) ^a	(0.002) ^a
ln(market pot.)	0.034	0.020	0.034	0.048	0.039	0.036	0.036	0.033	0.034
	(0.006) ^a	(0.005) ^a	(0.003) ^a	(0.007) ^a	(0.005) ^a	(0.005) ^a	(0.006) ^a	(0.010) ^a	(0.007) ^a
Instruments used:									
ln(1831 density)	Y	Y	Y	Y	Y	Y	Y	Y	Y
ln(1881 density)	Y	Y	Y	N	N	N	N	N	N
ln(1831 m. pot.)	Y	Y	Y	N	N	N	N	N	N
Erodibility	N	N	N	Y	N	N	N	N	Y
Soil carbon content	N	N	N	Y	Y	N	N	N	N
Subsoil water capacity	N	N	N	N	Y	Y	N	N	N
Depth to rock	N	N	N	N	N	Y	Y	N	N
Ruggedness	N	N	N	N	N	N	Y	Y	N
Soil differentiation	N	N	N	N	N	N	N	Y	Y
First stage statistics	298.0	298.0	298.0	8.3	17.0	23.0	19.8	8.3	10.5
Over-id test <i>p</i> -value	0.57	0.36	0.67	0.62	0.19	0.36	0.54	0.11	0.14

306 observations for each regression.

All regressions include a constant and three amenity variables (sea, lake, and mountain).

Standard errors in parentheses. *a*, *b*, *c*: corresponding coefficient significant at 1, 5, 10%.

6. TFP

Firm and establishment data

To construct our establishment-level data, we proceed as follows. We first put together two firm-level data sets: the BRN ('Bénéfices Réels Normaux') and the RSI ('Régime Simplifié d'Imposition'). The BRN contains the balance sheet of all firms in the traded sectors with a turnover above 730,000 euros. The RSI is the counterpart of the BRN for firms with a turnover below 730,000 euros. Although the details of the reporting differs, for our purpose these two data sets contain essentially the same information. Their union covers nearly all French firms.

For each firm we have a firm identifier and detailed annual information about its output and its consumption of intermediate goods and materials. This allows us to construct a reliable measure of value added. To estimate TFP (see below), we use a measure of capital stock based on the sum of the reported book values of productive and financial assets.²⁴ We also experimented with TFP estimations using the cost of capital rather than assets values following the detailed methodology developed by Boutin and Quantin (2006).

Since firms can have many establishments at many locations, we also use the SIREN data ('Système d'Identification du Répertoire des ENtreprises'). It is an exhaustive registry of all establishments in the traded sectors. For each establishment and year, SIREN reports both a firm and an establishment identifier, a municipality code, and total employment. Note finally that BRN, RSI, and SIREN only report total employment and not hours worked.

To obtain information about hours, we return to the DADS which report them after 1993. Hence, for 1994-2002, we use another, this time exhaustive, DADS dataset.²⁵ Using the individual information about hours and two-digit occupation that this source contains, we can aggregate it at the establishment level to obtain the hours for all employees and by skill group. We emphasise this because of the suspected importance of labour quality. To avoid estimating too many coefficients for different types of labour, we aggregate two-digit occupational categories into 3 groups: high-, intermediate- and low-skill workers

following the classification of Burnod and Chenu (2001).

To merge these four data sets, we extend the procedure of Aubert and Crépon (2003). At the establishment level, we first match SIREN with DADS using the establishment identifier present in both datasets. This establishment-level data set (sector and hours by skill group) is needed below to create a number of local characteristics. Next, we aggregate this establishment dataset at the firm level using the firm identifier. Finally, we merge this firm data with RSI and BRN to recover firm-level information. For each firm between 1994 and 2002, we end up with its value added, the value of its assets, and total hours worked by establishment and skill group. The total number of observations for 1994 is 942,506. This number rises slowly over the period.

Finally, and to avoid dealing with the complications of TFP estimation for multi-establishment firms for which capital and output are known only at the firm level, we restrict our attention to single-establishment firms to estimate TFP.²⁶ Because the information about very small firms tends to be noisy, we only retain firms with more than 5 employees.

Constructing area-year measures of TFP

We now turn to TFP and start by constructing productivity measures for each employment area and year from TFP regressions. We estimate TFP for 114 sectors separately. For simplicity, we ignore sector subscripts for the coefficients. For firm i in a given sector, its value-added va_{it} is specified as:

$$\ln va_{it} = \alpha \ln k_{it} + \beta \ln l_{it} + \sum_m \beta_m^S q_{imt} + \phi_t + \varepsilon_{it}, \quad (18)$$

where k_{it} is the capital of firm i , l_{it} its labour (in hours), q_{imt} the share of labour hours in skill group m , ϕ_t a year fixed effect, and ε_{it} an error term measuring firm TFP. The way we introduce skill shares is justified in Hellerstein, Neumark, and Troske (1999).

Three important issues are worth highlighting at this stage. First, we face the same problem as with wages regarding input quality and more particularly labour. Unfortunately, workers characteristics are typically scarce in firm- or establishment-level data. We

use the strategy used in (18) based on occupational categories to control for labour quality.²⁷ This is obviously a less powerful set of controls than the individual fixed effects used in the wage regressions above.

Second, we can hope to control for the two main factors of production, capital and labour, but not for other factors, land in particular.²⁸ As argued above, the price of land is expected to affect the consumption of land and thus production while, at the same time, be correlated with other local characteristics. Again, instrumenting for these local characteristics is the solution we consider here. Furthermore, output prices are unobserved and are likely to be correlated with local characteristics as well. To the extent that we think of our work as looking into the determinants of local value added rather than pure productivity, this need not bother us much here.²⁹

The third issue about TFP estimation is related to the fact that input choices are expected to be endogenous. This issue has received a lot of attention in the literature (see Akerberg, Caves, and Frazer, 2006, for a recent contribution). For our purpose, this endogeneity bias matters only when it differs across areas. Our main TFP results were estimated using Olley and Pakes (1996). See Appendix A for details about the OP approach. This approach allows us to recover r_{it} , an estimator of ε_{it} . We then average it within sectors, areas, and years:

$$r_{ast} \equiv \frac{1}{L_{ast}} \sum_{i \in (a,s,t)} l_{i,t} r_{i,t}, \quad (19)$$

where $L_{ast} \equiv \sum_{i \in (a,s,t)} l_{i,t}$ is the total number of hours worked in area a , sector s , and year t . A first measure of the local productivity of the average firm in area a and year t denoted TFP_{at}^1 is obtained by averaging equation (19) across sectors within areas and years with weights equal to the number of firms:

$$\text{TFP}_{at}^1 \equiv \frac{1}{n_{at}} \sum_{s \in (a,t)} n_{ast} r_{ast}, \quad (20)$$

where n_{ast} and n_{at} are the total numbers of firms for area a , sector s , and year t and for area a and year t , respectively.

This measure of TFP does not control for the local sector structure. To control for the fact that high productivity sectors may have a propensity to locate in particular areas, we

regress r_{ast} on a full set of sector fixed effect, γ_s :

$$r_{ast} = \gamma_s + \iota_{ast}. \quad (21)$$

This equation is estimated with WLS where the weights are the numbers of establishments associated with each observation.³⁰ To estimate a productivity index TFP_{at}^2 , we average the estimated residuals of (21) for each area and year:

$$\text{TFP}_{at}^2 \equiv \frac{1}{n_{at}} \sum_{s \in (a,t)} n_{ast} \hat{\iota}_{ast}. \quad (22)$$

TFP_{at}^2 can thus be interpreted as a productivity index net of sector effects.

We finally compute a third local productivity index TFP_{at}^3 controlling for variables at the area and sector level, X_{ast} . For that purpose, we estimate the equation:

$$r_{ast} = \text{TFP}_{at}^3 + \gamma_s + X_{ast}\varphi + \epsilon_{ast}. \quad (23)$$

This equation is estimated with WLS where weights are once again the numbers of establishments associated with each observation. It mimics equation (12) for wages and uses the same (centred) local characteristics (same sector specialisation, number of firms, share of professionals, average age, and average squared-age). The main difference however is that these characteristics are constructed using the TFP data and not the wage data.

For comparison, we also estimate equation (18) with OLS. Denote $\hat{\epsilon}_{it}$ the estimated residual for firm i . We then define:

$$r_{ast}^{\text{OLS}} \equiv \frac{1}{L_{ast}} \sum_{i \in (a,s,t)} l_{it} \hat{\epsilon}_{it}, \quad (24)$$

the OLS counterpart to (19). It is possible to recompute our three measures of local productivity TFP_{at}^1 , TFP_{at}^2 , and TFP_{at}^3 using (24) rather than (19). Below we compare the coefficients in our main regressions using local productivity indices computed from OP and OLS.

One aspect of the simultaneity bias at the area level is that establishments may produce more and grow larger in areas where the local productivity is higher. It is possible to

control for that by introducing area and year fixed effects g_{at} in equation (18):

$$\ln va_{it} = \alpha \ln k_{it} + \beta \ln l_{it} + \sum_m \beta_m^S q_{imt} + \phi_t + g_{at} + \varepsilon_{it}. \quad (25)$$

This equation is estimated with OLS. Since this equation is estimated for each sector, the area-year fixed effects depend on the sector and can be rewritten with a sector subindex, g_{ast} . We can then define $r_{ast}^{\text{FE}} \equiv g_{ast}$, the fixed effect counterpart to (19) and (24), and construct once again our three measures of local productivity.³¹

Finally, we average our estimates across years as we did for wages to avoid identifying out of the temporal variation.³² Before going to our results, note that our local productivity variables are strongly correlated with one another. Using OP estimates, the correlation between TFP^1 and TFP^2 is 0.93, the correlation between TFP^1 and TFP^3 is also 0.93, and the correlation between TFP^2 and TFP^3 is 0.98. For TFP^3 , the correlation between OP and OLS estimates is 0.96, the correlation between OP and fixed effects estimates is 0.91, and the correlation between OLS and fixed effects is also 0.91. Finally the correlation between TFP^3 estimated with OP in and mean wages (W^1) is 0.77.³³ This correlation raises to 0.88 after correcting wages of sector effects (W^2) or to 0.87 after correcting wages of sector and worker effects (W^3).

Results

Table 11 presents the results of three regressions for our three measures of local OP productivity: TFP^1 , the mean productivity computed in (20), TFP^2 , the local productivity controlling for sector fixed-effects as estimated in (21), and TFP^3 , the local productivity estimated in (23) which conditions out a broader set of sector effects. This table mirrors the ‘wage’ table 6 for productivity. In columns 1, 2, and 3, these three measures of local productivity are regressed on log employment density controlling for amenities using OLS. The mean elasticity of TFP with respect to density is at 0.04 for mean productivity, 0.041 when taking out sector effects, and 0.047 when also controlling for the local sector structure. In columns 4, 5, and 6, we instrument employment density with 1831 urban population density. The TSLS coefficients for employment density are marginally lower

Table 11. Local TFP (OLLEY-PAKES) as a function of density: OLS and historical instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
	TFP ¹	TFP ²	TFP ³	TFP ¹	TFP ²	TFP ³	TFP ¹	TFP ²	TFP ³
	OLS	OLS	OLS	TOLS	TOLS	TOLS	TOLS	TOLS	TOLS
ln(density)	0.040	0.041	0.047	0.031	0.034	0.038	0.033	0.035	0.039
	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.003) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a
Instruments used:									
ln(1831 density)	-	-	-	Y	Y	Y	Y	Y	Y
ln(1881 density)	-	-	-	-	-	-	Y	Y	Y
First stage statistics	-	-	-	371.4	371.4	371.4	429.1	429.1	429.1
Over-id test <i>p</i> -value	-	-	-	-	-	-	0.07	0.18	0.37
R-squared	0.63	0.70	0.75	-	-	-	-	-	-

306 observations for each regression.

All regressions include a constant and three amenity variables (sea, lake, and mountain).

Standard errors in parentheses. *a*, *b*, *c*: corresponding coefficient significant at 1, 5, 10%.

than in columns 1, 2, and 3. In columns 7, 8, and 9, we add 1881 population density to instrument for contemporaneous employment density. Although the Sargan test of over-identification marginally fails in column 7 with a p -value of 7%, the results are very close to those of columns 4, 5, and 6.

Comparing these results to those of table 6 for wages, we note the following. First, instrumenting for contemporaneous employment density with deep historical lags lowers the coefficients in roughly the same proportion in both cases. This confirms our finding of a mild simultaneity bias regarding the quantity of labour. Second, controlling for sector effects in TFP^3 compared to TFP^1 raises the coefficient on employment density just like it does when considering W^2 instead of W^1 (although the increase is slightly more important here).³⁴ A stronger effect of density after conditioning out sector effects is consistent with the notion that sectors located in less dense areas may gain less from overall density, and perhaps more from same sector specialisation or another sector characteristics that is conditioned out in TFP^3 .³⁵ Third, it is also interesting to note that, when a direct comparison is possible, the density elasticities for wages tend to be above those for TFP. From the theoretical framework developed above (and particularly equations 5 and 6), we actually expect the coefficients on employment density to be higher for wages by a factor equal to the inverse of the labour share ($\frac{1}{1-\alpha}$). With labour coefficients typically between 0.5 and 0.75, the difference between the two sets of estimates is of the right magnitude, although a bit smaller than expected.

To assess the sensitivity of our results to the approach used to estimate TFP, we reproduce in table 14 of Appendix B some of the regressions of table 11 using alternative local productivity indices. These measures of local TFP are constructed from the OLS estimates of (18) and from (25) which computes local productivity fixed effects. When TFP is estimated with OLS instead of OP, the coefficients on density are close, though not exactly the same.³⁶ When TFP is estimated with local fixed effects instead of OP, we find lower coefficients on density. At this stage, our best estimate of the density elasticity of TFP is at 0.04.³⁷

Table 12. Local TFP (OLLEY-PAKES) as a function of density: geological instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	TFP ¹	TFP ²	TFP ³	TFP ³	TFP ³	TFP ³	TFP ³	TFP ³
	TOLS	TOLS	TOLS	LIML	LIML	LIML	LIML	LIML
ln(density)	0.054	0.041	0.045	0.045	0.048	0.047	0.045	0.046
	(0.009) ^a	(0.007) ^a	(0.007) ^a	(0.007) ^a	(0.005) ^a	(0.005) ^a	(0.007) ^a	(0.005) ^a
Instruments used:								
Subsoil mineralogy	Y	Y	Y	Y	Y	N	N	N
Ruggedness	Y	Y	Y	Y	N	N	N	N
Depth to rock	N	N	N	N	N	Y	N	N
Soil carbon content	N	N	N	N	N	Y	Y	N
Topsoil water capacity	N	N	N	N	N	N	Y	Y
Dominant parent material (6 cat.)	N	N	N	N	Y	N	N	Y
First stage statistics	5.5	5.5	5.5	5.5	5.9	7.4	5.2	6.0
Over-id test <i>p</i> -value	0.58	0.60	0.38	0.38	0.26	0.19	0.15	0.22

306 observations for each regression.

All regressions include a constant and three amenity variables (sea, lake, and mountain).

Standard errors in parentheses. *a*, *b*, *c*: corresponding coefficient significant at 1, 5, 10%.

Turning to geological instruments, table 12 mirrors for TFP what table 7 does for wages.³⁸ Columns 1-3 use subsoil mineralogy and ruggedness to instrument for employment density using our three measures of TFP as dependent variables. The coefficients on density are higher than with historical instruments in table 11. Such a difference between geological and historical instruments is also observed with wages.³⁹ To repeat, it reflects the fact that geological instruments have a larger correlation with market potential than local density. The results of columns 3 are confirmed in column 4 when LIML rather than TSLS is used and in column 5-8 when different sets of instruments are used. It is interesting to note that the over-identification tests are passed for the same specifications as with wages (and they also fail for the same unreported regressions as well).

To mirror again the analysis performed with wages, table 15 of Appendix B performs the regressions of table 8 with TFP rather than wages using historical and geological instruments at the same time. The results are again extremely consistent with the wage results. The coefficients on employment density in table 15 with both sets of instruments are the same as those that use historical instruments only in table 11. This near-equality also holds with wages. Furthermore, over-identification tests appear to be passed or failed with the same combinations of instruments again. An exception is column 8 with dominant parent material and topsoil water capacity. The test is passed with wages with a p -value of 15%, while it is failed with TFP (p -value of 5%).

In table 16 of Appendix B we add market potential as explanatory variable as we do with wages in table 9. We again use the exact same specifications as with wages. Adding market potential to the OLS specifications lowers the coefficient on employment density for TFP. The elasticity of TFP with respect to market potential is about half the density elasticity. These two results closely mirror what happens in our wage regressions when we add market potential as an explanatory variable. In the second part of table 16 we instrument employment density with 1831 density and a range of soil characteristics. The coefficient on density declines by about 0.01 point to 0.033 while that on market potential increases by about the same amount to 0.027. This again is very close to what happens

in the wage regressions. Interestingly, the same combinations of instruments pass the over-identification tests with both wages and TFP. The failure of the Sargan test in the last column of table 16 is an exception.

Finally, in table 13, market potential is also assumed to be endogenous. As with wages in table 10, we instrument density and market potential with historical and soil variables. The main result is that instrumenting for market potential leaves its coefficient unchanged. The IV coefficient on employment density is also unchanged. This is the same outcome as with wages. In tables 17 and 18 of Appendix B, we repeat the same exercise but use TFP indices estimated with OLS and with local fixed effects as in equation (25). As in previous comparisons, the results for OLS and OP TFP are very close. With (local) fixed-effect TFP, the density and market potential elasticities are lower than with OLS and OP TFP. However we observe the same stability in the coefficients across regressions. This suggests that the method used to estimate TFP matters with respect to the point estimates for the density and market potential elasticities (though by only 0.01). However, the choice of TFP estimation does not matter otherwise.

7. Conclusions

We revisit the estimation of local scale effects using large scale French wage and TFP data. To deal with the ‘endogenous quantity of labour’ bias (i.e., urban agglomeration is consequence of high local productivity rather than a cause), we take an instrumental variable approach and introduce a new set of geological instruments in addition to standard historical instruments. To deal with the ‘endogenous quality of labour’ bias (i.e., cities attract skilled workers so that the effects of skills and urban agglomeration are confounded), we take a worker fixed effect approach.

Our first series of findings relates to the endogenous quantity of labour bias. (i) Long lags of our endogenous explanatory variables make for strong instruments. (ii) Geological characteristics are more complicated instruments to play with. (iii) Nevertheless, geological and historical instruments lead to similar answers once the regression is properly

Table 13. Local TFP (OLLEY-PAKES) as a function of density and (endogenous) market potential: historical and geological instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
	TFP ¹	TFP ²	TFP ³	TFP ³	TFP ³	TFP ³	TFP ³	TFP ³	TFP ³
	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs
ln(density)	0.028	0.030	0.035	0.034	0.034	0.034	0.034	0.034	0.035
	(0.003) ^a	(0.002) ^a	(0.002) ^a	(0.003) ^a	(0.003) ^a	(0.003) ^a	(0.003) ^a	(0.004) ^a	(0.003) ^a
ln(market pot.)	0.025	0.027	0.026	0.021	0.023	0.021	0.022	0.024	0.018
	(0.005) ^a	(0.004) ^a	(0.004) ^a	(0.009) ^b	(0.007) ^a	(0.006) ^a	(0.008) ^a	(0.013) ^c	(0.009) ^b
Instruments used:									
ln(1831 density)	Y	Y	Y	Y	Y	Y	Y	Y	Y
ln(1881 density)	Y	Y	Y	N	N	N	N	N	N
ln(1831 m. pot.)	Y	Y	Y	N	N	N	N	N	N
Erodibility	N	N	N	Y	N	N	N	N	Y
Soil carbon content	N	N	N	Y	Y	N	N	N	N
Subsoil water capacity	N	N	N	N	Y	Y	N	N	N
Depth to rock	N	N	N	N	N	Y	Y	N	N
Ruggedness	N	N	N	N	N	N	Y	Y	N
Soil differentiation	N	N	N	N	N	N	N	Y	Y
First stage statistics	230.6	230.6	230.6	8.2	16.4	22.8	20.2	8.0	10.5
Over-id test <i>p</i> -value	0.16	0.43	0.17	0.68	0.90	0.60	0.29	0.04	0.16

306 observations for each regression.

All regressions include a constant and three amenity variables (sea, lake, and mountain).

Standard errors in parentheses. *a*, *b*, *c*: corresponding coefficient significant at 1, 5, 10%.

specified: The simultaneity problem between employment density and local wages / productivity is relatively small. It reduces the impact of density by around a fifth.

Our second finding relates to the endogenous quality of labour bias. (iv) Better workers are located in more productive areas. This sorting of workers by skills (observed and unobserved) is quantitatively more important than the endogenous quantity of labour bias. (v) In our regressions, we address sorting using the panel dimension of our wage data. The density elasticity is divided by almost 2. Applying this type approach to TFP is problematic. We thus put more weight on our wage results than we do on our TFP results. Nonetheless, the high degree of consistency between wage and TFP results is reassuring.

We believe the priority for future work should be to develop more sophisticated approaches to deal with the sorting of workers across places. Awaiting progress on this issue, our preferred estimates for the elasticity of wages to density is at 0.02 and around 0.035 for the density elasticity of TFP. For market potential, we find elasticities around 0.035 for wages and 0.025 for TFP. Finally, our result about the relative importance of the two biases raises an interesting question. To which extend does it reflect particular features of the French housing and labour market institutions? One may imagine that in a country like the US with greater labour mobility and a much flatter housing supply curve (in at least part of the country), the endogenous quantity of labour bias might dominate. Further research should inform this question.

Appendix A. Implementation of Olley and Pakes (1996)

The error term in (18) is rewritten as $\varepsilon_{it} \equiv v_{it} + \zeta_{it}$ where v_{it} is the part of the error term that influences the decision of the firm regarding its factors and ζ_{it} is an independent noise. The crucial assumption is that capital investment, I_{it} , can be written as a function of the error term, v_{it} , and current capital: $I_{it} \equiv f_t(k_{it}, v_{it})$ with $\partial f_t / \partial v_{it} > 0$. The investment function can be inverted to yield: $v_{it} = f_t^{-1}(k_{it}, I_{it})$. Equation (18) can then be rewritten as:

$$\ln va_{it} = \alpha \ln k_{it} + \beta \ln l_{it} + \sum_m \beta_m^S q_{imt} + \phi_t + f_t^{-1}(k_{it}, I_{it}) + \zeta_{it}. \quad (\text{A } 1)$$

This equation can be estimated in two stages. Denote $\Phi_t(k_{it}, I_{it}) \equiv f_t^{-1}(k_{it}, I_{it}) + \alpha \ln k_{it} + \phi_t$. Equation (A 1) becomes:

$$\ln va_{it} = \beta \ln l_{it} + \sum_m \beta_m^S q_{imt} + \Phi_t(k_{it}, I_{it}) + \zeta_{it}. \quad (\text{A } 2)$$

This equation can be estimated with OLS after approximating $\Phi_t(k_{it}, I_{it})$ with a third-order polynomial, crossing k_{it} , I_{it} , and year dummies. Its estimation allows us to recover some estimators for the labour and skill share coefficients ($\hat{\beta}$ and $\hat{\beta}_m^S$). It is then possible to construct $z_{it} \equiv \ln va_{it} - \hat{\beta} \ln l_{it} - \sum_m \hat{\beta}_m^S q_{imt}$. Furthermore, the error v_{it} is rewritten as the projection on its lag and an innovation: $v_{it} \equiv h(v_{it-1}) + \zeta_{it-1}$. Using $v_{it-1} = f_{t-1}^{-1}(k_{it-1}, I_{it-1}) = \Phi_{t-1}(k_{it-1}, I_{it-1}) - \alpha \ln k_{j,t-1} - \phi_{t-1}$, the value-added equation then becomes:

$$z_{it} = \alpha \ln k_{it} + \phi_t + h\left(\hat{\Phi}(k_{it-1}, I_{it-1}) - \alpha \ln k_{j,t-1} - \phi_{t-1}\right) + \psi_{it}, \quad (\text{A } 3)$$

where ψ_{it} is a random error. The function $h(\cdot)$ is approximated by a third-order polynomial and equation (A 3) is estimated with non-linear least squares. It allows us to recover some estimators of the capital coefficient $\hat{\alpha}$ and the year dummies $\hat{\phi}_t$. Firm TFP is then defined as $r_{it} \equiv z_{it} - \hat{\alpha} \ln k_{it} - \hat{\phi}_t$. It is an estimator of ε_{it} . For further details about the implementation procedure in stata used in our paper, see Arnold (2005).

Although the OP method allows us to control for simultaneity, it has some drawbacks. In particular, we need to construct investment from the data: $I_{it} = k_{it} - k_{it-1}$. As a consequence it can be computed only for firms that are present in two consecutive

years. Other observations must be dropped. Furthermore, the investment equation $I_{it} = f_t(k_{it}, v_{it})$ can be inverted only if $I_{it} > 0$. Hence, we can keep only observations for which $I_{it} > 0$. This double selection may introduce a bias, for instance, if (i) there is greater ‘churning’ (i.e. entry and exits) in denser areas, and (ii) age and investment affect productivity positively. Then, more establishments with a low productivity may be dropped in high density areas. In turn, this may increase the measured difference in local productivity between areas of low and high density. Re-estimating OLS TFP on the same sample of firms used for OP shows that this is, fortunately, not the case on French data.

Appendix B. Further results

Table 14. Local TFP (OLS and FIXED-EFFECTS) as a function of density: OLS and historical instruments

Variable	TFP estimated with OLS				TFP with fixed effects			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	TFP ¹	TFP ³	TFP ¹	TFP ³	TFP ¹	TFP ³	TFP ¹	TFP ³
	OLS	OLS	TSLS	TSLS	OLS	OLS	TSLS	TSLS
ln(density)	0.035	0.049	0.029	0.042	0.027	0.040	0.018	0.033
	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a
Instruments used:								
ln(1831 density)	-	-	Y	Y	-	-	Y	Y
ln(1881 density)	-	-	Y	Y	-	-	Y	Y
First stage statistics	-	-	432.3	432.3	-	-	432.3	432.3
Over-id test <i>p</i> -value	-	-	0.21	0.10	-	-	0.85	0.75
R-squared	0.63	0.75	-	-	0.45	0.67	-	-

306 observations for each regression.

All regressions include a constant and three amenity variables (sea, lake, and mountain). Standard errors in parentheses. *a*, *b*, *c*: corresponding coefficient significant at 1, 5, 10%.

Table 15. Local TFP (OLLEY-PAKES) as a function of density: historical and geological instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	TFP ¹	TFP ²	TFP ³	TFP ³	TFP ³	TFP ³	TFP ³	TFP ³
	TOLS	TOLS	TOLS	TOLS	GMM	TOLS	TOLS	TOLS
ln(density)	0.031	0.034	0.038	0.038	0.039	0.039	0.039	0.038
	(0.003) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.002) ^a
Instruments used:								
ln(1831 density)	Y	Y	Y	Y	Y	Y	Y	Y
Subsoil mineralogy	Y	Y	Y	Y	N	Y	N	N
Ruggedness	N	N	N	N	Y	Y	Y	N
Hydrogeological class	N	N	N	N	N	N	Y	N
Topsoil water capacity	N	N	N	N	N	N	N	Y
First stage statistics	129.7	129.7	129.7	103.3	194.1	100.3	64.5	125.7
Over-id test <i>p</i> -value	0.14	0.56	0.78	0.66	0.14	0.46	0.33	0.05

306 observations for each regression.

All regressions include a constant and three amenity variables (sea, lake, and mountain).

Standard errors in parentheses. *a*, *b*, *c*: corresponding coefficient significant at 1, 5, 10%.

Table 16. Local TFP (OLLEY-PAKES) as a function of density and (exogenous) market potential: historical and geological instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
	TFP ¹	TFP ²	TFP ³	TFP ³	TFP ³	TFP ³	TFP ³	TFP ³	TFP ³
	OLS	OLS	OLS	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs
ln(density)	0.036	0.037	0.043	0.033	0.033	0.034	0.033	0.033	0.033
	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.003) ^a	(0.003) ^a	(0.003) ^a	(0.003) ^a	(0.003) ^a	(0.003) ^a
ln(market pot.)	0.017	0.019	0.017	0.027	0.027	0.027	0.028	0.027	0.028
	(0.004) ^a	(0.004) ^a	(0.004) ^a	(0.004) ^a	(0.004) ^a	(0.004) ^a	(0.004) ^a	(0.004) ^a	(0.004) ^a
Instruments used:									
ln(1831 density)	-	-	-	Y	Y	Y	Y	Y	Y
Subsoil mineralogy	-	-	-	Y	N	N	N	N	N
Ruggedness	-	-	-	N	Y	N	N	N	N
Subsoil water capacity	-	-	-	N	N	Y	N	N	N
Depth to rock	-	-	-	N	N	N	Y	N	N
Erodibility	-	-	-	N	N	N	N	Y	N
Soil differentiation	-	-	-	N	N	N	N	N	Y
First stage statistics	-	-	-	115.0	170.6	71.4	86.6	70.7	116.1
Over-id test <i>p</i> -value	-	-	-	0.27	0.50	0.69	0.32	0.42	0.05
R-squared	0.64	0.72	0.76	-	-	-	-	-	-

306 observations for each regression.

All regressions include a constant and three amenity variables (sea, lake, and mountain).

Standard errors in parentheses. *a*, *b*, *c*: corresponding coefficient significant at 1, 5, 10%.

Table 17. Local TFP (OLS) as a function of density and (endogenous) market potential: historical and geological instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
	TFP ¹	TFP ²	TFP ³	TFP ³	TFP ³	TFP ³	TFP ³	TFP ³	TFP ³
	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs
ln(density)	0.023	0.022	0.037	0.035	0.035	0.036	0.037	0.036	0.035
	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.003) ^a	(0.003) ^a	(0.003) ^a	(0.003) ^a	(0.004) ^a	(0.003) ^a
ln(market pot.)	0.030	0.030	0.024	0.022	0.026	0.019	0.012	0.020	0.020
	(0.004) ^a	(0.004) ^a	(0.004) ^a	(0.010) ^b	(0.007) ^a	(0.007) ^a	(0.008)	(0.013)	(0.009) ^b
Instruments used:									
ln(1831 density)	Y	Y	Y	Y	Y	Y	Y	Y	Y
ln(1881 density)	Y	Y	Y	N	N	N	N	N	N
ln(1831 m. pot.)	Y	Y	Y	N	N	N	N	N	N
Erodibility	N	N	N	Y	N	N	N	N	Y
Soil carbon content	N	N	N	Y	Y	N	N	N	N
Subsoil water capacity	N	N	N	N	Y	Y	N	N	N
Depth to rock	N	N	N	N	N	Y	Y	N	N
Ruggedness	N	N	N	N	N	N	Y	Y	N
Soil differentiation	N	N	N	N	N	N	N	Y	Y
First stage statistics	232.2	232.2	232.2	8.2	16.4	22.9	20.2	8.0	10.5
Over-id test <i>p</i> -value	0.54	0.56	0.04	0.73	0.29	0.05	0.14	0.38	0.87

306 observations for each regression.

All regressions include a constant and three amenity variables (sea, lake, and mountain).

Standard errors in parentheses. *a*, *b*, *c*: corresponding coefficient significant at 1, 5, 10%.

Table 18. Local TFP (FIXED EFFECTS) as a function of density and (endogenous) market potential:
historical and geological instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
	TFP ¹	TFP ²	TFP ³	TFP ³	TFP ³	TFP ³	TFP ³	TFP ³	TFP ³
	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs	TSLs
ln(density)	0.011	0.013	0.028	0.030	0.029	0.030	0.030	0.030	0.030
	(0.002) ^a	(0.002) ^a	(0.002) ^a	(0.003) ^a	(0.003) ^a	(0.003) ^a	(0.003) ^a	(0.004) ^a	(0.003) ^a
ln(market pot.)	0.032	0.027	0.022	0.013	0.015	0.011	0.008	0.012	0.010
	(0.004) ^a	(0.004) ^a	(0.004) ^a	(0.009)	(0.007) ^b	(0.007)	(0.008)	(0.013)	(0.009)
Instruments used:									
ln(1831 density)	Y	Y	Y	Y	Y	Y	Y	Y	Y
ln(1881 density)	Y	Y	Y	N	N	N	N	N	N
ln(1831 m. pot.)	Y	Y	Y	N	N	N	N	N	N
Erodibility	N	N	N	Y	N	N	N	N	Y
Soil carbon content	N	N	N	Y	Y	N	N	N	N
Subsoil water capacity	N	N	N	N	Y	Y	N	N	N
Depth to rock	N	N	N	N	N	Y	Y	N	N
Ruggedness	N	N	N	N	N	N	Y	Y	N
Soil differentiation	N	N	N	N	N	N	N	Y	Y
First stage statistics	232.1	232.1	232.1	8.2	16.4	22.9	20.2	8.0	10.5
Over-id test <i>p</i> -value	0.63	0.15	0.89	0.68	0.79	0.56	0.38	0.33	0.57

306 observations for each regression.

All regressions include a constant and three amenity variables (sea, lake, and mountain).

Standard errors in parentheses. *a*, *b*, *c*: corresponding coefficient significant at 1, 5, 10%.

References

- Akerberg, Daniel A., Kevin Caves, and Garth Frazer. 2006. Structural identification of production functions. Processed, UCLA.
- Andrews, Donald W.K. and James H. Stock. 2007. Inference with weak instruments. In Richard Blundell, Whitney K. Newey, and Torsten Persson (eds.) *Advances in Econometrics: Proceedings of the Ninth World Congress of the Econometric Society, volume 3*. Cambridge: Cambridge University Press.
- Arnold, Jens Matthias. 2005. Productivity estimation at the plant level: A practical guide. Processed, Bocconi University.
- Aubert, Patrick and Bruno Crépon. 2003. La productivité des salariés âgés : une tentative d'estimation. *Economie et Statistique* 0(368):65–94.
- Bacolod, Marigee, Bernardo S. Blum, and William C. Strange. 2009. Skills in the city. *Journal of Urban Economics* 65:forthcoming.
- Bayer, Patrick, Stephen L. Ross, and Giorgio Topa. 2008. Place of work and place of residence: Informal hiring networks and labor market outcomes. *Journal of Political Economy* 116(6):1150–1196.
- Boutin, Xavier and Simon Quantin. 2006. Une méthodologie d'évaluation comptable du coût du capital des entreprises françaises : 1984–2002. Working paper, INSEE-DESE G2006-09.
- Briant, Anthony, Pierre-Philippe Combes, and Miren Lafourcade. 2007. Does the size and shape of geographical units jeopardize economic geography estimations? Processed, Paris-jourdan Sciences Economiques.
- Burnod, Guillaume and Alain Chenu. 2001. Employés qualifiés et non-qualifiés: Une proposition d'aménagement de la nomenclature des catégories socioprofessionnelles. *Travail et Emploi* 0(86):87–105.
- 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.
- Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2008a. Spatial wage disparities: Sorting matters! *Journal of Urban Economics* 63(2):723–742.
- Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Sébastien Roux, and Diego Puga. 2007. The productivity advantages of large markets: Distinguishing agglomeration from firm selection. Memeographed, University of Toronto.
- Combes, Pierre-Philippe, Thierry Mayer, and Jacques Thisse. 2008b. *Economic Geography*. Princeton (NJ): Princeton University Press.

- Dahl, Gordon B. 2002. Mobility and the return to education: Testing a Roy model with multiple markets. *Econometrica* 70(6):2367–2420.
- Davis, Donald R. and David E. Weinstein. 2008. A search for multiple equilibria in urban industrial structure. *Journal of Regional Science* 48(1):29–65.
- Duranton, Gilles. 2007. Urban evolutions: The fast, the slow, and the still. *American Economic Review* 97(1):197–221.
- 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, 2063–2117.
- Eaton, Jonathan and Zvi Eckstein. 1997. Cities and growth: Theory and evidence from France and Japan. *Regional Science and Urban Economics* 27(4–5):443–474.
- Evans, William N., Wallace E. Oates, and Robert M. Schwab. 1992. Measuring peer group effects: A study of teenage behavior. *Journal of Political Economy* 100(5):966–991.
- Fox, Jeremy T. and Valérie Smeets. 2007. Do input quality and structural productivity estimates drive measured differences in firm productivity? Processed, University of Chicago.
- Glaeser, Edward L. and David C. Maré. 2001. Cities and skills. *Journal of Labor Economics* 19(2):316–342.
- Guo, L.B. and R.M. Gifford. 2002. Soil carbon stocks and land use change: A meta analysis. *Global Change Biology* 8(4):345–360.
- Head, Keith and Thierry Mayer. 2004. The empirics of agglomeration and trade. In Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North-Holland, 2609–2669.
- Head, Keith and Thierry Mayer. 2006. Regional wage and employment responses to market potential in the EU. *Regional Science and Urban Economics* 36(5):573–594.
- Hellerstein, Judith K., David Neumark, and Kenneth R. Troske. 1999. Wages, productivity, and worker characteristics: Evidence from plant-level production functions and wage equations. *Journal of Labour Economics* 17(3):409–446.
- Henderson, J. Vernon. 2003. Marshall's economies. *Journal of Urban Economics* 53(1):1–28.
- Levinsohn, James and Amil Petrin. 2003. Estimating production functions using inputs to control for unobservables. *Review of Economic Studies* 70(2):317–342.
- Marchand, Olivier and Claude Thélot. 1997. *Le travail en France (1800-2000)*. Paris: Nathan.
- 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.

- Moretti, Enrico. 2004. Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics* 121(1):175–212.
- Moulton, Brent R. 1990. An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *Review of Economics and Statistics* 72(2):334–338.
- Olley, G. Steven and Ariel Pakes. 1996. The dynamics of productivity in the telecommunication equipment industry. *Econometrica* 64(6):1263–1297.
- Rappaport, Jordan. 2007. Moving to nice weather. *Regional Science and Urban Economics* 37(3):375–398.
- Redding, Stephen J. and Daniel M. Sturm. 2008. The costs of remoteness: Evidence from German division and reunification. *American Economic Review* 98(5):1766–1797.
- 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, 2119–2171.
- Rosenthal, Stuart S. and William C. Strange. 2008. The attenuation of human capital spillovers. *Journal of Urban Economics* 64(2):373–389.
- Seybold, C.A., J.E. Herrick, and J.J. Brejda. 1999. Soil resilience: A fundamental component of soil quality. *Soil Science* 164(4):224–234.
- Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: Printed for W. Strahan, and T. Cadell.
- Stock, James H. and Motohiro Yogo. 2005. Testing for weak instruments in linear IV regression. In Donald W.K. Andrews and James H. Stock (eds.) *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge: Cambridge University Press, 80–108.
- Syverson, Chad. 2004. Market structure and productivity: A concrete example. *Journal of Political Economy* 112(6):1181–1222.

1. Even if the overall effect is positive, they may also be many negative effects of density on productivity due to crowding or congestion.
2. Combes, Mayer, and Thisse (2008*b*, chapter 11) provide a more complete model of local productivity and a precise discussion of a number of issues including those that relate to the prices of factors, intermediates and final output.
3. In addition, when estimating (6), factors might be endogenous as well. This issue is discussed in section 6.
4. Using the results from (9), it is easy to show that the estimated coefficient for log workforce will be: $\frac{\lambda \text{Var}(\ln B_a) + \delta \text{Var}(\ln C_a) - (\delta + \lambda) \text{Cov}(\ln B_a, \ln C_a)}{\text{Var}(\ln B_a) + \text{Var}(\ln C_a) - \text{Cov}(\ln B_a, \ln C_a)}$. With zero covariance between B_a and C_a and equal variance, this reduces to $\frac{\delta + \lambda}{2}$.
5. Alternative approaches may include focusing on groups of workers for which there is an element of exogeneity in their location decision. One could think for instance of spouses of military personnel. However such groups are likely to be very specific. Another alternative may be to look at ‘natural experiments’ that led to large scale population and employment changes. Such experiments are very interesting to explore a number of issues. For instance, Davis and Weinstein (2008) estimate the effects of the US bombing of Japanese cities during World War II on their specialisation to provide some evidence about multiple equilibria. Redding and Sturm (2008) use the division of Germany after World War II to look at the effects of market potential. However such natural experiments are not of much relevance to study productivity since the source of any such large scale perturbation (e.g., the bombing of Japanese cities) is also likely to affect productivity directly and there is no natural exclusion restriction.
6. The issue with instrumenting is that the number of possible instruments is small while there are potentially dozens of (endogenous) variables that can describe a local economy. In view of this problem, our strategy is to consider parsimonious specifications with no more than one or two potentially endogenous variables. The drawback is that the exclusion restriction for the instruments (i.e., lack of correlation between the instruments

and the error) is more difficult to satisfy than with a greater number of controls. Despite this, we think that a more demanding exclusion restriction is preferable to the addition of inappropriate, and possibly endogenous, controls.

7. Opposite spatial identifying assumptions are made. In Evans *et al.* (1992), the choice of the more aggregate area is assumed to be exogenous while location choice at a lower spatial level is not. Bayer *et al.* (2008) assume instead that randomness prevails at the lower level of aggregation and not at the higher level of aggregation.

8. We view this level of aggregation as a reasonable compromise. On the one hand, we need finely defined sectors in wage regressions and for TFP estimation. On the other hand, localisation economies are expected to be driven by similarities in customers, suppliers, workers, and technology and thus take place at a fairly broad level of sectoral aggregation.

9. We retain a simple specification for market potential and do not aim to derive it from a 'New Economic Geography' model (Head and Mayer, 2004). Alternative specifications for market potential are highly correlated with the one we use. See Head and Mayer (2006) for further evidence and discussion of this fact.

10. We keep in mind that the years are not the same for the wage and TFP regressions. For each set of regressions, the explanatory variables are constructed from the corresponding data sources.

11. Each employment area contains on average more than 100 municipalities.

12. One further (minor) issue need to be mentioned. We take the log of mean wages rather than the mean of log (individual) wages. When viewing local wages as an aggregate of individual wages, the log of mean wages is not the proper aggregate to consider. Mean log wages should be used instead. However, the former is easier to implement than the latter, especially for those who do not have access to micro-data. In any case, this issue is empirically unimportant since the correlation between log mean wages and mean log wages is 0.99.

13. The shares of each skill in local sector employment capture the effects of both individual characteristics at the worker level and the interactions between workers. The two cannot be separately identified with aggregate data.

14. Equation (13) is identified from both the movers (to identify the difference between W_{at}^3 and $W_{a't+1}^3$) and the stayers (to identify the difference between W_{at}^3 and W_{at+1}^3).

15. It is also true that using as dependent variable a coefficient estimated in a previous step introduces some measurement error. The procedure used in Combes *et al.* (2008a) to control for this problem shows that it makes no difference because the coefficients are precisely estimated at the first step.

16. These averages are weighted by by the number of workers in the area for each year to obtain a wage index for the average worker in the area over time. By contrast, our final regressions for the cross-section of employment areas assess whether denser areas make their average firm more productive. There is no longer any reason to weigh the observations (by the number of workers) in these regressions.

17. Because they are in log, using these two variables together allows us to instrument for both past 1831 level and past growth between 1831 and 1881.

18. To aggregate the information from 1 km by 1 km pixels to employment areas, the zonal statistics tool from ArcGIS 9 was used. The tool uses the zones defined in the zone dataset (in our case French employment areas), and internally converts the vectors into a zone raster, which it aligns with the value raster dataset for soils.

19. The ESDB (v2 Raster Archive) contains many more characteristics. For France, some of them like the soil code according to the standard FAO classification are poorly reported. A large number of characteristics also contain categories that refer to land use (e.g., 'urban' or 'agriculture') and are thus not appropriate here. More generally, characteristics *a priori* endogenous to human activity were discarded. Finally, some characteristics such as the secondary dominant parent material stroke us as anecdotal and unlikely to yield relevant

instruments.

20. We use the measures of density used for our wage regressions (1976-1996). Our measures of density for the TFP regressions differ slightly since they are calculated from a slightly different source and cover different years.

21. Stock and Yogo (2005) provide two tests for weak instruments. They are both based on a single F -statistic of the instrumental variables but use different thresholds. The first one tests the hypothesis that two-stage least square (TSLS) small sample bias is small relative to the OLS endogeneity bias ('bias test'). Second, an instrument is considered strong if, from the perspective of the Wald test, its size is 'close' to its level for all possible configurations of the IV regression ('size test'). Note that instruments may be weak in one sense but not another, and instruments may be weak in the context of TSLS but not when using limited information maximum likelihood (LIML).

22. In the other regressions, the differences in the point estimates and standard errors between TSLS and LIML remain small. The differences with respect to the over-identification tests are sometimes more important. This is due to the greater power of the Anderson-Rubin test under LIML relative to the Sargan test used with TSLS.

23. This is consistent with the fact that over-identification tests are passed only for the small set of regressions reported in the table.

24. In this respect, we proceed like Syverson (2004). Nevertheless, valuing assets at their historical costs is not without problems. We minimise them by estimating TFP at the three-digit level with 114 sectors. Indeed, the capital stocks of firms within the same sector are likely to be of the same vintage when sectors are more narrowly defined. We also use year dummies. An alternative would be to deflate assets using economic criteria. However, our panel is rather short which makes it difficult to trace the original investments. Our procedure also differs from that of Olley and Pakes (1996) who use a permanent inventory method.

25. Unfortunately this data cannot be used for our wage regressions because the different years have not been linked up.

26. With multi-establishment firms, we need to impute the same residual estimated from a firm-level production function to all establishments of the same firm. This is a strong assumption that we would rather not make. In results not reported here, we nonetheless experimented with TFP estimated from multi-establishment firms.

27. An obvious way to deal with the unobserved quality of the workforce is to use fixed-effects but unfortunately their use is often problematic with firm-level data because of the sluggish adjustment of capital. See Fox and Smeets (2007) for a more thorough attempt to take (observable) input quality into account when estimating TFP. Like us they find that measures of labour quality are highly significant but taking labour quality into account does not reduce the large dispersion of TFP across firms.

28. We also expect the functional form to matter although we limit ourselves to simple specifications here.

29. In a different context where one is interested in distinguishing between price and productivity effects, such benign neglect may not be warranted. See for instance Combes, Duranton, Gobillon, Roux, and Puga (2007). Note that this issue also applies to wages.

30. These weights give more importance to sectors and areas for which a larger number of $r_{i,t}$ are considered when constructing $r_{s,a,t}$. For these area-sector-years, the sampling error on $r_{s,a,t}$ is usually smaller. Weighing should thus reduce the impact of the sampling error on the dependent variable that comes from the first-stage estimation.

31. We also experimented with a number of alternative TFP approaches such as GMM, cost shares, iv cost shares, Levinsohn and Petrin (2003), etc.

32. Like with wages, these averages are now unweighted.

33. Recall that the years over which TFP and wages are computed are not the same.

34. While TFP^1 may be taken to be the counterpart of W^1 , TFP^3 corresponds to W^2 . Because we cannot control for input quality well, there is no TFP concept that corresponds to W^3 .

35. This higher coefficient on density with TFP^3 is also consistent with possible correlations between unobserved input quality and the local structure of production.

36. When TFP is estimated with OP we must drop the first year of data and firms with no investment. Estimating TFP with OLS on the same sample of firms as with OP makes no difference with respect to OLS estimates of local productivity.

37. Comparing these results to the main study about agglomeration effects using TFP data in the literature, Henderson (2003), is not easy. First, Henderson (2003) uses very different US data for which value added cannot be measured directly and focuses on five industries only. Second, he focuses on sector effects and uses as key independent variable the number of plants in the local industry. We focus instead on total local employment conditioning out local industry shares (among others) in some TFP measures. Third, he estimates TFP and the effects of local characteristics in one stage using a different specification for productivity, which includes firm fixed effects. Finally, he tackles endogeneity problems using a GMM approach. Despite these differences, his findings of strong heterogeneity across industries and modest to high scale effects at the industry level are consistent with ours.

38. That is, aside from the difference in dependent variables, the regressions are exactly the same. The values taken by employment density differ very slightly because of the differences in years between the wage and TFP data and the difference in data source.

39. As previously, the coefficients on density are also slightly above those obtained with wages for similar regressions.