Estimating agglomeration effects:
History and Geology

Pierre-Philippe Combes[*†]
University of Aix-Marseille

Gilles Duranton[*‡]
University of Toronto

Laurent Gobillon[*§]
Institut National d'Etudes Démographiques

Sébastien Roux[*‖]
Centre de Recherche en Économie et Statistique

Work in progress, printed 14 December 2007

ABSTRACT: We revisit the estimation of urbanisation economies 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 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, urbanisation economies, instrumental variables, wages, TFP

JEL classification: R12, R23

[*]We are grateful to Alejandra Castrodad for superb research assistance with the GIS work. Thanks also to Jason Faberman, Stuart Rosenthal, Will Strange, and NARSA conference participants in Savannah (GA). The second author also gratefully acknowledges financial support from the Canadian Social Science and Humanities Research Council.

[†]GREQAM, 2 Rue de la Charité, 13236 Marseille cedex 02, France (e-mail: ppcombes@univmed.fr; website: http://www.vcharite.univ-mrs.fr/PP/combes/). Also affiliated with the Centre for Economic Policy Research.

[‡]Department of Economics, University of Toronto, 150 Saint George Street, Toronto, Ontario M5S 3G7, Canada (e-mail: gilles.duranton@utoronto.ca; website: http://individual.utoronto.ca/gilles/default.html). Also affiliated with the Centre for Economic Policy Research.

[§]Institut National d'Etudes Démographiques, 133 Boulevard Davout, 75980 Paris cedex 20, France (e-mail: laurent.gobillon@ined.fr; website: http://laurent.gobillon.free.fr/).

[‖]Centre de Recherche en Économie et Statistique (CREST-INSEE), 15 Boulevard Gabriel Péri, 92245 Malakoff cedex, France (e-mail: roux@ensae.fr).

1. Introduction

Productivity and wages are higher in larger cities. 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). Typically, a doubling of city size is associated with a 4 to 8% increase in productivity. Is it that a larger city size causes various measures of productivity to increase or, instead, that more productive places tend to grow larger?

To address this question, we take an instrumental variable (IV) approach using both history and geology as sources of exogenous variation for population. Since Ciccone and Hall's (1996) pioneering work, using historical variables such as long lags of population density to instrument for the size or density of local population is a standard way to address endogeneity issues in this context. 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 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 economic activity. More fertile lands may have also attracted people in greater number, etc. To the extent that geology affects the distribution of population and does not otherwise cause productivity, either because the nature of soils affects construction costs and not much else or because fertile lands are no longer a relevant driver of local wealth, geology can provide reasonable instruments to explain the distribution of employment. Except by Rosenthal and Strange (2006) in a slightly different context, geology has not been used to instrument for the distribution of population.

Using both history and geology offers a number of advantages. First, although history and geology may be related to some extent (e.g., geology may have caused history), they arguably affect contemporaneous population in different ways (and this can be tested). Soils are a first-nature determinant of population current patterns and as such are *one* fundamental cause behind current settlement patterns among many. On the other hand, past population patterns, which are correlated with contemporaneous patterns, are the outcome of many factors including historical accidents. Put differently, historical lags are *the* main proximate factors for the current distribution of population. Thus, our approach, which relies on both historical and geological variables, allows us to compare two alternative answers to the same question. Second, and more formally, having two sets of instrumental variables allows us to conduct meaningful diagnostic tests about our instruments. More specifically, we can perform powerful over-identification tests to jointly assess the orthogonality of our instruments with the error term.

Our implementation of this approach on two large scale French data sets for total factor productivity (TFP) and a closely related quantity, wages, yields the following results. Using ordinary least square (OLS), we find that the elasticity of mean wages to employment density in French employment areas is close to 5%. For TFP, we find a similar result. Instrumenting contemporaneous employment density by a measure of population density in 1831 and a number of other historical instruments lowers those numbers by up to one fifth. This is consistent with previous results in the literature and hints at a small endogeneity bias. There is more variability in the answers given by geological variables but they typically confirm the estimates obtained with historical instruments. This suggests that, although the two sets of instruments do not yield exactly the same answer, they are sufficiently close to each other for us to conclude that the endogeneity of population size or density is only a minor issue. We perform a number of robustness tests to confirm these findings. The caveat is, of course, that if the two sets of instruments we use are correlated in the same way with a variable that is both missing from our regressions and affects our local economic outcomes, our results might be unreliable. Our results are thus not the last word on this issue but they represent progress since they rely on much weaker assumptions. Finally, we also explore agglomeration effects at a greater spatial scale using a market potential variable. We find no evidence of simultaneity regarding this variable.

We may refer to the endogeneity problem just described as the 'endogenous quantity of labour' problem. There is also the possibility that denser areas attract better quality labour. We can refer to this problem as the 'endogenous quality of labour' problem. In many respects, this is an extremely hard problem to deal with. When using aggregate wage data, one may be tempted to control for a number of observable characteristics of the local workforce but such characteristics are likely to be endogenous because of sorting. One can then attempt to address this endogeneity issue using again some instrumental variables. They are much scarcer. 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, controlling for observable characteristics of the workforce and finding some source of exogenous variation for their spatial distribution is unlikely to be enough to overcome this 'endogenous quality of labour' problem. This is because we also expect unobservables such as ambition or work discipline to matter and be spatially unevenly distributed (Bacolod, Blum, and Strange, 2007). French university professors may have similar observable characteristics everywhere but a disproportionate fraction of the better ones are working in or around Paris.

To deal with this possible sorting bias, we follow Glaeser and Maré (2001) and Combes, Duranton, and Gobillon (2008) and use the longitudinal dimension of our wage data. We impose individual fixed effects and local time-varying fixed effects in a wage regression. This allows us to separate local and individual effects and reconstruct some local wages net of individual effects. Without any other control, the elasticity of these wages corrected of individual effects to employment density is 3.4% against close to 5% with raw wages. Correcting for the endogenous quantity of labour bias lowers this elasticity to 2.3%. Together with other results, this suggests that the endogenous quality bias is larger than the endogenous quantity bias.

We draw a number of conclusions from this work. First, even though we control for two major sources of bias in local wage and productivity regressions 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 a number of estimation issues. Section 3 presents the data and the details of our instrumentation strategy. Our results for wages are presented in section 4 while those for productivity follow in section 5. Finally, section 6 concludes.

2. Issues when estimating agglomeration effects

A simple model

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

$$y_i = A_i k_i^{\alpha} l_i^{1-\alpha}, \tag{1}$$

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

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

Taking logs directly leads to:

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

The whole focus of the agglomeration literature is then on how the local characteristics of area a where firm i is located determine productivity.[1] We thus assume that TFP depends

[1] See Duranton and Puga (2004) for a review of the theoretical literature.

on a vector of local characteristics X_a and (observed and unobserved) firm characteristics μ_i:

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

Inserting into (3) implies:

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

This equation can in principle be estimated using wage data and local characteristics. An alternative strategy is to estimate (4) directly using:

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

Hence both wage and firm level (TFP) data can be used to estimate the coefficients of interest, φ.

Issues about agglomeration

To estimate equations (5) and (6), we must first choose a level of spatial and sectoral aggregation and the vector of local characteristics X_a that should be considered.

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 boundaries 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 km^2) 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.

With respect 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 industry level (*localisation economies*). Although in this chapter we want to focus on overall scale effects, sector level effects cannot be discarded. Previous results for France suggest that localisation effects matter although they are economically far less important than urbanisation effects (Combes *et al.*, 2008). Since localisation effects are expected to be driven by similarities in customers, suppliers, workers, and technology, using highly disaggregated sectors risks missing important input-output, labour market, or technology linkages. We should thus consider instead relatively broadly defined sectors when controlling for sector effects. On the other hand, sectoral aggregation is also an important issue when estimating TFP. Appropriate TFP estimates require the greatest possible level of disaggregation since TFP is estimated as a residual and thus can be greatly affected by a lack of precision regarding the coefficients on capital and labour (as shown by equation 6).

In light of these constraints, we estimate TFP for 114 three-digit sectors. This is a reasonable compromise between the need for finely defined sectors when estimating TFP and the fact that localisation effects are expected to take place within more broadly defined sectors. We use the same level of sectoral aggregation when working on wage data.

Turning to the explanatory variables that affect local productivity in equation (4), a large number of local characteristics could be considered. Short of large scale experiments in the spatial allocation of population, simultaneity is the fundamental problem when trying to identify the determinants of local TFP or local wages. Virtually any variables that describes the employment and production structure of an area can be suspected of being endogenous. More specifically, the estimation of agglomeration effects is plagued by the problem that high productivity and high wages could be a cause of a high level of local employment as much as a consequence. This is the fundamental identification issue that we are concerned with here.

Although we postpone the discussion of the details of our empirical strategy until later, note that the use of instrumental variables is natural in our context.[2] The issue with instrumenting is that the number of possible instruments may be 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

[2]Alternative approaches may include focusing on groups of workers or firms for which there is an element of exogeneity in the 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 (2007) 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.

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 with parsimonious specifications 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.

The main explanatory variables we are interested in are employment density and market potential. Employment density is our favourite measure of local scale. Since Ciccone and Hall (1996), density-based measures have often been used to assess urbanisation 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. Experimenting with other measures leads to very similar results.

Issues about wages and TFP

We now discuss issues regarding our two main dependent variables, wages and TFP. As made clear by the theoretical framework developed above (and particularly equations (5) and (6)), we expect these two outcomes to be closely related. This result relies on a number of assumptions. Importantly, the production function is taken to be Cobb-Douglas, agglomeration effects are assumed to be Hicks neutral, and the wage setting competitive. Hence, we treat the prediction from equation (5) that the coefficients on local characteristics should be higher for wages by a factor equal to the inverse of the labour share $(\frac{1}{1-\alpha})$ more as an hypothesis than anything else.

There are then issues specific to each dependent variable. Starting with wages, note that to derive equation (2) we also use the first-order condition for labour as well as that for the other factor of production. If this other factor, k, represents physical capital for which the 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, it is also possible to think of this other factor as being land for which the price varies across areas.

This missing variable can have important implications for the estimation. Following Roback (1982), we expect better consumption amenities (which may be entirely unrelated to production) to draw in more population and in turn imply higher land prices. Since land is also a factor of production, firms will use less of it. In turn, this lowers the marginal product of labour when land and labour are imperfect substitutes in the production function (as in the framework above). Put differently, non-production variables may affect both population patterns and be capitalised into wages. To deal with this problem, we can attempt to control for local variables that directly affect consumer utility and thus land prices. However, our range of controls is limited and, to repeat, 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 missing variables that potentially affect both wages and the density of employment, our strategy is to rely again on instrumental variables. Hence, we are asking to our instrument to deal with both the reverse-causality problem described above and the missing variable issue highlighted here.

There is an even more serious issue with wages. 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.*, 2008) leads us to be believe that this is a first-order issue.

Conditioning out the quality of the workforce is in part easy since many individual characteristics are observed. Unfortunately, there are also many unobserved characteristics which are known to play an important role in the determination of wages on the labour market. To the extent that these characteristics do not vary over time, they can be conditioned out with a fixed effect strategy. The hard part is to condition out unobserved individual characteristics separately from unobserved local characteristics with which they may be correlated. The details of our empirical strategy to deal with unobserved quality bias are discussed below.

Turning to TFP, a number of issues must also be kept in mind. First, we can hope to control for the two main factors of production, capital and labour, but not for other factors, land in particular.[3] 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. Second, 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.[4] The main caveat here is that we are not able to disentangle between price and pure productivity effects.[5]

The third issue about TFP estimation is related to the fact that input choices are expected to be endogenous. This potential correlation between the error term used to measure TFP and input choices has received a lot of attention in the literature (see Ackerberg, Caves, and Frazer, 2006, for a recent contribution). Before going any further, it is important to note that this endogeneity bias matters to us only to the extent that it differs across areas. This being said, we experimented with a variety of approaches such as instrumenting with lagged variables and using the more structural methodologies developed by Olley and Pakes (1996) and Levinsohn and Petrin (2003). Our main TFP results were actually estimated using Olley and Pakes (1996). We also report some results for a number of alternative approaches.

The fourth issue with TFP estimation relates to unobserved input quality, more particularly labour. This is the same problem as the endogenous labour quality bias discussed above for wages. Although we postpone our discussion of how to deal with this problem until later, note that workers characteristics are typically much less richly described by firm- or establishment-level data than by standard individual wage data. Note further that 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. This suggests that while we can hope to make good progress on that problem of unobserved quality with labour market data we expect more limited progress with firm level data.[6] This is why we think our wage results carry more weight than our TFP results.

---

[3]We also expect the exact specification to matter although we limit ourselves to simple specifications here.

[4]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).

[5]And the same issue applies to wages.

[6]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.

3. Data and instruments

Main variables

To obtain our two dependent variables, we use four large-scale, French, administrative data sets from the French statistical institute (INSEE).

For wages, we use an extract from the Déclarations Annuelles des Données Sociales (DADS) or Annual Social Data Declarations database. 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.* (2008). 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.

To construct our measures of TFP, we proceed as follows. We first put together two firm-level data sets: the BRN ('Bénéfices Réels Normaux') and the RSI ('Régime Social des Indépendants'). 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 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.[7] 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 a firm 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. For 1994-2002, we use another, exhaustive, DADS dataset. 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 procedure 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 an establishment identifier present in both datasets. This establishment-level data (sector and hours by skill group) is needed to create a number of local characteristics below.

Next, we aggregate this establishment data at the firm level using a 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.

Using the above data, we can readily construct a number of variables for each year. The two explanatory variables we focus on are the density of employment and the market potential for each employment area. Employment density can be readily calculated from the DADS.[8] The market potential of an area is computed as the employment density of all

[7]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.

[8]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.

other areas weighted by the inverse of their distance to the area under consideration.[9] For each area and sector, we 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 four amenities variables. These amenities variables are the share of population located on a sea shore, mountains, lakes, and endowed with 'outstanding cultural or architectural heritage' (as recognised by an inventory of monuments made by the central government).[10] 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 reports a number of descriptive statistics for French employment area. 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.[12] Because the information about very small firms tends to be noisy, we only retain firms with more than 5 employees.

Our first set of instruments is composed of historical populations from early French censuses. For 26 French censuses prior to our earliest observations (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. 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 good half 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

[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]This last explanatory variable is of course unlikely to have a direct effect on local productivity. However, it is likely to determine land prices and thus affect the output of firms indirectly. We note that in the regression below this variable is often significant (though not always). Nonetheless, it is always very small in magnitude (at most a couple of percents), suggesting that the wage capitalisation of amenities is small.

[11]Employment area contains on average more than 100 municipalities.

[12]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 prefer not to make. In results not reported here, we nonetheless experimented with TFP estimated from multi-establishment firms.

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

	Mean	Std. dev.
1976 Mean wage (in 1980 French Francs, per day)	177.7	15.1
1996 Mean wage (in 1980 French Francs, per day)	238.0	18.9
1976 Employment density (workers per sq. km)	64.4	543.0
1976 Log employment density	2.4	1.2
1976 Market potential (workers km per sq. km)	108.1	139.9
1976 Log market potential	4.4	0.7
1831 Urban population density (per sq. km)	38.2	419.8
1881 Urban population density (per sq. km)	106.8	1232.3
Sea (% municipalities on a coast line)	8.8	21.1
Lake (% municipalities on a lake coast)	17.2	12.9
Mountain (% municipalities on a mountain)	9.8	19.7
Heritage (% municipalities)	48.6	16.5

Source: DADS for the first six lines, historical censuses for the next two and 1988 municipal inventory for the last four.

employment area.[13] Given that soil variables are usually discrete, we use the majority, or the value that appears more often in each area. To take an illustrative example, the initial and transformed data for the water capacity of the subsoil are represented in figure 1. 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 non-missing soil variables). 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.[14] 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

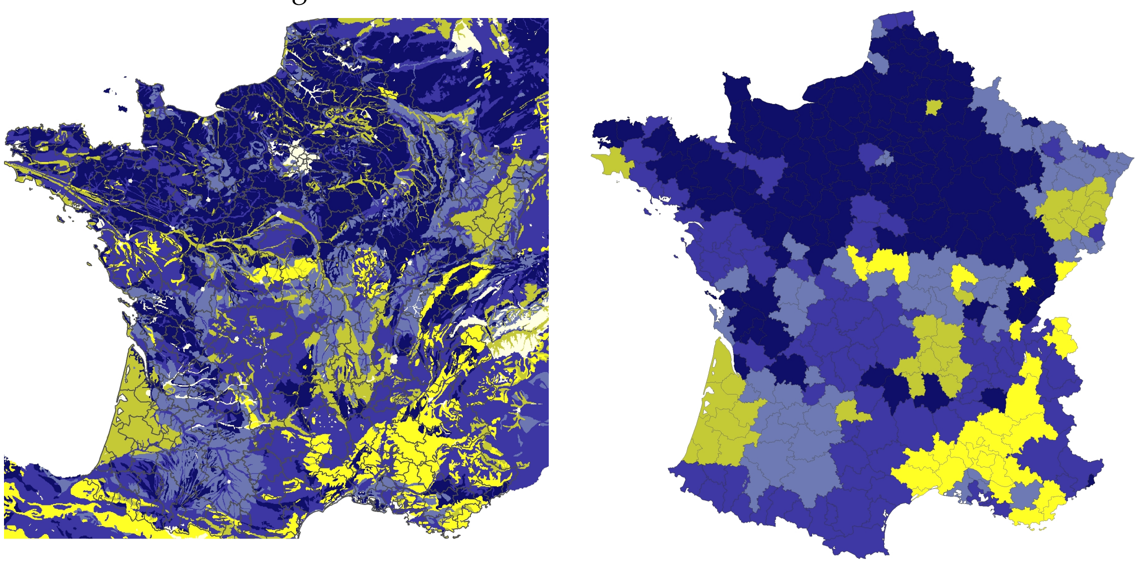
[13]To aggregate the information from 1 km by 1 km pixels to employment areas, the zonal statistics tool from ArcGIS 9 was used. The tool uses the zones defined in the zone dataset (in our case French employment areas), and internally converts the vectors into a zone raster, which it aligns with the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

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

Figure 1. Geological variables: Water capacity of the subsoil

Panel A. Original data

Panel B. Transformed data



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 are in white.

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 variables 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

The regressions we want to estimate are:

$$\ln W_{at} = \text{Constant} + X_{at}\varphi^{W} + \mu_{at}^{W} \tag{7}$$

and

$$\ln TFP_{at} = \text{Constant} + X_{at}\varphi^{TFP} + \mu_{at}^{TFP}, \tag{8}$$

where the vector of dependent variables X_{at} contains the four amenity variables discussed above, employment density and 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 using instrumental variables can yield unbiased estimates provided that the instruments satisfy two conditions, relevance and exogeneity. Formally, these conditions are

$$Cov(Density, Z|.) \neq 0, \qquad Cov(MarketPotential, Z|.) \neq 0, \tag{9}$$

and

$$Cov(\mu, Z) = 0. \tag{10}$$

where Z denotes the set of instruments. Similar conditions apply when local TFP, rather than wages, is the dependent variable. We begin by discussing the ability of our instruments to predict contemporaneous employment density and market potential.

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

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

	ln(1996 employment density)	ln(1996 market potential)
ln(1831 density)	0.56 (0.74)	0.06 (0.24)
ln(1881 density)	0.76 (0.87)	0.11 (0.33)
ln(1831 market potential)	0.20 (0.47)	0.96 (0.98)
ln(1881 market potential)	0.21 (0.48)	0.98 (0.99)

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

France (Eaton and Eckstein, 1997). The raw data confirms this. Table 2 presents pairwise correlations between our four historical instruments and current employment density and market potential. For the sake of comparison with geology variables below, we also report the adjusted 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 variables, we expect the nature of soils and their characteristics to be fundamental drivers of population settlements. Soil characteristics arguably determine their fertility. Since soil characteristics are discrete variables, it is not meaningful to run pairwise correlations as with historical variables. Instead, table 3 reports the R-squared when regressing 1996 employment density and 1996 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 appear promising. 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 (9) makes clear that the validity 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, mountain, and heritage) and year effects. 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

Table 3. R-squareds when regressing 1996 density and market potential on soil characteristics

	ln(1996 emp. density)	ln(1996 market pot.)
Subsoil mineralogy (2 dummies)	0.02	0.06
Topsoil mineralogy (3 dummies)	0.02	0.06
Dominant parent material (5 dummies)	0.10	0.31
Dominant parent material (19 dummies)	0.13	0.48
Topsoil water capacity (2 dummies)	0.03	0.22
Subsoil water capacity (3 dummies)	0.01	0.32
Depth to rock (3 dummies)	0.09	0.35
Soil differentiation (2 dummies)	0.06	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

Adjusted R-squareds. 306 observations.

Table 4. First stage: Density

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
ln(1831 density)	0.901 (0.053)[a]								
ln(1881 density)		0.908 (0.037)[a]							
Subsoil mineralogy	N	N	Y	N	N	N	N	N	N
Dominant parent material (20 categories)	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
Amenities	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
R-squared	0.58	0.76	0.07	0.24	0.20	0.06	0.10	0.14	0.11
F-test (H_0 − All instruments zero)	287.0	617.8	4.2	3.5	4.8	1.4	6.0	8.6	6.1
Partial R-squared	0.57	0.75	0.03	0.21	0.16	0.02	0.07	0.10	0.07

Dependent variable: ln(employment density).
All regressions include a constant and the four amenity variables are sea, lake, mountain, and outstanding architectural heritage. Standard errors clustered by employment area in parentheses.
1836 observations for each regression (306 employment areas for 1976, 1980, 1984, 1988, 1992, and 1996).
a, b, c: 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.)	0.999								
	(0.014)[a]								
ln(1881 market pot.)		0.946							
		(0.012)[a]							
Subsoil mineralogy	N	N	Y	N	N	N	N	N	N
Dominant parent material (20 categories)	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
Amenities	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
R-squared	0.96	0.99	0.24	0.57	0.50	0.41	0.29	0.44	0.31
F-test (H_0 − All instruments zero)	5156.7	6142.7	6.1	34.8	41.8	19.5	9.7	23.3	14.6
Partial R-squared	0.95	0.98	0.04	0.46	0.37	0.25	0.10	0.29	0.13

Dependent variable: ln(market potential).
All regressions include a constant and the four amenity variables are sea, lake, mountain, and outstanding architectural heritage. Standard errors clustered by employment area in parentheses.
1836 observations for each regression (306 employment areas for 1976, 1980, 1984, 1988, 1992, and 1996).
a, *b*, *c*: significant at 1%, 5%, 10%.

on a series of soil dummies concerning their mineralogy, dominant parent material, water capacity, carbon 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 variables. 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).[15] Table 4 the relevant F-statistics. The two lagged density instruments in columns 1 and 2 have F-statistics close to 300 and 600, respectively. This makes them very strong in light of the critical values reported by Stock and Yogo (2005) in their tables 1-4. The soils instruments are weaker by comparison and fall below the critical values of Stock and Yogo (2005) with TSLS. To avoid the pitfalls of weak instruments, a number of possible strategies can be envisioned. First, it would be possible to increase the strength

[15] 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).

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 variables. 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, 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 time variation and naturally yields higher R-squareds. Also, distances used to compute the market potential do not vary over time. The facts that in column 1 the coefficient on 1831 market potential is essentially one and the partial R-squared is 95% indicate that we should not expect much difference between OLS and TSLS below. This result is also a first strong argument in favour of the exogeneity of market potential. Despite the construction of railroads, highways or airports, European integration, and radical technological change over 150 years, 1831 market potential still very accurately predicts contemporaneous market potential. Soils variables are also stronger instruments for market potential than they were for density. To repeat, the reason is that the spatial scale at which most soil variables vary is greater than that of French employment areas. This is illustrated by panel B of figure 1 for subsoil water capacity. In a nutshell, soil variables are better replicating the smooth evolution of market potential than the spikes of employment density.

One may worry that the good explanatory power of soil variables 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 (10) gives the second condition that must be satisfied by a valid instrument: the instrument must be orthogonal 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. We now develop *a priori* arguments that our instruments satisfy this 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 1831 was very different from what it is today. First, the structure of the French economy 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 dependance 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. 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 South-Western Mediterranean coast have been 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 in such circumstances.

We now consider geological variables. The *a priori* case for thinking that geological variables are good instruments hinges first on the fact that soil characteristics have been decided mostly by nature and do not result from human activity. This argument applies very strongly to a number of soil variables we use. For instance, soil mineralogy and their dominant parent material were determined millions of years of 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 variables 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 precisely.

Ruling out reverse-causality does not, however, ensure that any soil characteristics will automatically satisfy condition (10) and be a valid instrument. Any correlation between a soil characteristic and a missing variable in (7) or (8) would make it invalid as an instrument. The main argument 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 variables 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. Soil variables are narrower but more 'fundamental' determinants of population location. Put differently, although we expect soils to have determined history, they were not the sole determinants of population patterns in 1831. As will become clear below, geological variables can also explain current patterns of employment density over and above past employment density. Consequently we can meaningfully compare the answers given by these two groups of instruments since, if one group of instruments fail, 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 selection 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.

4. Main wage results

Three wages

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

$$W^1_{at} \equiv \ln \overline{w}_{at} \equiv \ln \left(\frac{1}{N_{at}} \sum_{j \in (a,t)} w_{jt} \right). \tag{11}$$

We can then use W^1_{at} as dependent variable to be explained by local employment density and other local characteristics in equation (7). Using a simple log mean like W^1_{at} throws a number of problems. First, when using mean wages we do not condition out localisation economies nor, more generally, sectoral effects. Second, it is also important to keep in mind that W^1_{at} does nothing regarding the endogenous quality of labour bias.[16]

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:

$$\overline{w}_{ast} \equiv \frac{1}{N_{ast}} \sum_{j \in (a,s,t)} w_{jt}. \tag{12}$$

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:

$$\ln \overline{w}_{ast} = W^2_{at} + \gamma_s + X_{ast}\varphi + \epsilon_{ast}. \tag{13}$$

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 localisation effects we use in X_{ast} the 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 is each of 6 skill groups.[17] In equation (13), the coefficient of interest is W^2_{at}, a fixed effect for each employment area and year. W^2_{at} can be interpreted as a wage index for area a and year t after conditioning out the effects of sectors, localisation economies, and observable characteristics.[18] When estimating (13), all local sector and mean individual characteristics are centred and the observations are weighted by the square-root of the number of workers in each cell to avoid heteroscedasticity.

The coefficients W^2_{at} 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 (13) are given in Combes *et al.* (2008), 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 that will be made clear below.

[16]Two further (minor) issues need to be mentioned. First, 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 in absence of micro-data. In any case, this issue is empirically unimportant since the correlation between log mean wages and mean log wages is 0.99. Second, and since we generate an observation for each employment area and year, an 'individual perspective' on local wages demands that each observation be weighted by the number of workers and properly clustered. In the results reported below, we cluster our standard errors by employment areas but do not weigh our observations by local employment. Weighing (or not) has only a small effect on the results due to a mild non-linearity of the effects of employment density. We prefer to report un-weighted results because they are more immediately comparable to our TFP results in section 5.

[17]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.

[18]W^2 needs to be properly re-normalised to be interpreted directly as a wage.

Next, estimating (13) 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 whether we condition out sector effects or not does not affect our final results. In turn, this is because although the coefficients for local specialisation and the number of establishments are significant, they only explain a very small part of the variation in (13) (Combes *et al.*, 2008).

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. To deal with this problem of endogenous labour quality, a number of approaches can be envisioned. The first would be to instrument labour quality at the area level just like we instrument for labour quantity. 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 (2005) use location at the block level and assume an absence of sorting conditional on neighbourhood choice. The problem is that 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.* (2005) assume instead that randomness prevails at the lower level of aggregation and not at the higher level of aggregation. 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 to the complications introduced by the discrete nature of the choice among many locations, this is a difficult exercise. Dahl (2002) proposes a new approach in this direction 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 Moretti (2004) and Combes *et al.* (2008). Rather than (13), we can instead estimate:

$$\ln w_{it} = W^3_{a(it)t} + \gamma_{s(it)} + X^1_{a(it)s(it)t}\varphi^1_{s(it)} + X^2_{it}\varphi^2 + \theta_i + \epsilon_{it}. \tag{14}$$

This equation is estimated at the level of individual workers and contains a worker fixed effect θ_i which controls for all fixed individual characteristics.[19] The use of individual data also allows us to control for individual characteristics X^2_{it} (age and its square) separately

[19] Equation (14) is identified from both the movers (to identify the difference between $W^3_{\mathbf{a}t}$ and $W^3_{\mathbf{a'}t+1}$) and the stayers (to identify the difference between $W^3_{\mathbf{a}t}$ and $W^3_{\mathbf{a}t+1}$).

from (centred) local industry characteristics X^1_{ast}. As previously, 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 (14) is the wage index W^3_{at} for each area and year after conditioning out sector effects, localisation economies, 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 (14) regards the endogeneity of location choices. However, because we have time-varying local effects, W^3_{at}, problems only arise when we have spatial or industry 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 fixed effects, i.e., when they make their location decision on the basis of their expected wages. See Combes *et al.* (2008) for further discussion.

Finally, note that we prefer this two-step approach, which first estimates (14) before regressing W^3_{at} on local characteristics, to its corresponding one-step counterpart. It is true that the error structure with two steps is marginally more restrictive. However, Combes *et al.* (2008) show that it has no significant bearing on the results. 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.* (2008) to control for this problem shows that it makes no difference because the coefficients are precisely estimated at the first step. On the other hand, our two-step approach offers three significant benefits. 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.

Before turning to our results for employment density and market potential, it is interesting to note that these three local wage variables are highly correlated with one another. For 1996, the correlation between W^1 and W^2 is 0.87. That between between W^1 and W^3 is also equal to 0.87. Finally, the correlation between W^2 and W^3 is 0.92.

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

Variable	[1] W^1 OLS	[2] W^2 OLS	[3] W^3 OLS	[4] W^1 OLS	[5] W^2 OLS	[6] W^3 OLS	[7] W^1 TSLS	[8] W^2 TSLS	[9] W^3 TSLS
ln(density)	0.047	0.051	0.034	0.041	0.048	0.027	0.035	0.042	0.023
	(0.0029)^a	(0.0022)^a	(0.015)^a	(0.0031)^a	(0.0025)^a	(0.0016)^a	(0.0035)^a	(0.0022)^a	(0.0016)^a
ln(market pot.)	-	-	-	0.024	0.011	0.026	0.032	0.018	0.030
	-	-	-	(0.0059)^a	(0.0044)^b	(0.0030)^a	(0.0064)^a	(0.0047)^a	(0.0031)^a
Instruments used:									
ln(1831 density)	-	-	-	-	-	-	Y	Y	Y
ln(1881 density)	-	-	-	-	-	-	Y	Y	Y
ln(1831 m. pot.)	-	-	-	-	-	-	Y	Y	Y
First stage statistics	-	-	-	-	-	-	1761	1761	1761
Over-id test p-value	-	-	-	-	-	-	0.91	0.11	0.17
R-squared	0.82	0.89	0.94	0.83	0.89	0.95	-	-	-

All regressions include a constant, 4 amenity controls, and year effects.
Standard errors clustered by employment area in parentheses.
1836 observations for each regression.
a, b, c: significant at 1%, 5%, 10%. Cragg-Donald statistic reported for the first stage (this statistic is not robust but the first-stage F statistics reported above are).

Results

Table 6 presents the results of three simple regressions for our three wages: W^1, the mean local wage as computed in (11), W^2, the wage index after conditioning out sector effects and observable individual characteristics as estimated in (13), and W^3, the wage index from (14) which also conditions out individual fixed effects. In columns 1, 2, and 3, these three wages are regressed on log employment density controlling for year effects and four amenity variables using OLS. The density elasticity of mean wages is at 4.7%. This is very close to previous results in the literature (Ciccone and Hall, 1996; Ciccone, 2002). Controlling for sector effects in column 2 yields a marginally higher estimate of 5.1% for the density elasticity. Controlling also for unobserved individual characteristics gives a significantly lower elasticity of 3.4%. This suggests that a good share of measured agglomeration effects are in fact attributable to the unobserved characteristics of the workforce.

Adding market potential among explanatory variables in columns 4, 5, and 6 slightly lowers the density elasticity of wages. This is unsurprising since the correlation between employment density and market potential is rather high at 0.5. With our preferred measure of wages, W^3, the elasticity of wages to market potential is at 2.6%. It is about the same as the density elasticity.

In columns 7, 8, and 9, we perform the same regressions as in columns 4, 5, 6 but we instrument density and market potential with 1831 and 1881 urban population densities and 1831 market potential. Compared to their corresponding OLS coefficients, the TSLS coefficients for employment density are about 0.5% point lower. On the other hand,

Table 7. Local wages as a function of density and market potential: 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	LIML	LIML	LIML	LIML	LIML
ln(density)	0.017	0.026	0.032	0.033	0.011	0.006	0.008	0.004
	(0.025)	(0.018)	$(0.012)^a$	$(0.015)^a$	(0.014)	(0.014)	(0.011)	(0.054)
ln(market pot.)	0.051	0.032	0.022	0.021	0.060	0.066	0.049	0.068
	(0.032)	(0.024)	(0.015)	(0.019)	$(0.023)^a$	$(0.025)^a$	$(0.015)^a$	(0.104)
Instruments used:								
Subsoil mineralogy	Y	Y	Y	Y	Y	N	N	Y
Depth to rock	Y	Y	Y	Y	N	N	N	N
Soil carbon content	N	N	N	N	Y	Y	N	N
Ruggedness	N	N	N	N	N	Y	Y	Y
Erodibility	N	N	N	N	N	N	Y	N
First stage statistics	5.3	5.3	5.3	5.3	5.4	7.0	10.8	1.6
Over-id test p-value	0.19	0.91	0.69	0.19	0.63	0.41	0.28	0.17

All regressions include a constant, 4 amenity controls, and year effects.
Standard errors clustered by employment area in parentheses.
1836 observations for each regression.
a, b, c: significant at 1%, 5%, 10%. Cragg-Donald statistic reported for the first stage.

instrumenting raises the coefficient on market potential by about the same amount. The fact that the partial R-squared for the market potential in the first stage is 96% suggests that the main issue is an upward bias on the density coefficient in OLS. The increase in the coefficient on market potential simply follows from the correlation between these two explanatory variables.

To summarise, the density elasticity of mean wages is 4.7% (column 1). Improving the specification to control for market potential lowers it to 4.1% (column 4). Instrumenting for density and market potential lowers this coefficient further to 3.5% (column 7). Controlling for sector effects and the unobserved quality of labour lowers it even further to 2.3%, that is to say less than half the initial coefficient.[20] Overall the results in this table suggest a small change in the density coefficient when introducing the market potential, a small endogenous quantity of labour bias for employment density, and a larger endogenous quality of labour bias.

Table 6 also reports the p-values for an over-identification test (Hansen's J statistic) for the three IV regressions. For our preferred specification in column 9, the p-value of this statistic is 0.17. We fail to reject our over-identifying restriction. Column 8 reports a slightly lower value while column 7 reports a higher value.

Next, table 7 reports results for number of regressions which all use geological variables as instruments. As made clear above, most geological variables appear to explain market

[20]The wage index W^3 is estimated only from the workers who appear at least twice in the data whereas W^1 and W^2 are computed from all workers. Computing these two quantities from the same sample as that used to estimate W^3 only generates minimal differences.

potential better than employment density. As a result, IV estimations that rely *only* on geological variables to instrument for both market potential and employment density are not expected to perform very well. In this table, we focus on only five sets of geological variables for which the over-identification tests are passed. In column 1, mean local wages, W^1, are regressed on the same variables as in table 6 using subsoil mineralogy (2 dummies) and depth to rock (3 dummies) to instrument for employment density and market potential. The coefficients on employment density and market potential are slightly different from those obtained in table 6 column 7 when instrumenting with historical variables. The standard errors are also bigger, unsurprisingly. As result the coefficients on market potential and employment density are not significant at conventional thresholds. Nor are they statistically different from those obtained with historical instruments. Columns 2 and 3 repeat the same regression for the wage net of sector effects, W^2, and the wage net of sector effects and individual fixed effects, W^3. Because W^3 controls for the unobserved quality of labour, we focus on this dependent variable in what follows.

The absence of significant difference between the coefficients obtained with historical instruments and those with the geological instruments used in columns 1-3 is also frequently obtained when experimenting with alternative sets of geological instruments. Before taking these results at face value, note that the low first stage statistics in columns 1-3 raises some questions about the strength of these 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. Their comparison shows that the TSLS and LIML coefficients for employment density and market potential are very close. However, at 0.19 the *p*-value for the over-identification test is much lower with LIML in column 4 than with TSLS in column 3. This is due to the much greater power of the (robust) Anderson-Rubin test used with LIML compared to the Hansen's J statistics used with robust TSLS.[21]

In columns 5 to 8, we report LIML results for further combinations of instruments. The coefficient on employment density is positive but insignificant. However it does not differ significantly from our preferred estimate of 2.3% with historical instruments. The coefficient on market potential is higher than our preferred estimate of 3.0% with historical instruments. We suspect that the better predictive abilities of the instruments for market

[21]More generally, with TSLS the over-identification tests are passed most of the times for most instruments. This does not hold under LIML. However the coefficients obtained with LIML and TSLS are nearly always in the same ballpark and we can never reject that historical and geological instruments give the same result. At some level, failing to pass over-identification tests need not worry us to much. With a powerful enough test and sufficient data, it would be unlikely have instruments that are *exactly* orthogonal to the error. More generally the combination of collinear instruments and weak over-identification tests (such as Hansen'J) can easily lead to false claims about instrument validity. On the other hand, the interpretation of *p*-values on the low side obtained very different instruments and powerful over-identification tests may be not be so clear-cut.

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

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	W^3	W^3	W^3	W^3	W^3	W^3	W^3	W^3
	TSLS	LIML	TSLS	TSLS	TSLS	TSLS	TSLS	TSLS
ln(density)	0.019	0.019	0.021	0.022	0.023	0.021	0.020	0.021
	$(0.0042)^a$	$(0.0042)^a$	$(0.0019)^a$	$(0.0025)^a$	$(0.0018)^a$	$(0.0020)^a$	$(0.0022)^a$	$(0.0022)^a$
ln(market pot.)	0.044	0.045	0.035	0.028	0.022	0.034	0.039	0.033
	$(0.018)^b$	$(0.018)^b$	$(0.0054)^a$	$(0.0089)^a$	$(0.0062)^a$	$(0.0053)^a$	$(0.0077)^a$	$(0.0063)^a$
Instruments used:								
ln(1831 density)	Y	Y	Y	Y	Y	Y	Y	Y
Subsoil mineralogy	Y	Y	N	N	N	N	N	N
Depth to rock	N	N	Y	N	N	N	N	N
Soil differentiation	N	N	N	Y	N	N	N	N
Topsoil water capacity	N	N	N	N	Y	N	N	N
Subsoil water capacity	N	N	N	N	N	Y	N	N
Erodibility	N	N	N	N	N	N	Y	N
Dominant parent material (6 cat.)	N	N	N	N	N	N	N	Y
First stage statistics	16.0	16.0	148.2	62.2	129.3	141.1	58.1	15.0
Over-id test p-value	0.85	0.68	0.29	0.15	0.44	0.37	0.46	0.10

All regressions include a constant, 4 amenity controls, and year effects.
Standard errors clustered by employment area in parentheses.
1836 observations for each regression.
a, b, c: significant at 1%, 5%, 10%. Cragg-Donald statistic reported for the first stage.

potential imply that the prediction for this variable captures the effect of both market potential and employment density. Taken together these results suggest two conclusions. First, as could be suspected from tables 4 and 5 above, it is hard to instrument for both market potential and employment density using only geological variables. Second, none of the results generated here contradicts the conclusions reached above using historical instruments.

We now consider in table 8 historical and geological instruments simultaneously. A number of issues must be kept in mind. First, 1831 market potential is such a strong instrument that geological variables have no explanatory power left when past market potential is also used as an instrument. As a result, we only report results that exclude past market potential from the set of instruments.[22]

[22]It is nonetheless interesting to note that when past market potential and past population density are used together with a set of geological variables, the over-identification test is usually passed. Second, we use 1831 population density as sole historical instrument in every regression. Adding 1881 population density changes nothing because of the strong correlation between these two instruments. Third, we only reports results for which the first stage F statistics for market potential is above 10. Depending on the exact number of dummy variables for the set of geological variables that is considered, this ensures that the first stage F statistics for market potential is around or above the critical values for weak instruments given by Stock and Yogo (2005). It should be noted that in all the reported regressions, at least one soil dummy is significant at 1% in the first stage for market potential. We check this to avoid situations where identification would only come from 1831 population density. As for the first stage F for employment density, it is always well above 10. Finally, we only report results for local wages corrected of individual effects, W^3, for reasons mentioned above.

In table 8 column 1, we use 1831 density and dummies for the mineralogy of the subsoil to instrument for density and market potential.[23] The density elasticity of wages is at 1.9% while the market potential density is at 4.4%. Compared to our preferred estimation using only historical variables (column 9 of table 6), the coefficient on employment density is slightly lower while that on market potential is slightly higher. However the differences are not significant. Column 2 re-estimates the same regression using LIML. The coefficients are virtually undistinguishable. So are the standard errors. This is reassuring in light of any weak instrument concern. Although we do not report further LIML results, this absence of difference between TSLS and LIML holds for all the other regressions in the table. In column 3, we replace subsoil mineralogy by the depth to rock. This yields very similar results. Columns 4, 5, 6, and 7 repeat the same exercise using soil differentiation, topsoil and subsoil water capacity, and erodibility to obtain very similar results again. In column 8, the results are yet again the same using the dominant parent material (measured in 6 categories). However for this latter set of instruments, the over-identification test is only marginally passed. Using the more detailed measure of dominant parent material leads to a p-value marginally below 10% for this test.

In Appendix A, table 14 reproduces the same set of regressions as table 8 except that it takes market potential as exogenous on the ground that it is very highly correlated with its past values from 150 years ago and that there is no evidence above of any simultaneity bias. In the table, the coefficient on employment density is rock solid at 2.1% while that on market potential is equally stable at 3.2%. Overall, these results of tables 8 and 14 strongly confirm those of the two previous tables.

5. Main TFP results

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:

$$\ln va_{it} = \alpha \ln k_{it} + \beta \ln l_{it} + \sum_j \beta_j^S q_{ijt} + \theta_t + \varepsilon_{it}, \qquad (15)$$

[23] One might worry that geological variables may no longer have any predictive power after including 1831 density as excluded instrument. This is not the case. Because lagged density is imprecisely measured (or because it only measures urban population), geological variables are actually more strongly correlated with and explain a greater share of the variation of contemporaneous density than of 1831 density. Importantly, geological variables remain correlated with current density conditional on 1831 density.

where k_{it} is the capital of firm i, l_{it} its labour (in hours), q_{ijt} the share of labour hours in skill group j, θ_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).

As discussed above, capital and labour are likely to be endogenous in (15). A standard approach to deal with this problem is proposed by Olley and Pakes (1996). See Appendix B 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:

$$r_{ast} \equiv \frac{1}{L_{ast}} \sum_{i \in (a,s,t)} l_{i,t} r_{i,t}, \tag{16}$$

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 local productivity denoted TFP^1_{at} is obtained by averaging equation (16) across sectors within areas and years:

$$\text{TFP}^1_{at} \equiv \frac{1}{L_{at}} \sum_{s \in (a,t)} L_{ast} r_{ast}, \tag{17}$$

where L_{at} is total hours for area a and year t.

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:

$$r_{ast} = \gamma_s + \iota_{ast}. \tag{18}$$

This equation is estimated with WLS where the weights are the number of establishments associated with each observation.[24] To estimate a productivity index TFP^2_{at}, we average the residuals of (18) for each area and year:

$$\text{TFP}^2_{at} \equiv \frac{1}{L_{at}} \sum_{s \in (a,t)} L_{ast} \iota_{ast}. \tag{19}$$

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

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

$$r_{ast} = \text{TFP}^3_{at} + \gamma_s + X_{ast}\varphi + \epsilon_{ast}. \tag{20}$$

This equation is estimated with WLS where weights are once again the number of establishments associated with each observation. It mimics equation (13) for wages and uses

[24]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.

the same (centred) local characteristics. A difference however is that these characteristics are constructed using the TFP data and not the wage data.

For comparison, we also estimate equation (15) with OLS. Denote $\widehat{\varepsilon}_{it}$ the residual for firm i. We then define:

$$r_{ast}^{\text{OLS}} \equiv \frac{1}{L_{ast}} \sum_{i \in (a,s,t)} l_{it} \widehat{\varepsilon}_{it}, \tag{21}$$

the OLS counterpart to (16). It is possible to recompute our three measures of local productivity TFP_{at}^{1}, TFP_{at}^{2}, and TFP_{at}^{3} using (21) rather than (16). 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 (15):

$$\ln va_{it} = \alpha \ln k_{it} + \beta \ln l_{it} + \sum_{j} \beta_{j}^{S} q_{ijt} + \theta_t + g_{at} + \varepsilon_{it}. \tag{22}$$

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 g_{ast}. We can then define $r_{ast}^{\text{FE}} \equiv g_{ast}$, the fixed effect counterpart to (16) and (21), and construct once again our three measures of local productivity.

Before going to our results, note that our local productivity variables are fairly strongly correlated with one another. For 1996 and using OP estimates, the correlation between TFP^1 and TFP^2 is 0.73, the same as the correlation between TFP^1 and TFP^3 while that between TFP^2 and TFP^3 is 0.99. For TFP^3, the correlation between OP and OLS estimates is 0.78. That between OP and fixed effects estimates is 0.72 while it is 0.89 between OLS and fixed effects. Finally the correlation between TFP^3 estimated with OP in 1996 and mean wages (W^1) is 0.67. This correlation goes to 0.74 after correcting wages of sector effects (W^2) and 0.76 after correcting wages of sector and worker effects (W^3).

Results

Table 9 presents the results of three regressions for our three measures of local OP productivity: TFP^1, the mean productivity computed in (17), TFP^2, the local productivity controlling for sector fixed-effects as estimated in (18), and TFP^3, the local productivity estimated in (20) which conditions out sector fixed-effects and localisation economies. 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 year effects and amenities using OLS. The mean elasticity of TFP with respect to wages is at 5.1% for mean productivity, 4.1% when taking out sector effects, and 4.6% when also

Table 9. Local productivity (OP) as a function of density and market potential: OLS and historical instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
	TFP^1 OLS	TFP^2 OLS	TFP^3 OLS	TFP^1 OLS	TFP^2 OLS	TFP^3 OLS	TFP^1 TSLS	TFP^2 TSLS	TFP^3 TSLS
ln(density)	0.051	0.041	0.046	0.046	0.036	0.040	0.038	0.029	0.034
	(0.0038)^a	(0.0019)^a	(0.0020)^a	(0.0040)^a	(0.0019)^a	(0.0019)^a	(0.0057)^a	(0.0020)^a	(0.0022)^a
ln(market pot.)	-	-	-	0.021	0.020	0.021	0.026	0.028	0.028
	-	-	-	(0.0072)^a	(0.0039)^a	(0.0041)^a	(0.0083)^a	(0.0043)^a	(0.0045)^a
Instruments used:									
ln(1831 density)	-	-	-	-	-	-	Y	Y	Y
ln(1881 density)	-	-	-	-	-	-	Y	Y	Y
ln(1831 m. pot.)	-	-	-	-	-	-	Y	Y	Y
First stage statistics	-	-	-	-	-	-	1863	1863	1863
Over-id test p-value	-	-	-	-	-	-	0.71	0.52	0.73
R-squared	0.44	0.58	0.62	0.46	0.61	0.64	-	-	-

All regressions include a constant, 4 amenity controls, and year effects.
Standard errors clustered by employment area in parentheses.
2448 observations for each regression.
a, b, c: significant at 1%, 5%, 10%. Cragg-Donald statistic reported for the first stage.

controlling for the local sector structure. Controlling for sector effects seem to make very little difference although it appears to improve the precision of the estimates.

Adding market potential among explanatory variables in columns 4, 5, and 6 lowers the density elasticity by 0.5% point. The elasticity of local productivity to market potential is around 2% and very stable across all three measures of local productivity. In columns 7, 8, and 9, we repeat the same regression but we instrument density and market potential with 1831 and 1881 urban population densities and 1831 market potential. Compared to their corresponding OLS coefficients, the TSLS coefficients for employment density are 0.5-0.8% point lower while those on market potential increase by roughly the same amount.

Comparing these results to those of table 6, we note the following. First, the coefficient on employment density does not decrease as much as with wages when we consider more 'purged' measures of productivity. To a large extent, we believe that this is because we do not control for unobserved labour quality in TFP as we do with wages.[25] In other words, we cannot confirm the importance of the endogenous quality of labour bias because of a lack of a reliable method to do so. Then, adding market potential as control has the same effect as with wages. That is, we confirm the importance of different spatial scales at which agglomeration effects take place. Finally, instrumenting employment density and market potential by long lags also has the same effect as with wages. Hence, we also confirm the existence of a mild endogenous quantity of labour bias.

[25]We could do so by introducing firm fixed effects when estimating TFP. However, this approach usually leads to implausible factor coefficients because capital adjusts slowly. In turn this is likely to create problems when regressing TFP on local characteristics.

Table 10. Local productivity (OLS) as a function of density and market potential: OLS and historical instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
	TFP[1] OLS	TFP[2] OLS	TFP[3] OLS	TFP[1] OLS	TFP[2] OLS	TFP[3] OLS	TFP[1] TSLS	TFP[2] TSLS	TFP[3] TSLS
ln(density)	0.031	0.030	0.033	0.026	0.025	0.027	0.019	0.018	0.020
	(0.0017)[a]	(0.0017)[a]	(0.0017)[a]	(0.0016)[a]	(0.0016)[a]	(0.0016)[a]	(0.0019)[a]	(0.0018)[a]	(0.0019)[a]
ln(market pot.)	-	-	-	0.023	0.023	0.023	0.029	0.029	0.029
	-	-	-	(0.0030)[a]	(0.0030)[a]	(0.0031)[a]	(0.0034)[a]	(0.0034)[a]	(0.0035)[a]
Instruments used:									
ln(1831 density)	-	-	-	-	-	-	Y	Y	Y
ln(1881 density)	-	-	-	-	-	-	Y	Y	Y
ln(1831 m. pot.)	-	-	-	-	-	-	Y	Y	Y
First stage statistics	-	-	-	-	-	-	2340	2340	2340
Over-id test p-value	-	-	-	-	-	-	0.55	0.56	0.73
R-squared	0.51	0.49	0.52	0.55	0.53	0.60	-	-	-

All regressions include a constant, 4 amenity controls, and year effects.
Standard errors clustered by employment area in parentheses.
3060 observations for each regression.
a, *b*, *c*: significant at 1%, 5%, 10%. Cragg-Donald statistic reported for the first stage.

Table 11. Local productivity (FIXED EFFECT) as a function of density and market potential: OLS and historical instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
	TFP[1] OLS	TFP[2] OLS	TFP[3] OLS	TFP[1] OLS	TFP[2] OLS	TFP[3] OLS	TFP[1] TSLS	TFP[2] TSLS	TFP[3] TSLS
ln(density)	0.027	0.027	0.029	0.020	0.020	0.022	0.014	0.013	0.015
	(0.0017)[a]	(0.0018)[a]	(0.0018)[a]	(0.0015)[a]	(0.0015)[a]	(0.0016)[a]	(0.0019)[a]	(0.0019)[a]	(0.0020)[a]
ln(market pot.)	-	-	-	0.029	0.029	0.030	0.034	0.036	0.036
	-	-	-	(0.0027)[a]	(0.0028)[a]	(0.0029)[a]	(0.0031)[a]	(0.0032)[a]	(0.0033)[a]
Instruments used:									
ln(1831 density)	-	-	-	-	-	-	Y	Y	Y
ln(1881 density)	-	-	-	-	-	-	Y	Y	Y
ln(1831 m. pot.)	-	-	-	-	-	-	Y	Y	Y
First stage statistics	-	-	-	-	-	-	2340	2340	2340
Over-id test p-value	-	-	-	-	-	-	0.87	0.76	0.72
R-squared	0.51	0.49	0.52	0.58	0.57	0.58	-	-	-

All regressions include a constant, 4 amenity controls, and year effects.
Standard errors clustered by employment area in parentheses.
3060 observations for each regression.
a, *b*, *c*: significant at 1%, 5%, 10%. Cragg-Donald statistic reported for the first stage.

It is also interesting to note that the coefficients on density and market potential take approximately the same values in tables 6 and 9 when either wages or productivity are used as dependent variable. In light of equation (5), this is puzzling since a higher coefficient was expected for wages. A first reason for this difference may be due to the fact that estimating TFP with OP selects firms which appear during two consecutive years (at least) and make positive investments. Although the OP approach corrects for the endogenous choice of inputs, it may bias our results if firms' survival rates and investment behaviour differs across areas. In tables 10 and 11, we reproduce the same regressions as in table 9 using local productivity indices constructed from OLS estimates of (15) and from (22) which computes local productivity fixed effects. The comparison between tables 9, 10, and 11 confirms all the findings of table 9. The main difference is that the OP approach yields density elasticities which are systematically 1.5 to 2% points higher than those obtained with OLS while the latter are always about 0.5% point higher than those estimated through local productivity fixed effects. Using equation (5) and the fact that the labour shares we estimate are typically around 0.6 implies that the density elasticities of wages are roughly consistent with our productivity results when TFP is estimated through fixed effects or by OLS.

[[More on the difference between OP and OLS/FE results very soon]]

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 (recognising that log employment in an industry is the sum of log total employment and the log of the share).

Next, table 12 reports results for a number of instrumental regressions using geological variables as instruments. As with wages, this is a difficult exercise because geological variables tend to be on the weak side. As a result, we expect the coefficient on density to be imprecisely estimated. Because employment density and market potential are strongly correlated, we also expect some instability in the coefficients of instrumented density and market potential.

In columns 1, 2, and 3, we regress our three measures of productivity on employment density and market potential instrumenting them with the dummies for the water capacity

Table 12. Local productivity (OP) as a function of density and market potential: geological instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	TFP^1 TSLS	TFP^2 TSLS	TFP^3 TSLS	TFP^3 LIML	TFP^3 LIML	TFP^3 LIML	TFP^3 LIML	TFP^3 LIML
ln(density)	0.033	0.024	0.029	0.029	0.039	0.036	0.004	0.076
	(0.016)^b	(0.009)^a	(0.009)^a	(0.010)^a	(0.014)^a	(0.016)^b	(0.026)	(0.046)
ln(market pot.)	0.030	0.025	0.024	0.024	0.009	0.013	0.058	-0.065
	(0.017)^c	(0.009)^a	(0.009)^b	(0.010)^b	(0.025)	(0.026)	(0.032)^c	(0.096)
Instruments used:								
Subsoil water capacity	Y	Y	Y	Y	N	N	N	N
Soil carbon content	Y	Y	Y	Y	Y	Y	N	N
Ruggedness	N	N	N	N	Y	N	N	Y
Depth to rock	N	N	N	N	N	N	Y	N
Subsoil mineralogy	N	N	N	N	N	N	N	Y
First stage statistics	14.0	14.0	14.0	14.0	8.4	9.4	8.1	2.7
Over-id test p-value	0.99	0.99	0.96	0.31	0.35	0.31	0.57	0.85

All regressions include a constant, 4 amenity controls, and year effects.
Standard errors clustered by employment area in parentheses.
2448 observations for each regression.
a, b, c: significant at 1%, 5%, 10%. Cragg-Donald statistic reported for the first stage.

of the subsoil and carbon content. The results are very close to those of the last three columns of table 9 when instrumenting with historical variables. Given our concerns about the geological instruments being weak when instrumenting for employment density, column 4 reproduces the regression of column 3 using LIML rather than TSLS. The results are exactly the same. Columns 5 to 8 perform the same estimation as column 4 but use different sets of geological instruments. The results in these columns do not differ significantly from those of column 4. These columns also make it clear that, as the instruments get weaker, the coefficients become less precisely estimated. Other combinations of instruments give similar results but the over-identification tests often fail.

These results obtained with geological instruments are very similar to those obtained with wages in table 7. It is also interesting to note that the sets of instruments for which the Anderson-Rubin over-identification test is passed are roughly similar in table tables 7 and 12.

In table 13, we consider historical and geological instruments simultaneously. We proceed as in table 8 above with wages and retain 1831 density as historical instrument in all regressions. Across all columns in table 13 we can see that the coefficient on employment density is very stable between 2.8 and 3.9%. The results for this coefficient confirm previous results showing that the IV coefficients on employment density are below their OLS counterparts. In turn this suggests a small endogenous quantity of labour bias. The coefficient on market potential in table 13 is less stable than that on employment density. This instability is manifest when using topsoil or subsoil mineralogy in columns 7 and 8.

Table 13. Local productivity (OP) as a function of density and market potential: geological and historical instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	TFP[3] TSLS	TFP[3] LIML	TFP[3] TSLS	TFP[3] TSLS	TFP[3] TSLS	TFP[3] TSLS	TFP[3] TSLS	TFP[3] TSLS
ln(density)	0.036	0.036	0.034	0.034	0.035	0.036	0.039	0.028
	(0.0039)[a]	(0.0039)[a]	(0.0026)[a]	(0.0027)[a]	(0.0029)[a]	(0.0025)[a]	(0.0060)[a]	(0.0046)[a]
ln(market pot.)	0.012	0.013	0.020	0.025	0.021	0.016	-0.004	0.059
	(0.015)	(0.015)	(0.0071)[a]	(0.0079)[a]	(0.0094)[b]	(0.0063)[a]	(0.029)	(0.021)[a]
Instruments used:								
ln(1831 density)	Y	Y	Y	Y	Y	Y	Y	Y
Soil carbon content	Y	Y	N	N	N	N	N	N
Subsoil water capacity	N	N	Y	N	N	N	N	N
Depth to rock	N	N	N	Y	N	N	N	N
Dominant parent material (6 cat.)	N	N	N	N	Y	N	N	N
Dominant parent material (20 cat.)	N	N	N	N	N	Y	N	N
Subsoil mineralogy	N	N	N	N	N	N	Y	N
Topsoil mineralogy	N	N	N	N	N	N	N	Y
First stage statistics	48.1	48.1	183.9	200.6	109.0	88.0	20.8	27.0
Over-id test p-value	0.89	0.60	0.78	0.20	0.28	0.10	0.31	0.22

All regressions include a constant, 4 amenity controls, and year effects.
Standard errors clustered by employment area in parentheses.
1836 observations for each regression.
a, *b*, *c*: significant at 1%, 5%, 10%. Cragg-Donald statistic reported for the first stage.

It is due in these two cases to the instruments being on the weak side for market potential with a first-stage F statistic between 8 and 10. Nonetheless, none of coefficients on market potential in this table differs significantly from an elasticity of TFP with respect to market potential of 2.8% as obtained in the last column of table 9. Finally in Appendix A, table 15 replicates a number of these regressions but using TFP estimated with OLS and fixed effects to obtain similar results.

6. Conclusions

We revisit the estimation of urbanisation economies 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 fixed effect approach.

Our findings are the following. The first two relate to problems of specification. (i) Agglomeration effects matter at different spatial scales. The specification should reflect this and use measures of size both at a very local level and at a higher level of spatial aggregation. Because it is more robust to the spatial zoning, our preferred measure of local size is employment density. To capture scale at broader spatial scales, we use a measure of market potential. (ii) Although we are interested in urbanisation effects, we control for sector effects. We find that sector effects matter. However, controlling or not for them does not seem to affect the estimation of urbanisation economies.

Our second series of findings relate to the endogenous quality of labour bias. (iii) Long lags of our endogenous explanatory variables make for strong instruments. (iv) Geological variables are more complicated instruments to play with. As instruments, they are weaker than historical variables. For the geological instruments we use here, this is particularly true for employment density. Instrumenting market potential with geological variables is much easier. (v) Nevertheless, geological and historical instruments lead to similar answers for the coefficients on employment density and market potential. The simultaneity problem between employment density and local wages / productivity is small. We find no evidence of simultaneity for market potential.

Finally, our last finding relates to the endogenous quality of labour bias. (vi) 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 or the specification issues discussed above. In our regressions, we address sorting using the panel dimension of our wage data. Applying this type approach to TFP is problematic. We thus put more weight on our wage results than we do on our TFP results. 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 productivity to density is at 1.5% and at 2-2.5% for the density elasticity of wages.

Appendix A. Further results

Table 14. Local wages as a function of density and exogenous market potential: geological and historical variables

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	W^3	W^3	W^3	W^3	W^3	W^3	W^3	W^3
	TSLS	LIML	TSLS	TSLS	TSLS	TSLS	TSLS	TSLS
ln(density)	0.021	0.021	0.021	0.021	0.022	0.021	0.021	0.021
	(0.0018)[a]	(0.0018)[a]	(0.0017)[a]	(0.0018)[a]	(0.0018)[a]	(0.0018)[a]	(0.0018)[a]	(0.0017)[a]
ln(market pot.)	0.032	0.032	0.032	0.032	0.032	0.032	0.032	0.032
	(0.0033)[a]	(0.0033)[a]	(0.0032)[a]	(0.0033)[a]	(0.0032)[a]	(0.0033)[a]	(0.0033)[a]	(0.0032)[a]
Instruments used:								
ln(1831 density)	Y	Y	Y	Y	Y	Y	Y	Y
Subsoil mineralogy	Y	Y	N	N	N	N	N	N
Depth to rock	N	N	Y	N	N	N	N	N
Soil differentiation	N	N	N	Y	N	N	N	N
Topsoil water capacity	N	N	N	N	Y	N	N	N
Subsoil water capacity	N	N	N	N	N	Y	N	N
Dominant parent material (6 cat.)	N	N	N	N	N	N	Y	N
Erodibility	N	N	N	N	N	N	N	Y
First stage statistics	108.2	108.2	117.1	129.3	115.8	71.4	55.9	68.3
Over-id test p-value	0.75	0.32	0.45	0.29	0.22	0.48	0.16	0.47

All regressions include a constant, 4 amenity controls, and year effects.
Standard errors clustered by employment area in parentheses.
1836 observations for each regression.
a, *b*, *c*: significant at 1%, 5%, 10%. Cragg-Donald statistic reported for the first stage.

Appendix B. Implementation of Olley and Pakes (1996)

The error term in (15) is rewritten as $\varepsilon_{it} \equiv v_{it} + \xi_{it}$ where v_{it} is the part of the error term that influences the decision of the firm regarding its factors and ξ_{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 (15) can then be rewritten as:

$$\ln va_{it} = \alpha \ln k_{it} + \beta \ln l_{it} + \sum_j \beta_j^S q_{ijt} + \theta_t + f_t^{-1}(k_{it}, I_{it}) + \xi_{it}. \tag{A 1}$$

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

$$\ln va_{it} = \beta \ln l_{it} + \sum_j \beta_j^S q_{ijt} + \Phi_t(k_{it}, I_{it}) + \xi_{it}. \tag{A 2}$$

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

Table 15. Local productivity (OLS and FE) as a function of density and market potential: geological and historical instruments

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	$\text{TFP}^3_{\text{OLS}}$ TSLS	$\text{TFP}^3_{\text{OLS}}$ TSLS	$\text{TFP}^3_{\text{OLS}}$ TSLS	$\text{TFP}^3_{\text{OLS}}$ TSLS	TFP^3_{FE} TSLS	TFP^3_{FE} TSLS	TFP^3_{FE} TSLS	TFP^3_{FE} TSLS
ln(density)	0.023	0.021	0.022	0.023	0.018	0.017	0.018	0.018
	(0.0034)[a]	(0.0022)[a]	(0.0025)[a]	(0.0045)[a]	(0.0033)[a]	(0.0023)[a]	(0.0024)[a]	(0.0041)[a]
ln(market pot.)	0.012	0.027	0.015	0.011	0.019	0.030	0.022	0.019
	(0.013)	(0.0061)[a]	(0.0067)[b]	(0.020)	(0.012)	(0.0056)[a]	(0.062)[a]	(0.018)
Instruments used:								
ln(1831 density)	Y	Y	Y	Y	Y	Y	Y	Y
Soil carbon content	Y	N	N	N	Y	N	N	N
Subsoil water capacity	N	Y	N	N	N	Y	N	N
Depth to rock	N	N	Y	N	N	N	Y	N
Subsoil mineralogy	N	N	N	Y	N	N	N	Y
First stage statistics	60.3	230.4	250.8	200.6	26.2	230.4	250.8	26.2
Over-id test p-value	0.96	0.09	0.13	0.42	0.80	0.15	0.20	0.13

All regressions include a constant, 4 amenity controls, and year effects.
Standard errors clustered by employment area in parentheses.
1836 observations for each regression.
a, b, c: significant at 1%, 5%, 10%. Cragg-Donald statistic reported for the first stage.

some estimators for the labour coefficient, $\widehat{\beta}$, and the skill share coefficients, $\widehat{\beta}^S_j$. It is then possible to construct $z_{it} \equiv \ln va_{it} - \widehat{\beta} \ln l_{it} - \sum_j \widehat{\beta}^S_j q_{ijt}$. Furthermore, the error v_{it} is rewritten as a the projection on its lag and an innovation: $v_{it} \equiv h(v_{it-1}) + \zeta_{it-1}$. Using the fact that $v_{it-1} = f^{-1}_{t-1}(k_{it-1}, I_{it-1}) = \Phi_{t-1}(k_{it-1}, I_{it-1}) - \alpha \ln k_{j,t-1} - \theta_{t-1}$, the value-added equation then becomes:

$$z_{it} = \alpha \ln k_{it} + \theta_t + h\left(\widehat{\Phi}(k_{it-1}, I_{it-1}) - \alpha \ln k_{jt-1} - \theta_{t-1}\right) + \psi_{it}, \qquad \text{(A 3)}$$

where ψ_{it} is a random error. The function $h(.)$ 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 $\widehat{\alpha}$ and the year dummies $\widehat{\theta}_t$. Firm TFP is then defined as $r_{it} \equiv z_{it} - \widehat{\alpha} \ln k_{it} - \widehat{\theta}_t$. 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 and the only relevant way to do so is to define it as $I_{it} = k_{it} - k_{it-1}$. As a consequence investment can be computed for a given firm only when it is present in two consecutive years. Observations for which it is not the case 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. This can happen, for instance, if there is greater 'churning' (i.e. entry and exits) in denser areas, and age and investment affect productivity positively. Then, the double selection associated with the implementation of OP leads more establishments with a low productivity to be dropped in high density areas. In turn, this increases the measured difference in local productivity between areas of low and high density.

References

Ackerberg, 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. 2007. Skills in the city. Processed, University of Toronto.

Bayer, Patrick, Stephen L. Ross, and Giorgio Topa. 2005. Place of work and place of residence: Informal hiring networks and labor market outcomes. Processed, University of Connecticut.

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. 2008. Spatial wage disparities: Sorting matters! *Journal of Urban Economics* 63(forthcoming).

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Sébastien Roux, and Diego Puga. 2007. The productivity advantages of large markets: Distinguishing agglomeration from firm selection. Work in progress, University of Toronto.

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. Forthcoming.

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.

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. 2007. The costs of remoteness: Evidence from German division and reunification. Processed, London School of Economics.

Roback, Jennifer. 1982. Wages, rents and the quality of life. *Journal of Political Economy* 90(6):1257–1278.

Rosenthal, Stuart S. and William C. Strange. 2004. Evidence on the nature and sources of agglomeration economies. In Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North-Holland, 2119–2171.

Rosenthal, Stuart S. and William C. Strange. 2006. The attenuation of human capital spillovers. Processed, Univserity of Toronto.

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.