SERC
Spatial Economics Research Centre

SERC DISCUSSION PAPER 31

# Agglomeration Economies and Labour Productivity: Evidence from Longitudinal Worker Data for GB's Travel-to-Work Areas

Patricia C. Melo (Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College London)
Daniel J. Graham (Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College London)

**October 2009**

This work was part of the research programme of the independent UK Spatial Economics Research Centre funded by the Economic and Social Research Council (ESRC), Department for Business, Enterprise and Regulatory Reform (BERR), the Department for Communities and Local Government (CLG), and the Welsh Assembly Government. The support of the funders is acknowledged. The views expressed are those of the authors and do not represent the views of the funders.

# Agglomeration Economies and Labour Productivity: Evidence from Longitudinal Worker Data for GB's Travel-to-Work Areas

**Patricia C. Melo* and Daniel J. Graham****

**October 2009**

* Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College London
** Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College London

Acknowledgements
Please note that this work contains statistical data from ONS which is Crown copyright and reproduced with the permission of the controller of HMSO and Queen's Printer for Scotland. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

Abstract

This paper analyzes the impact of agglomeration externalities on hourly earnings using longitudinal worker micro-level data from the Annual Survey of Hours and Earnings over the period 2002- 2006. We find that the effect of agglomeration externalities on wages is sensitive to the estimator used. Controlling for nonzero correlation between workers' unobservable skills and other covariates halves the size of the wage elasticity of agglomeration externalities. On the contrary, accounting for firms' unobservable heterogeneity has only a weak contribution to the explanation of wage differentials. Another interesting result is that correcting for reverse causality between productivity and agglomeration does not appear to have a substantial impact on the magnitude of the parameter estimates. Our best estimate for the effect of labour market density (market potential) is 0.8% (5.8%). This means that doubling labour market's employment density can raise hourly earnings by nearly 1%, while halving the distances to other markets produces an increase of hourly wages of nearly 3%. The last piece of evidence refers to the spatial attenuation of agglomeration externalities. We estimate that a 100,000 increase in the number of jobs within 5 kilometres raises hourly wages by approximately 1.19%; the effect falls sharply thereafter.

Keywords: Agglomeration externalities, wages, endogeneity, sorting.
JEL Classifications: J31, R12, R23

# 1. Introduction

In this paper we look at individual hourly wage rates to investigate the importance of agglomeration economies. We ask four questions. The first question relates to the degree to which spatial wage disparities result from agglomeration externalities: we test for the effect of labour market scale and access to product markets on workers' hourly earnings. The second question concerns the sensitiveness of the effects from agglomeration externalities to both workers' and firms' unobservable heterogeneity. Thirdly, we are also interested in the extent to which correcting for reverse causality between earnings and agglomeration affects the magnitude of the wage elasticities of agglomeration externalities. Finally, we ask whether agglomeration externalities decay with increasing distance, and investigate the spatial scale over which its effects are likely to operate. It is worth mentioning that our regression analyses contemplate the whole aggregate economy as well as various industry groups.

The first question is addressed through a regression of individual worker wages on a set of variables that proxy for agglomeration externalities, controlling for time-varying worker characteristics and other factors that influence wages. The initial results for the effect of labour market scale and market potential indicate that doubling the employment density of a given labour market raises hourly wages by about 2.2%, whereas doubling the market potential of a given labour market increases hourly wages of around 10.0%.

To evaluate the impact of workers' and firms' unobservable heterogeneity on the parameter estimates for agglomeration externalities, we use a linear fixed-effects model that accounts for both worker- and firm-specific heterogeneity. Controlling for workers unobservable heterogeneity halves the size of the effects from labour market scale and market potential to 1.0% and 5.4% respectively. In contrast, further accounting for firm level unobservable heterogeneity has only a very weak impact on the estimates: the wage elasticity of employment density (market potential) further reduces to 0.72% (5.1%). Another interesting finding is that correcting for the endogeneity between earnings and agglomeration through instrumental variables estimators appears to affect the magnitude of agglomeration externalities only slightly: the preferred elasticity of employment density (market potential) is approximately 0.8% (5.8%).

Finally, to address the fourth question we examine the spatial attenuation of agglomeration externalities by testing for the effect of proximity to jobs in concentric distance bands around the centroid of worker's workplace. We find that a 100,000 increase in the number of jobs within 5 kilometres raises hourly wages by approximately 1.19%; this effect falls sharply (0.38%) if the increase in the number of jobs occurs 10 kilometres away, and remains around 0.15% if the increase in the number of jobs occurs between 10 to 20 kilometres.

This paper attempts to contribute to the empirical literature in several ways. Previous studies have generally paid scant attention to the importance of workers' unobservable skills. Our analysis considers the role of both workers' and firms' unobservable heterogeneity. We also provide evidence on the spatial sorting of more able workers across space. Another contribution from our work is that it provides further evidence on the nature and magnitude of the endogeneity bias from reverse causality between productivity and agglomeration economies not only for the aggregate economy but also for a comprehensive set of economic sectors. Finally, we add to the empirical literature by generating new evidence on the spatial decay and scope agglomeration externalities.

The paper is organized as follows. Section 2 briefly summarises the main results and characteristics of the empirical research on the effects of agglomeration economies on productivity through estimation of wage models. The theoretical framework and econometric estimation are described in section 3. Section 4 describes the data and variables used in our analysis. Section 5 discusses the results of our estimations and section 6 provides some concluding remarks.

## 2. Overview of previous empirical evidence

There is general agreement on positive effects from agglomeration externalities but magnitudes can differ quite substantially both within and between studies (see Rosenthal and Strange, 2004 for a descriptive review of the empirical literature; and Melo et al., 2009 for a meta-analysis of the empirical literature). Although the empirical evidence is mostly produced by the agglomeration literature within the field of regional and urban economics, other strands of literature have contributed with estimates for the productivity benefits from agglomeration economies (e.g. labour economics with estimates of the "urban wage premium" from *Mincerian* wage models augmented with measures of agglomeration economies).

Studies generally build on one or several alternative theories for the observation of spatial differences in workers' productivity levels and wages. For the purpose of our analysis, the theory we wish to test for is that workers are paid more in large and denser markets because they are more productive there due to the presence of agglomeration economies. From the point of view of individual firms, the relevant question is why they prefer to be located in areas where input costs are higher. The theory of agglomeration economies answers that it is because firms enjoy efficiency gains from more productive workers. Alternative theories propose that larger urban areas pay higher wages because they attract more skilled workers. If more able workers self-select into more productive markets, then part of the effect from urban agglomeration externalities is due to workers unobserved skills: the "omitted ability bias" problem, which reflects a process of spatial self-selection into more educated and more productive areas (see Glaeser and Mare, 2001; Wheeler, 2001; Yankow, 2006). There is also the hypothesis that workers are paid more because of knowledge spillovers in large and denser areas, which arise from externalities in learning and human capital investment (see Moretti, 2004 for a recent survey of this literature). This theory is related to the agglomeration economies hypothesis because knowledge spillovers are one of the *Marshallian* mechanisms through which agglomeration economies arise (Rosenthal and Strange, 2006, p.18).

Table 1 provides a summary of previous evidence for the effect of urban agglomeration externalities on wages. The great majority of existing evidence is for the USA. Sveikauskas (1975) estimates that doubling the population of metropolitan areas in the USA causes earnings to raise between 1.2% to 8.6%. Segal (1976) estimates that income in the largest USA metropolitan areas (with two million or more inhabitants) is 8% higher than in the remaining metropolitan areas. Wheeler (2001) finds that USA metropolitan area population increases hourly wages by 2.7%. Glaeser and Mare (2001) find that earnings in dense (non-dense) USA metropolitan cities can be up to 28% (15%) higher than earnings in non-urban areas. Yankow (2006) reaches similar conclusions: earnings in USA big cities (small cities) can be up to 22% (10%) higher than in non-urban areas.

Nevertheless, there is also some evidence for Europe. Evidence from regional level studies also proposes positive effects from agglomeration on regional economic performance. Fingleton (2003) and Fingleton (2006) estimate wage elasticities of employment density between 1.4% and 4.9% for Local Authority Districts (LAUD) in GB. Rice et al. (2006) estimate a wage elasticity with respect to economic mass of about 5% for GB NUTS3 regions. Evidence from worker level data suggest somewhat smaller elasticity values. Combes et al. (2008a,b) estimate that doubling employment density of French employment areas raises wages by between 2-3%. Mion and Naticchioni (2005) estimate smaller effects for Italian Provinces: doubling employment density raises wages by only 0.22%.

The most common approach in the empirical literature has been to measure agglomeration economies with total population/employment or density. The main limitation of using measures of total size/density is their failure to tell something about the spatial distribution of the effects from agglomeration externalities. To overcome this gap, some of the more recent studies have experimented with "market potential" type measures that capture both the size of and the proximity to economic activity. Market potential measures allow for the effects of agglomeration externalities to be realised over space and diminish with increased distance. The standard definition of the market potential of location *r* can be written as follows:

$$MP_r = \sum_j \frac{emp_j}{d_{rj}{}^{\alpha}}, \tag{1}$$

where $emp_j$ measures the economic size of area $j$, $d_{rj}$ represents the spatial separation between any two locations $r$ and $j$, and $\alpha$ is the spatial decay gradient. Most of these studies, however, tend to assume a one-unit distance decay parameter, implying that the effect from increasing distance is linear. Some studies include own-area size to obtain a full measure of agglomeration effects (e.g. Graham and Kim, 2007), whereas others exclude own-area size to capture effects related to access to product markets (e.g. Mion and Naticchioni, 2005; Combes et al., 2008a; Combes et al., 2008b).

Graham and Kim (2007) estimate an elasticity of wages with respect to market potential of 0.84% for manufacturing; the values for individual industries range between -13.3% for the manufacture of basic metals and fabricated metal products and 14.3% for the manufacture of office machinery and computers. The average elasticity for the service sectors is about 3.97%, while the values for the disaggregated sectors range between -13% for the motion picture, video, radio, and television activities and 17.3% for the wholesale and retail trades activities. Combes et al. (2008a) and Combes et al. (2008b) suggest that doubling the size of a given employment area market potential increases wages by about 2.4% and 3.4% respectively. Mion and Naticchioni (2005) estimate an elasticity of wages with respect to market potential of similar size:3.19%.

The more common alternative approach is to divide total market potential defined above into a set of consecutive distance bands (e.g. Di Addario and Patacchini, 2007; Rosenthal and Strange, 2008). This method is attractive because it can inform on the spatial scale within which agglomeration externalities are likely to be significant and the relative importance of these effects across different distance intervals.

Rosenthal and Strange (2008) investigate the spatial reach of agglomeration externalities by regressing US worker wages on the total employment contained within concentric distance rings from workers' place of work. They find that wages increase by 1.5% to 2.14% for an additional 100,000 full-time workers within 5 miles (about 8 kilometres) from the workers' place of work and fall sharply thereafter: 0.52% (5-25 miles/8-40 kilometres), 0.84% (25-50 miles/40-80 kilometres), and 0.20% (50-100 miles/80-160 kilometres). The elasticity of wages with respect to jobs within the first distance ring is between 0.031 and 0.047: doubling total employment within 5 miles increases wages by 3.1% to 4.7%.

Di Addario and Patacchini (2007) and Fu (2007) follow Rosenthal and Strange's approach. Di Addario and Patacchini (2007) fit wage models based on worker level data for Italy and estimate that an increase of 100,000 inhabitants within 4 kilometres from workers' residence raises wages by 0.1%-0.2%, but the increase falls sharply thereafter. Moreover, the impact is found to be significant only up to 12 kilometres, which is less than the average radius of Italian local labour markets (14.7 kilometres); this suggested to the authors that agglomeration economies are likely to occur within local labour markets. Fu (2007) also uses individual level data from the 1990 Massachusetts Census to test for the spatial attenuation of human capital related externalities. The findings indicate that human capital externalities in cities exhibit a very localized pattern (generally within 6 miles/9.7 kilometres) and can decay sharply with increasing distance from workers' location.

Only very few papers estimate the decay gradient α (e.g. Rice et al., 2006; Amiti and Cameron, 2007; and Graham et al., 2009). Graham et al. (2009a) is the only study that provides decay gradients for various industry sectors.

Rice et al. (2006) investigate the spatial scope of agglomeration forces by using a measure of economic mass that considers access to population of working age within various driving time intervals. Using an exponential function, they estimate the rate of decay to be 1.37 (using an instrumental variables (IV) estimator) and 1.51 (using a nonlinear least squares estimator (NLS)). These values imply that moving the population of working age 30 (60) minutes further away decreases the impact of economic mass on productivity by about 75% (94%) and 78% (95%) respectively.

They also consider the decay of the effect of economic mass on wages: the rates of decay estimated range between 1.20 and 1.41, which implies that moving the population of working age 30 minutes further away decreases the impact of economic mass on average hourly earnings by 70% to 76%. This means that an individual of working age between 60-70 minutes away captures only 24% to 30% of the effects of economic mass on its average hourly wage, compared to an individual of working age within 30-40 minutes of driving time away.

Graham et al. (2009a) estimate firm level production functions and obtain a decay gradient for the Great Britain economy of about 1.66. The decay gradient of the market potential measure differs across economic sectors; the value is higher for service industries and smaller for manufacturing. Business services and consumer services have a decay gradient of 1.75 and 1.82 respectively, whereas for manufacturing the value is 1.10. This supports a steeper spatial attenuation of agglomeration externalities for the service sectors, which are generally more dependent on urbanisation levels.

Amiti and Cameron (2007) focus on the spatial attenuation of the agglomeration effects arising from access to suppliers and markets. They use NLS estimators and find that the distance-decay parameter for the supplier and market access, using an exponential function, is 1.79 and 2.81 respectively. The findings indicate that only 10% of the benefits of supplier and market access extend beyond 129 and 82 kilometres. The benefits from market access are substantially more localized than those from supplier access, but comparing the distance parameters across years reveals that the market access externality appears to have become less localized over time (fall of the distance-decay parameter), whereas the supply access externality appears to have become more localized over time (increase of the distance-decay parameter).

To summarise, the existing evidence generally suggests a localised geographic scope of agglomeration externalities. The findings on the geographic scope of agglomeration effects can differ depending on which particular mechanism is being considered. Studies looking at knowledge spillovers and human capital externalities tend to agree on a very short geographic scope, while studies looking at input sharing linkages find that the geographic scope for these interactions is much wider. Unfortunately, there is not enough evidence on the actual spatial attenuation of the effects from each mechanism that we can use to make definite conclusions.

**Table 1: Evidence on the effect of agglomeration economies on wages**

| Author(s) | Measure of agglomeration | Country | Spatial unit | Industry | Estimates |
|---|---|---|---|---|---|
| Lewis and Prescott (1974) | Employment | USA | SMSAs | M | β=9.4%[a] |
| Sveikauskas (1975) | Population | USA | SMSAs | M | ε=1.2-8.6% |
| Segal (1976) | Population | USA | SMSAs | E | β=8%[b] |
| Fogarty and Garofalo (1978) | Population | USA | SMSAs | E | β=9.93%;β=9.24%[b] |
| Moomaw (1981) | Population | USA | SMSAs | M;E | ε=5.98%;2.68% |
| Diamond and Simon (1990) | Population | USA | Cities | M/NM | β=0.6-2.0%[c] |
| Carlino and Voith (1992) | Share population in metropolitan area | USA | States | E | β=0.22-1.68%[d] |
| | Square of share of population in metropolitan area | | | | β=-1.74-(-0.08%[d] |
| | Share population in metropolitan area | USA | States | M | β=-0.196-(-0.230)%[d] |
| | Square of share of population in metropolitan area | | | | β=-0.0004-(-0.001)%[d] |
| De Lucio et al. (1996) | Population | Spain | Provinces | M | β=0.09 (in $w_{ij}/w_i$:i-industry, j-provice)[c] |
| | Share of population in large cities (>20,000) | | | | β=0.063/0.065 (in wij/wi:i-industry, j-provice)[d] |
| Adserà (2000) | MSA population | USA | States | M;F;ND;D;E | β=1.6%[e] |
| Graham (2000) | Inside - employment | UK | Counties | M | ε=76.4%;37.8% |
| | Inside - employment density | | | | ε=-1.8%;ε=-0.9% |
| | Outside - Market Potential (employment) | | | | ε=4.3%;ε=-6.8% |
| | Outside - Market Potential (employment density) | | | | ε=-9.6%;ε=-5.8% |
| Tabuchi and Yoshida (2000) | Population of SMEA | Japan | Cities | E | ε=-7%;-12% for real wages;ε= 10% for nominal wages |
| Wheeler (2001) | Population | USA | SMSAs | E | ε=2.7% |
| Glaeser and Mare (2001) | Dense metropolitan city (>500,000) *vs.* rural | USA | SMSAs | E | β=2.6-28.2% |
| | Nondense metropolitan city (<500,000) *vs.* rural | | | | β=7.0-15.3% |
| Costa and Kahn (2001) | Metropolitan area size | USA | SMSAs | E | ε=-0.2-9.1% |

**Table 1: Evidence on the effect of agglomeration economies on wages (*continued*)**

| Author(s) | Measure of agglomeration | Country | Spatial unit | Industry | Estimates |
|---|---|---|---|---|---|
| Wheaton and Lewis (2002) | Industry emp. specialisation (% SMA emp.) | USA | SMSAs | M | ε=2.78% |
| | Industry emp. concentration (% national emp.) | | | | ε=1.5% |
| | Occupation emp. specialisation (% SMA emp.) | | | | ε=3.66% |
| | Occupation emp. concentration (% national emp.) | | | | ε=0.59% |
| Fingleton (2003) | Employment density | GB | LAUD | E | ε=1.58-1.76% |
| Mion and Naticchioni (2005) | Employment density | Italy | Provinces | E | ε:0.22% |
| | Market potential | | | | ε:3.19% |
| Rosenthal and Strange (2008) | Employment in concentric distance rings from place of work | USA | Distance rings from WPUMA | E | ε=4.5% |
| Yankow (2006) | Population size (big city vs. non-urban) | USA | SMSAs | E | β=-1.6-22% |
| | Population size (small city vs. non-urban) | | | | β=1.9-9.5% |
| Wheeler (2006) | Population size | USA | Labour markets (MA/NMC) | E | ε=0.002pp (from +1%) |
| | Population density | | | | ε=0.003pp (from +1%) |
| Fingleton (2006) | Employment density | GB | LAUD | E | ε=1.4-4.9% |
| | Market potential | | | | ε=3.9-5.8% |
| Rice et al. (2006) | Economic mass | GB | NUTS3 | E(NA) | ε=4.96% |
| Di Addario and Patacchini (2007) | Population size | Italy | Local labour markets | E | β=0.1%[e] |
| Combes et al. (2008a) | Employment density | France | Employment areas | E | ε=3.0% |
| | Market potential | | | | ε=2.4% |
| Graham and Kim (2007) | Effective density (straight-line distance) | UK | Wards | M;S | ε=0.84%;ε=3.97% |
| Combes et al. (2008b) | Employment density | France | Employment areas | E | ε=2-4% |
| | Market potential | | | | ε=2-5% |

[a] for an increase in employment of 1thousand, [b] if population in SMSA exceeds 2 million, [c] for an increase in population/employment of 1million, [d] for +1percentage point (pp) in the share of population, [e] for +100,000 residents/workers

ε: elasticity, β: proportionate change, pp: percentage point

E: Whole economy, M: Manufacturing, S: Services, ND: Non-durable goods, D: Durable goods, F: Financial services; NM: Non-manufacturing

LAUD: Local Authority Districts, MA/NMC: Metropolitan Areas/Non-Metropolitan Counties, WPUMA: Work Public Use Micro Area, SMEA: Standard Metropolitan Employment Area, SMSA: Standard Metropolitan Statistical Area

# 3. Theoretical framework and econometric model

## 3.1. The theoretical model

In this section we present the theoretical basis on which we structure our econometric analysis. We draw upon standard economic theory to derive a wage model that allows testing for the importance of agglomeration economies. The theoretical model described below follows closely that used by Combes et al. (2008a). Suppose that the production of a firm located in region *r*, pertaining to industry sector *s*, and using as input factors labour $l_{rs}$ and other input factors $k_{rs}$ can be represented by the equation below.

$$y_{rs} = A_{rs}(s_{rs}l_{rs})^{\mu} k_{rs}^{1-\mu} , \tag{2}$$

where $A_{rs}$ is the total factor productivity and is defined in a Hicks-neutral form, and $s_{rs}$ represents the relative efficiency of labour. If the price of output $y_{rs}$ is $p_{rs}$, the remuneration of labour is $w_{rs}$ and the cost of the other input factors is $r_{rs}$, then, at the competitive equilibrium the price of the input factors should equal the value of their respective marginal products. These are the usual solutions from the firm profit maximisation strategy, where the profit function is given below.

$$\pi_{rs} = p_{rs}A_{rs}(s_{rs}l_{rs})^{\mu} k_{rs}^{1-\mu} - (w_{rs}l_{rs} + r_{rs}k_{rs}), \tag{3}$$

The first order conditions are given by the following expressions:

$$w_{rs} = \mu p_{rs}A_{rs}s_{rs}^{\mu}\left(\frac{k_{rs}}{l_{rs}}\right)^{1-\mu}, \tag{4}$$

and

$$r_{rs} = (1-\mu)p_{rs}A_{rs}s_{rs}^{\mu}\left(\frac{k_{rs}}{l_{rs}}\right)^{-\mu}, \tag{5}$$

Rearranging the second equation in terms of $k_{rs}/l_{rs}$ and substituting it into the wage equation, we obtain the demand curve for labour below:

$$w_{rs} = \mu(1-\mu)^{(1-\mu)/\mu} s_{rs} \left( \frac{p_{rs} A_{rs}}{r_{rs}^{1-\mu}} \right)^{1/\mu}, \tag{6}$$

where the remuneration of labour is positively determined by the quality of the labour force, $s_{rs}$, (e.g. qualification, education and other skills), the output price, $p_{rs}$, and the technological efficiency of the local economy, $A_{rs}$. On the other hand, the level of wages depends negatively on the cost of the other non-labour input factors, $r_{rs}$. It is through these latter three terms that agglomeration and dispersion forces are realized. While the output and input prices, $p_{rs}$ and $r_{rs}$, capture the effects from "pecuniary externalities", the local environmental efficiency, $A_{rs}$ captures the effects from "technological externalities".[1]

In particular, the term $A_{rs}$ captures the effects from "technological externalities" that are usually described in terms of the *Marshallian* externalities: (1) labour market pooling externalities from access to a large pool of skilled workers; (2) input-output linkages from access to large final- and intermediate product markets; and (3) knowledge spillovers from more intensive sharing of information and ideas. Here we can only estimate the combined net overall effect of all three mechanisms: $(p_{rs}A_{rs}/r_{rs}^{1-\mu})$^$(1/\mu)$.

## 3.2. The econometric model and estimation issues

We use the theoretical framework outlined above to investigate the effect of agglomeration economies on labour productivity as measured by individual worker wage rates. The wage equation below expresses the theoretical model above into the empirical model to be estimated:

$$lnw_{it} = \alpha_0 + X_{it}\beta + Z_{r(i)t}\alpha + \varphi lnY_{f(i,t)} + \delta_t + \lambda_{o(i,t)} + \sigma_{s(i,t)} + \omega_{r(i,t)} + \varepsilon_{it}, \tag{7}$$

where $i$ identifies the worker, $r$ identifies the region, $s$ refers to the economic sector, $o$ denotes the occupational group, and $t$ specifies the time period. The dependent variable is the logarithm of real net hourly earnings, that is, the gross hourly earnings discounted of any overtime pay and adjusted for the price level using the Average Earnings Index (AEI) conducted by the ONS. The term $X_{it}$ is a vector of individual worker characteristics commonly considered in labour economics, including worker's age, age squared, gender, and whether the worker is full-time or part-time. The data available allows us to compute for each individual worker the respective home-to-work distance. This is potentially a

[1] The main difference between the two types of externalities is that the former operate through prices, whereas the latter impact on the utility of workers and the production of firms through nonmarket interactions (Fujita and Thisse, (2002)).

valuable covariate because it captures the effect from some of worker's characteristics that we do not observe (e.g. household type, residential preferences, etc.).

Since we also have information about the size of the firm in which individual worker $i$ works, we include the number of workers in the respective firm $f$: $Y_{f(i,t)}$. $Z_{r(i)t}$ denotes a vector containing the variables that measure the net overall effects from agglomeration externalities, which we describe in the next section.

In addition to the explanatory variables described above, we add various sets of dummy variables to control for sources of heterogeneity that can lead to omitted variable bias and inconsistency of the model parameter estimates. To control for macro level changes in wage rates that are common to all individuals we include a time effect, $\delta_t$. To account for differences in workers' wages due to job heterogeneity, we also include indicator variables for occupational groups, $\lambda_{o(i,t)}$; the inclusion of occupation dummies is important because the ASHE does not provide any direct measure of workers' education. Similarly, we control for economic sector heterogeneity with a set of dummy variables for various industry groups, $\sigma_{s(i,t)}$. Because nominal wage levels can also reflect spatial differences in the living cost, the price (and rent) of land, the degree of market competition, etc., we incorporate a set of regional controls, $\omega_{r(i,t)}$. There are other sources of heterogeneity that can affect the consistency of the parameter estimates in the wage equation. Workers' unobserved heterogeneity (i.e. differences in worker skills and years of education) is one potential source of inconsistency. Finally, $\varepsilon_{it}$ is the residual error term assumed to be normally distributed while allowing for heteroskedasticity and clustering on individuals.

The estimation of the effects of agglomeration on productivity is subject to identification and specification problems. We are concerned with the possibility of inconsistent parameter estimates caused by endogenous regressors. In our analysis, this complication can arise from two main sources: (i) correlation between the measures of agglomeration, worker's quality as captured by their occupations, and the error term; (ii) reverse causality between agglomeration economies and worker productivity on one hand, and between home-to-work distances and earnings on the other hand.

The justification for inconsistent parameter estimation in the former case is that workers with higher unobserved ability (captured in the error term) will tend to have better quality and higher pay occupations; similarly, workers with higher unobserved ability will tend to self-select into larger labour markets that are more productive. As a result, the parameter estimate for the effect of urban agglomeration on wages will be composed of its direct effect but also an indirect effect that percolates through the correlation between agglomeration and workers' unobserved ability. The least squares estimator will therefore be both biased and inconsistent. Controlling for time-variant observable heterogeneity in labour force quality through inclusion of occupational dummy variables minimises problems of bias due to spatial sorting, but only partially. To control for endogeneity from self-selection of more skilled workers into more productive labour markets, we employ the within-groups fixed-effects estimator, as used by previous studies (e.g. Glaeser and Mare, 2001; Combes et al., 2008a; Combes et al., 2008b).

The second main estimation issue we face is that of reverse causality between wages and agglomeration, and between wages and home-to-work commutes. Reverse causality between productivity and agglomeration exists because there is a positive feedback between the two: firms and workers migrate to higher productivity areas and this reinforces agglomeration. As a result, agglomeration economies improve economic performance and economic performance reinforces agglomeration, making the latter endogenous. Graham et al. (2009b) estimate panel data vector autoregressions to test for bi-directional causality between productivity and agglomeration and find that agglomeration economies do Granger cause productivity.

The standard strategy implemented to correct for endogeneity consists of using some form of instrumentation technique. The most common approach is to use long-lagged values of population or population density to instrument contemporary values of agglomeration economies (e.g. Ciccone and Hall, 1996; Mion and Naticchioni, 2005; Rice et al., 2006; Combes et al., 2008a; Combes et al., 2008b). The motivation for selecting these variables is that historical values of population are correlated with current levels of urban size but are uncorrelated with current levels of productivity. Whereas the relevance of these instruments can be convincingly good, there is no absolute confidence that the instruments will be exogenous.

Land area and geological data (e.g. fraction of land covered by water, fraction of land underlain by sedimentary rock, fraction of land designated as seismic hazard, etc.) have also become popular instruments for agglomeration (e.g. Ciccone, 2002; Combes et al., 2008b; Rosenthal and Strange, 2008; Di Addario and Patacchini, 2007). The motivation here is that the current patterns of population densities are correlated to the underlying geological features of the various geographies, but such geological features are uncorrelated with contemporaneous productivity levels. Compared to the long-lagged demographic instruments, the relevance of geological data is less persuasive. On the other hand, geological instruments are more likely to fulfil the exogeneity condition.

Despite the previous attempts to solve the problem, the evidence has not proven conclusive in explaining the direction and magnitude of the bias that might arise from the simultaneity bias. While some studies obtain instrumental variables estimates that are smaller than the respective least squares ones, others seem to find the exact opposite result. In addition, some authors conclude that if agglomeration does have an endogenous component it does not appear to induce a substantial bias in estimates. Finally, there is lack of evidence on the extent to which endogeneity may predominate in some industries more than in others.

In addition to reverse causality, there is also a case for reverse causality between wages and home-to-work distance because the demand for travel depends on income levels. The main idea is that workers are only prepared to commute longer distances if they are compensated with higher wages. The reverse argument is also true: higher wage workers commute longer distances, other things (type of job, industry, etc.) remaining the same, because the have stronger preference for space and other amenities that tend to be located away from their workplace location.[2]

It is more difficult to find a valid instrument for home-to-work commuting distances, particularly one available at the worker level. To the best of our knowledge this has not been done in the context of estimation of agglomeration externalities on workers' productivity. We discuss the instrumental variables strategy for home-to-work distances in section 5.2.

# 4. Description of data and variables

## 4.1. Data sources

Our main data source is the Annual Survey of Hours and Earnings (ASHE), an employee level longitudinal survey conducted by the Office for National Statistics (ONS). The ASHE is mainly based on a 1% sample of National Insurance numbers (NINo) of employees on the Inland Revenue (IR) Pay-As-You-Earn (PAYE) register in February, and other supplementary samples with the purpose of improving the coverage of the survey. Additional samples include the IR PAYE register in April and the Inter Departmental Business Register that includes businesses that are not registered for PAYE but are registered for VAT. These two additional samples allow for the inclusion of employees that either entered the job market or moved jobs between the time of selection (February) and the date of the survey (April), as well as employees that earn below the PAYE limit.[3] [4]

The ASHE contains a rich set of individual information about workers. The main variables refer to workers earnings, number of hours worked, gender, age, occupation, industry, whether the worker is a part-timer or full-timer, whether earnings are affected by absence, whether the worker is paid an adult or trainee/junior rate, and various types of geographic boundaries that identify the home and workplace locations. Unfortunately, there is no direct measure of workers' education levels or

[2] The preference for space will also depend on workers' characteristics (e.g.: age, household size, etc), regardless of wage levels.

[3] The sample is based on NINo ending with a specific pair of digits generated by a 1 in 100 random sample of all jobs registered in the PAYE scheme, so each employee has an equal probability of being selected. The data coming from the VAT-only sample is obtained according to a random sample of registered businesses. The key difference to the IR PAYE scheme is that all employees in a selected business are included in the sample.

[4] The ASHE replaced the New Earnings Survey (NES) in 2004. The data pre-2004 was made consistent with the ASHE methodology by the ONS staff in order to ensure that the panel could be constructed.

qualifications. There are also some features related to the worker's employer, but the only variable consistent over time is that reporting the number of employees working in the firm.[5] One of the key advantages of micro level data is that it allows us to control for human capital attributes that determine wage levels, and which may be correlated with (work) location (e.g. years of education, occupation, and industry, etc).

The original dataset available contains 1,559,719 observations, covering the period from 1997 to 2006. After performing a series of actions concerned with data cleaning and data transformations, which we summarise in the paragraphs below, the final dataset consists of an unbalanced panel of 225,104 individuals (corresponding to 657,533 observations) over the 5 year period from 2002 to 2006. On average, we observe each individual worker 2.92 times.

The largest loss of data results from the inclusion of the home-to-work distances as one of the regressors in the wage equation. In order to calculate this measure we use the data for both home and workplace postcodes, which are only available from 2002 onwards. This essentially halves the size of our dataset. The remaining observations lost result mainly from cleaning the dataset of records for which there were either missing values or some form of misreport (e.g. negative wages, ages below 16 years, etc.). In addition to the cleaning of the dataset from erroneous and missing records, we also excluded the observations referring to employees whose earnings were affected by loss of pay due to absence and employees paid at junior/trainee rates.

Based on the Northings and Eastings of worker's respective home and work postcodes we calculate home-to-work distances. There are outlier values for home-to-work distances: about 2.2% of the home-to-work distances are greater than 200 kilometres. The mean home-to-work distance is around 21 kilometres, and the percentiles 75, 90, and 95 are 16, 38, and 86 kilometres respectively. There are various possible explanations for such long commutes and we apply a series of checks and tests to assess the extent to which errors could be affecting our data. Some of the long commutes can simply result from misreported postcodes. Alternatively, there may be unrealistic long commutes because of differences in workers' actual home address and the address in the register hold by the employer. Although many of the long commutes are likely to be some sort of error, others can actually be correct, particularly if they are associated with high wage earners who have top managerial and highly skilled professional occupations and do not commute daily, or work from home.

To investigate the potential for errors in postcode records and hence unrealistically high commute lengths, we examine workers' wage levels, occupation groups and industries simultaneously. We do not expect to find workers willing to commute long distances unless they are compensated with high wages. Inspection of the occupation group can also provide an indication of whether workers may only need to commute weekly and hence be able to live further away from their workplaces. After

[5] The questionnaires are completed by the employers, presumably reducing the incidence of both misreported and underreported records.

inspection of the data we applied two thresholds to obtain what we think is a more representative and clean dataset. We dropped the observations for workers with commutes lengthier than 120 kilometres, earn wage rates below the percentile 90, and work on occupation groups that do not reflect skilled jobs. Secondly, we also dropped the records for workers with commutes longer than 100 kilometres who earn less than the average wage rate.

The average home-to-work distance is 15 km and the percentiles 75, 90, and 95 are 14, 30, and 51 kilometres respectively. These Figures compare well with other data for commuting distances in Great Britain. Benito and Oswald (2000) use the British Household Panel Survey (BHPS) and show that the average commuting distance of full-time workers in Great Britain is 13 kilometres. Dent and Bond (2008) use data from the Census 2001 and show that the average commuting distance of full-time workers in Great Britain is also about 13 kilometres.

## 4.2. Measuring agglomeration externalities

To build measures for agglomeration economies we use data on yearly sectoral employment available from the Annual Business Inquiry (ABI) from 2002 to 2006. We use Travel-to-Work Areas (TTWAs) as the spatial unit of analysis because these are the best available approximations of self-contained labour markets; they are defined as regions where the proportion of people who live (work) in the area is at least 75% of the total number of people who work (live) in the area.[6] Similar geographic definitions can be found in the literature measuring the productivity benefits from agglomeration, including the French employment areas (e.g. Combes et al., 2008a; Combes et al., 2007; Combes et al., 2008b) and the Italian local labour markets (e.g. Blasio and Di Addario, 2005; Di Addario and Patacchini, 2007), both defined on the basis of daily commuting patterns. We base the analysis at the level of work TTWA instead of home TTWA on the grounds that this spatial unit is more relevant for the analysis of productivity gains from agglomeration economies.

The use of a territorial division that approximates labour markets can offer some important advantages. In particular, it can reduce the concerns about the endogeneity of agglomeration effects as a result of the low labour mobility between labour markets (Di Addario and Patacchini, 2007). In addition, labour market areas are well suited to capture the geographic scope within which agglomeration externalities take place, in particular those related to labour market pooling and knowledge spillovers. On the other hand, these boundaries are less appropriated to capture the spatial reach of the agglomeration effects arising from intermediate input-output linkages between firms which presumably take place between various labour market areas.

[6] We use TTWA as at 1998. For the list of names and codes see: http://www.statistics.gov.uk/geography/geographic_area_listings/other.asp.

Urban agglomeration economies are a function of the concentration of a diversity of activities, closely correlated with the size of a given area. These externalities derive from access to specialised inputs among diverse firms such as specialised business services, access of firms to public facilities, transport systems and other public infrastructure. To represent urban agglomeration externalities we therefore use a measure of the employment density of a given TTWA. Density measures have been widely used in the literature as a surrogate for external urban economies (e.g. Ciccone and Hall, 1996; Ciccone, 2002; Combes et al., 2008a; Combes et al., 2008b; Combes et al., 2007; Mion and Naticchioni, 2005). Employment density is simply defined as the ratio of total employment to the total land area:

$$dens_{rt} = \frac{emp_{rt}}{area_r}, \tag{8}$$

where $emp_{rt}$ is employment in TTWA area $r$ at time $t$ and $area_r$ is the respective land area in square kilometres.

To account for between-labour market interactions we also include a measure of the market potential accessible at a given TTWA $r$, as first introduced by Harris (1954). We follow Combes et al, 2008a,b and Mion and Naticchioni for this approach, who use a market potential function to capture product market interactions such as those related to input-output linkages. The market potential function can proxy for the potential demand for goods and services produced in a given location $r$ from all other locations $j$, taking into account the proximity of each location $j$ to $r$. We measure market potential as below.

$$MP_{rt} = \sum_{j}^{r \neq j} emp_{jt} d_{rj}^{-1}, \tag{9}$$

which is a distance discounted sum of employment of all other locations $j$ where $d_{rj}$ is the distance between the two locations, measured using the Pythagoras theorem and the *x-latitude* and *y-longitude* coordinates of each TTWA centroid.

To account for the importance from other characteristics of labour markets we include a set of controls commonly used in the agglomeration literature, which capture the extent to which labour markets are specialised in a one or some industry sectors and the extent to which they exhibit a diversified economic structure. To evaluate economic diversity we use an index that measures the similarity between the composition of economic activities of a given local economy to that of the national economy. The index of relative economic diversity is the inverse of the summed absolute differences between the local and national share of employment in economic sector *s*. We use the formula proposed by Duranton and Puga (2000).

$$div_{rt} = \left[ \sum_{s=1}^{S} \left| \frac{emp_{rst}}{emp_{rt}} - \frac{emp_{st}}{emp_{t}} \right| \right]^{-1} , \tag{10}$$

where $emp_{rst}$ is the employment in industry sector *s*, TTWA *r* at time *t*; $emp_{st}$ is the national employment in industry sector *s* at time *t*; and $emp_t$ is the total national employment at time *t*. The greater the value of the index the more similar is the composition of economic activities in the local and the national economies; the smaller the index the less diversified is the local economy relative to the whole economy.

We also consider measures that capture within-industry interactions. A simple measure of relative specialisation is that proposed by Duranton and Puga (2000), shown below:

$$spec_{rt} = \underset{s}{Max} \left[ \frac{emp_{rst}}{emp_{rt}} \Big/ \frac{emp_{st}}{emp_{t}} \right], \tag{11}$$

where the subscripts and variables are as before. In alternative, industrial specialisation can be measured with the Hirschmann-Herfindahl index (HHI), defined below.

$$HHI_{rt} = \sum_{s=1}^{S} \left( \frac{emp_{rst}}{emp_{rt}} \right)^{2} , \tag{12}$$

# 5. Results and discussion

In this section we present and discuss the results from the regression analysis of the wage equation below:

$$\ln w_{it} = \alpha_0 + X_{it}\beta + Z_{r(i)t}\alpha + \varphi \ln Y_{f(i,t)} + \delta_t + \lambda_{o(i,t)} + \sigma_{s(i,t)} + \omega_{r(i,t)} + \eta_i + \psi_{f(i,t)} + \varepsilon_{it} , \tag{13}$$

where the dependent variable is the logarithm of real net hourly earnings, $X_{it}$ is a vector of individual worker characteristics, $Y_{f(i,t)}$ measures the number of workers in the respective firm, and $Z_{r(i)t}$ includes the variables that measure agglomeration economies. The remaining terms are dummy variables that control for various sources of heterogeneity; we explain the rationale for their inclusion in the model specification below.

There are various sources of heterogeneity that if omitted can lead to inconsistency. Perhaps the most important source of consistency is workers' unobserved heterogeneity, which essentially captures differences in workers' ability and education: $\eta_i$. The inclusion of individual-specific effects is crucial to account for workers' unobservable attributes that can influence wage levels. The consistency of the model parameter estimates depends to large extent on the assumptions made about the behaviour of these effects and the econometric estimators employed.

Besides (time-invariant) worker heterogeneity, workers' wage levels may also be affected by firm characteristics such as the wage policy and other unmeasured attributes of firms. It has been argued elsewhere that firm size can capture, at least to some extent, firm heterogeneity in productivity (see Mion and Naticchioni, 2009). Nevertheless, there will be many other factors across which firms can differ regardless of their size, so we also include a set of dummy variables for each firm in the wage equation: $\Psi_{f(i,t)}$. The firm-specific effects are identified from workers' mobility across firms and capture the portion of wages that is specific to firm *f*. In this model, the identification of worker and firm fixed effects relies on there being at least one worker with at least one job change for each firm.[7]

Since the ASHE does not contain any direct measure of workers' educational qualifications we can not identify the effect on wage from individuals' skills acquired through education. Nevertheless, the ASHE collects data for occupational groups (SOC) up to the 4 digit level, which covers over 350 unit groups. This is very useful because it can provide a meaningful and detailed characterisation of the nature of the job performed by a given worker. Moreover, and in contrast to educational levels, the SOC accounts for competences acquired through non-school qualification, training and work experience. This means that the SOC can still offer a meaningful way of representing workers' skills. To account for wages differentials related to heterogeneity in workers' skills as represented by their respective occupation, we include dummy variables for the SOC at the one digit level: $\lambda_{o(i,t)}$,.

Similarly, there may be differences in wage levels of equivalent individuals that work in different industries and we wish to net that effect out of the effects from our variables of interest. Sectoral heterogeneity is accounted for through inclusion of a set of dummy variables: $\sigma_{s(i,t)}$. To account for time-specific events that are shared among all workers, we include a set of year dummies: $\delta_t$.

Also of great importance is the fact that inequality of wage levels can reflect spatial differences in the cost of living (including land/housing prices), the degree of market competition, etc. In particular, living cost tends to exhibit the same kind of positive relationship with agglomeration as wages: more agglomerated areas face higher living costs. In fact, one of the theories put forward to explain spatial wage differentials is that spatial wage differentials are only nominal and exist to

[7] We fit a linear fixed-effects model with worker- and firm-specific effects using the memory saving programme felsdvreg developed by Cornelissen, (2008).

compensate for higher costs of living. We do not have data for either the cost of living or house prices at the level of the TTWAs, and therefore add dummy variables to account for such differences. This strategy, however, raises some issues and finally we include dummy variables for the Governmental Office Regions (GORs) instead: $\omega_{r(i,t)}$. There are two main reasons supporting the option for GOR level controls. First, there were identification issues in the estimation of the panel data estimators because of multicollinearity between the individual effects and the TTWA dummies. Secondly, for many TTWAs inside the same GOR there may actually be no real differences in the average cost of living except for the TTWA that corresponds to the core region of the GOR. We acknowledge that this option may have some limitations but point out that it does not compromise the consistency of the estimates obtained. Finally, $\varepsilon_{it}$ is the error term and is assumed to be normally distributed while allowing for potentially heteroskedasticity and clustering on individuals.

### *Econometric estimators*

We compare the results obtained from various estimators that make very distinct assumptions about the correlation between the regressors and both the error term and the individual-specific effects. The estimators we use are the pooled OLS (POLS), the between estimator (BE), the random-effects estimator (RE), the fixed-effects estimator (FE), the first-differences estimator (FD), the Hausman and Taylor (1981) estimator (HT), and the instrumental-variables estimator (IV). Appendix B provides a detailed explanation of the statistical estimators used.

The main distinction between the estimators above concerns their ability to produce consistent and unbiased model parameter estimates. In the presence of nonzero correlation between the individual specific effects (and hence the error term) and (all) the covariates, only the FE and the FD estimators ensure consistency. On the contrary, if the individual-specific effects are uncorrelated with the other regressors the pooled OLS and the BE are consistent estimators but are less efficient than the RE estimator. While the FE estimator is also consistent in this case, the RE estimator is more efficient because it uses both “within-panel” and “between-panel” variation. Furthermore, the FE and the FD estimators will not be able to identify the coefficients of time-invariant variables (e.g. gender, education).

To assess the appropriateness of the RE or the FE model we perform the Hausman (1978) test of the null hypothesis of no correlation between the individual effects and the regressors; if the test fails to reject the null hypothesis the RE model should be preferred on the basis that it produces both consistent and efficient estimates of the parameters of the model, otherwise the FE should be preferred because it is the only estimator that produces consistent estimates. In addition, we also estimate the HT (1981), which, unlike the RE model, produces consistent estimates and, unlike the FE and the FD estimators, identifies the coefficients of time-invariant covariates. Finally, in the presence of reverse causality, some form of IV estimator needs to be applied to ensure consistency.

In section 5.1 we estimate the main baseline results for the effects from agglomeration externalities and investigate the degree to which workers' and firms' unobservable attributes affect the parameter estimates for agglomeration externalities. In addition, we undertake further regression analysis to test for the presence of spatial sorting of workers' skills across space. Section 5.2 corrects for endogeneity from reverse causality between productivity and agglomeration economies through instrumental variables estimation. In section 5.3 we examine the spatial decay pattern of agglomeration externalities.

## 5.1. Bias I: workers' self-selection

Table 2 reports the results for the whole economy, whereas Table 3 reports the results obtained for each of the 13 industry groups of two-digit industries. See Table A1 in Appendix A for a description of the economic sectors.

The size of the coefficients differs significantly across estimators. Wages increase with worker's age but the increases become smaller as workers become older. The marginal effect of age appears to decline with age, suggesting a concave age-earning relationship. This nonlinear relationship reflects some sort of diminishing returns of wage rates with respect to age. This can be because at early stages of their careers workers have lower wages and hence experience relatively larger increases than at later stages of the career where wage levels are higher. Alternatively, it can also reflect the fact that as individuals get older they learn new techniques more slowly and are less able to readjust to new tasks, resulting in a decline of their productivity levels. Female workers receive on average between -10% and -20% less than male workers. Full-timers only earn higher hourly wages according to the POLS and the BE estimators, while the panel regressions suggest that full-timers can earn hourly wages between -6% to -0.7% lower than part-timers.

The elasticity of wages with respect to firm size shows that doubling the number of employees raises wages by around 1.0%; this positive relation between firm size and workers' earnings can also suggest the presence of internal returns to scale. There is also some theoretical reasoning suggestive of a positive association between firm productivity and firm size (e.g. Kim, 1989; Helsley and Strange, 1990).

The trade-off between longer commute lengths and higher wages is reflected by the coefficient of home-to-work distance, which measures the (wage) compensation of commuting. The estimates indicate that increasing home-to-work distances by 10% requires compensation in hourly wages between about 0.14% and 0.61%, depending on the estimator. Taking the coefficient value from the

FE estimator as reference, and knowing that the average commute length (hourly wage rate) is about 15 kilometres (9 pounds), this means that an additional 150 metres requires a compensation of about 13 pence (one pound) in hourly (daily) wage rates. One additional kilometre requires compensation in hourly earnings of approximately 84 pence.[8]

As regards the effect from agglomeration externalities, the results indicate that the magnitude of the effects of labour market employment density and market potential on wages is sensitive to the estimator used. Controlling for unobserved worker heterogeneity while assuming that there is no correlation between individual-specific effects and other regressors, reduces the size of the elasticity of hourly wages with respect to employment density (market potential) from 2.2% (10.0%) to 2.0% (9.9%). Once we relax the assumption of no correlation between the individual effects and the regressors, the elasticity values further reduce to 1.0% (5.4%) respectively. The Hausman test rejects the null hypothesis of no correlation between the individual effects and the regressors, implying that the RE estimator provides inconsistent estimates and so we should prefer the FE estimator.

Like the FE, the FD estimator also ensures consistency because it eliminates the individual specific effects through the first-differences transformation. The results suggest that an increase of one percentage point in the annual growth of employment density (market potential) increases worker's hourly wage growth by about 0.37 (3.63) percentage points.

The FE estimator assumes that all covariates are correlated with the individual effects, and cannot estimate the coefficients of time-invariant regressors. The HT estimator can retrieve the coefficients of time-invariant covariates and also allows for nonzero correlation between the individual effects and some of the covariates. We fit the HT estimator allowing for nonzero correlation between the individual effects and three covariates: home-to-work distance, the proxy for experience (age squared), employment density, and market potential. The elasticities of hourly wages with respect to employment density and market potential are 1.6% and 10.1% respectively. Nevertheless, the Hausman test between the FE and the HT estimates still rejects the null hypothesis and we suspect this indicates that here may be other endogeneity issues in our model, namely reverse causality between worker productivity and agglomeration.

So far, the most robust coefficients are those obtained from the FE estimator. Besides depending on worker characteristics that affect their productivity (education, ability), wage can also be determined by firm characteristics, including firms' wage policies. Since our dataset has information for the enterprise in which workers work, we use the enterprise reference to control for firms' unobservable heterogeneity which should capture for firm (time-invariant) productivity differentials.

[8] Evidence for the Netherlands by Rouwendal, (1999) shows a smaller compensation level: on average workers are willing to accept an increase of 1 kilometre in commuting distances if there is an increase of 12 (96) cents of a Dutch guilder in hourly (daily) wage rates. Converting to pounds (NLG 1= £ 0.39), this means that on average workers are willing to accept commuting one kilometre more if their hourly (daily) wages increase by 5 (37) pence.

There are 5,814 firm effects identified and 225,103 worker fixed-effects identified in the model with worker- and firm-specific fixed-effects. Observation of the parameter estimates of model FE (2) reveals that further controlling for time-invariant firm level heterogeneity, besides unobserved time-invariant worker heterogeneity, appears to impact on the parameter coefficients only slightly. The wage elasticity of employment density (market potential) is 0.72% (5.1%) compared to 1.0% (5.4%) obtained from the model FE (1), which controls only for unobserved time-invariant worker heterogeneity.

Beneath the aggregate effect, there is substantial variation by industry group, as shown in Table 3. The results presented in the table are for the worker fixed-effects estimation, which according to the Hausman tests displayed at the bottom is the only consistent estimator. The effects are positive and significant for the renting, information technology, research and development industries, and other business activities. The elasticity values for employment density for these industry groups are 1.73% and 1.42% respectively. This is consistent with the theory that it is the knowledge and skill intensive activities that benefit the most from clustering in denser areas, where the benefits from face-to-face communication, access to a skilled workforce, and access to firms that provide complementary specialized services are maximised. Nevertheless, it is surprising to find that the effect from labour market density is not significant for the financial intermediation sector. One possible explanation for the lack of significance may be that this sector is composed of a variety of services (from commercial banks to investment banks and insurance) that do now all require being located in dense labour markets. The effect is not statistically significant for the remaining services industries, primary industries, manufacturing, electricity, gas and water, and construction.

The importance of access to product markets also differs across industry groups and exhibits a pattern distinct from the one outlined for labour market density. The market potential captures the importance of proximity to intermediate input suppliers and output costumers. Traditionally, proximity to intermediate suppliers of inputs has been of greater importance for manufacturing industries, for example the textiles industry. The effects from market potential are positive and significant for manufacturing (5.6%), electricity, gas and water (11%), and transport, storage and communication (7.8%).

If we consider the analysis at the more disaggregated level of two-digit SIC industries separately, we observe that increasing returns to labour market size are statistically important for a few service industries, in particular the research & development industry (SIC 73) and other business activities (SIC 74). Both the research & development and the business activities industries are strongly dependent on the availability of skilled workers, who are predominantly located in denser and larger labour markets. On the other hand, there is a negative association between labour market density and wages for the industries related to the extraction of crude oil and natural gas (SIC 11) and the manufacture of coke, refined petroleum products and nuclear fuels (SIC 23). These results are reasonable because the location of firms in these sectors, particularly the extraction of crude oil and natural gas, is vey much determined by the location of natural resources (crude oil, natural gas, etc.).

As for the role of access to product markets, the effects are mixed across industries. The effect tends to be positive for manufacturing, particularly for the manufacture of pulp, paper and paper products (SIC 21), the publishing, printing and reproduction of recorded media (SIC 22), the manufacture of coke, refined petroleum products and nuclear fuels (SIC 23), the manufacture of machinery, motor vehicles and other transport equipment (SIC 29, SIC 34, and SIC 35), and also the manufacture of furniture (SIC 36). Access to product markets (either intermediate and/or final) also appears to be important for the electricity, gas and water supply industry (SIC 40) and some service industries: retail trade, repair of personal and household goods (SIC 52), land transport and transport via pipelines (SIC 60), and public administration and defence, and compulsory social security (SIC 75). This suggests that access to a large final demand market may be important to some of the service sectors like retail trade and land transport, but also for some manufacturing sectors like the manufacture of machinery, motor vehicles and other transport equipment and publishing, printing and reproduction of recorded media.

On the contrary, market potential has a negative importance for the research & development industry (SIC 73), which is reasonable given that access to product markets is not the key factor affecting this industry's economic performance. Perhaps less expected is the negative coefficient for the importance of market potential for the agriculture, hunting and related service activities (SIC 01), the manufacture of wearing apparel, dressing and dyeing of fur, leather, and manufacture of handbags, saddlery, harness and footwear (SIC 18 and SIC 19). Despite the perishable nature of many of its products, the general reduction in transport costs may have made it less important for agriculture to be close to large final demand markets. As for the manufacturing industries SIC 18 and SIC 19, it may be that access to cheap labour dominates over access to final product markets.

Our baseline results for the effect of labour market scale and market potential indicate that doubling the employment density of a given TTWA can raise workers' hourly wages by no more than 1%. Halving the distances to other markets increases produces an increase of hourly wages of 2.7%. One important aspect of the estimation so far is that accounting for workers' unobservable heterogeneity appears to be important in obtaining a correct estimate for the effects economic scale and market potential, whereas accounting for firms' unobservable characteristics produces only slight changes. Worker fixed-effects account for nearly 75% of all variation in wages compared to the less that 2% of firm-effects. Controlling for workers' unobservable heterogeneity produces a reduction of about 55% in the size of the effects from labour market scale (from 2.2% to 1.0%) and about 46% in the magnitude of the effect from market potential (from 10.0% to 5.4%). Further controlling for firms unobservable characteristics produces wage elasticities of employment density and market potential of 0.72% and 5.1% respectively.

**Table 2: Regression results for the whole economy**

| Dependent variable: log of real hourly wage | POLS | BE | RE | FE (1) | FE (2) | HT | FD |
|---|---|---|---|---|---|---|---|
| Log of home-to-work distance | 0.0576** | 0.0609** | 0.0334** | 0.0143** | 0.0134*** | 0.0185** | 0.0090** |
| Age | 0.0387** | 0.0348** | 0.0475** | 0.0429** | 0.0418*** | 0.0694** | 0.0475** |
| Age2 | -0.0004** | -0.0004** | -0.0005** | -0.0006** | -0.0006*** | -0.0008** | -0.0006** |
| Female | -0.1198** | -0.0996** | -0.1668** | | | -0.1960** | |
| Full-timer | 0.0737** | 0.0862** | -0.0071** | -0.0633** | -0.0676*** | -0.0350** | -0.1017** |
| Log of firm size | 0.0139** | 0.0147** | 0.0100** | 0.0061** | 0.0053*** | 0.0083** | 0.0048** |
| Log of employment density | 0.0222** | 0.0219** | 0.0195** | 0.0096** | 0.0072*** | 0.0161** | 0.0037* |
| Log of market potential | 0.0999** | 0.0914** | 0.0990** | 0.0538** | 0.0519*** | 0.1072** | 0.0363** |
| Log of index of economic diversity | -0.0515** | -0.0568** | -0.0313** | -0.0092** | -0.0082*** | -0.0205** | 0.0058* |
| Log of index of industrial specialisation | 0.0012 | -0.0008 | -0.0010 | -0.0064 | -0.0042 | -0.0053 | 0.0088 |
| Log of area | 0.0410** | 0.0404** | 0.0351** | 0.0147** | 0.0131*** | 0.0324** | -0.0034 |
| Observations | 657533 | 657533 | 657533 | 657533 | 655924 | 657533 | 384889 |
| F-statistic | 7659.4920 | 9041.4471 | | 349.7688 | | 3965.1271 | 232.9263 |
| Chi2 | | | 1.91e+05 | | | 1.78e+05 | |
| R2 overall | 0.5989 | 0.5960 | 0.5855 | 0.3902 | | | |
| R2 within | | 0.0349 | 0.0633 | 0.0792 | 0.1195 | | 0.0515 |
| R2 between | | 0.6439 | 0.6219 | 0.4170 | | | |
| Test of individual effects:p-value | | | ** | | *** | | |
| Hausman (re vs fe) | | | | 66796.81 | | | |
| p-value | | | | ** | | | |
| Hausman (fe vs ht) | | | | 28977.54 | | | |
| p-value | | | | ** | | | |
| Time dummies | yes | yes | yes | yes | yes | yes | yes |
| Occupation dummies | yes | yes | yes | yes | yes | yes | yes |
| Industry dummies | yes | yes | yes | yes | yes | yes | yes |
| Region dummies | yes | yes | yes | yes | yes | yes | yes |
| Worker dummies | no | no | no | yes | yes | no | no |
| Firm dummies | no | no | no | no | yes | no | no |

Standard errors are cluster-robust. Clustering is on workers. Significance levels: + p<0.10, * p<0.05, ** p<0.01.

**Table 3: Regression results for individual industry groups**

| Dependent variable: log of real hourly wage | Primary | Manufacturing | Electricity, gas & water | Construction | Wholesale & retail | Hotels & restaurants |
|---|---|---|---|---|---|---|
| Log of home-to-work distance | 0.0086 | 0.0087** | 0.0075+ | 0.0082* | 0.0072** | 0.0029 |
| Age | 0.0180 | 0.0351** | 0.0778** | 0.0552** | 0.0350** | 0.0545** |
| Age2 | -0.0004** | -0.0005** | -0.0008** | -0.0006** | -0.0006** | -0.0004** |
| Full-timer | -0.1778** | -0.1292** | -0.1351* | -0.1917** | -0.0437** | -0.0632** |
| Log of firm size | -0.0165 | 0.0209** | -0.0047 | 0.0088* | 0.0020+ | -0.0012 |
| Log of employment density | 0.0214 | 0.0045 | -0.0010 | 0.0072 | 0.0007 | 0.0010 |
| Log of market potential | 0.0707 | 0.0561* | 0.1139* | 0.0207 | 0.0427 | 0.0623 |
| Log of index of economic diversity | -0.0506 | 0.0025 | 0.0304+ | 0.0139 | 0.0073 | 0.0089 |
| Log of index of industrial specialisation | -0.0157 | -0.0064 | -0.0236 | -0.0409 | 0.0343+ | 0.0468 |
| Log of area | -0.0347 | -0.0188** | -0.0479+ | -0.0212 | -0.0165+ | 0.0155 |
| Observations | 5206 | 96527 | 3288 | 23560 | 100576 | 21116 |
| R2 overall | 0.0202 | 0.1906 | 0.2158 | 0.1694 | 0.0707 | 0.1088 |
| R2 within | 0.0758 | 0.0507 | 0.1141 | 0.0847 | 0.0399 | 0.0574 |
| R2 between | 0.0398 | 0.1858 | 0.2404 | 0.1885 | 0.0588 | 0.1231 |
| Hausman (FE vs RE) | 549** | 10411** | 461** | 1460** | 10092** | 1216** |
| Hausman (FE vs HT) | 228** | 4584** | 175** | 860** | 3807** | 449** |

| Dependent variable: log of real hourly wage | Transport, storage & communication | Financial intermediation | Real estate | Renting, IT,R&D | Other business activities | Public services |
|---|---|---|---|---|---|---|
| Log of home-to-work distance | 0.0165** | 0.0122** | 0.0135+ | 0.0150** | 0.0096** | 0.0069** |
| Age | 0.0076 | 0.0753** | 0.0513** | 0.0678** | 0.0808** | 0.0396** |
| Age2 | -0.0002** | -0.0009** | -0.0005** | -0.0007** | -0.0008** | -0.0005** |
| Full-timer | -0.1213** | -0.0430** | -0.0847** | -0.1282** | -0.0822** | -0.1086** |
| Log of firm size | 0.0102** | 0.0052** | 0.0088+ | 0.0194** | 0.0045+ | 0.0036** |
| Log of employment density | 0.0018 | 0.0081 | 0.0292 | 0.0173+ | 0.0142* | 0.0006 |
| Log of market potential | 0.0780* | 0.0170 | -0.0750 | 0.0151 | 0.0647 | 0.0025 |
| Log of index of economic diversity | 0.0032 | 0.0160+ | 0.0044 | -0.0160 | 0.0200+ | 0.0044 |
| Log of index of industrial specialisation | 0.0489+ | 0.0708* | -0.0248 | -0.0427 | -0.0195 | -0.0122 |
| Log of area | 0.0036 | 0.0021 | -0.0364 | -0.0312 | -0.0021 | 0.0107* |
| Observations | 42072 | 36659 | 8526 | 19191 | 54216 | 223085 |
| R2 overall | 0.1666 | 0.3283 | 0.0267 | 0.2582 | 0.2104 | 0.3197 |
| R2 within | 0.0490 | 0.0643 | 0.0846 | 0.0472 | 0.0738 | 0.0939 |
| R2 between | 0.1411 | 0.3672 | 0.0330 | 0.2771 | 0.2148 | 0.3243 |
| Hausman (FE vs RE) | 4460** | 5134** | 695** | 1608** | 5315** | 22214** |
| Hausman (FE vs HT) | 1848** | 2017** | 311** | 870** | 2639** | 10553** |
| Time dummies | yes | yes | yes | yes | yes | yes |
| Occupation dummies | yes | yes | yes | yes | yes | yes |
| Region dummies | yes | yes | yes | yes | yes | yes |
| Worker dummies | yes | yes | yes | yes | yes | yes |
| Firm dummies | no | no | no | no | no | no |

Standard errors are cluster-robust. Clustering is on workers. Significance levels: + p<0.10, * p<0.05, ** p<0.01.

### *Workers' spatial sorting*

To investigate the extent to which workers' unobservable skills sort across space we compute the average of worker fixed-effects for (i) high-density (HD) and low-density (LD) labour markets[9], and (ii) urban and rural areas[10]. These averages measure percentage wage differences from the overall mean of the worker fixed-effects, which equals zero. If there is spatial sorting of workers' skills towards urban and more dense labour markets we should find a positive value for the average of workers' unobservable skills in HD labour markets and urban areas, against a negative value for the average of workers' unobservable skills in LD labour markets and urban areas respectively.

The average of worker fixed-effects in HD labour markets is 0.0139, compared to the -0.0515 for workers whose workplace is in LD labour markets. Looking at the urban versus rural workplace location, the averages of workers' unobservable skills are 0.0062 and -0.0383 respectively. This indicates that more able individuals work in HD labour markets and urban areas, whereas less able individuals work in LD labour markets and rural areas.

To further inspect the spatial sorting of workers' unobservable skills we calculate the average skill for workers moving workplace area. The averages of workers' unobservable skills for workers moving workplace between LD/rural and HD/urban labour markets/areas are shown in Table 5. We observe that the average is negative (positive) for workers moving workplace within low-density (high-density) labour markets: the values are-0.025 and 0.048 respectively. The average is positive and equal to 0.065 for movers from LD to HD labour markets. This is indicative of a positive spatial sorting of workers' unobservable skills. However, we observe that the average skill of movers from HD to LD labour markets is 0.073. This runs in contrary to the spatial sorting argument, but can result from the fact that workers can move workplace from a rural area in a HD labour market to an urban area in a LD labour market. We therefore calculate the average skill for workers moving between urban and rural workplace areas. Table 4 shows the average skill of movers between rural areas (from urban-to-rural areas) is negative, whereas the average skill of movers between urban areas (from rural-to-urban areas) is positive.

**Table 4: Average of workers' unobservable skills by workplace area**

| | Low-density | High-density | | Rural | Urban |
|---|---|---|---|---|---|
| Low-density | -0.025 | 0.065 | Rural | -0.0062 | 0.0044 |
| High-density | 0.073 | 0.048 | Urban | -0.0023 | 0.0403 |

[9] HD and LD are defined by reference to the mean labour market employment density over the period from 2002 to 2006. Labour markets with density values greater (lower) that the mean value are classified as HD (LD).

[10] Workplace postcodes are classified as urban or rural according to the National Statistics Rural and Urban Classification of Output Areas. If a given postcode is located in an urban Output Area (OAs), it will be classified as urban, and vice-versa. Output Areas are designated as urban or rural depending on their size. For England and Wales, OAs with 10,000 or more people are designated as urban. In Scotland, OAs with 3,000 or more people are classified as urban.

## 5.2. Bias II: endogeneity of agglomeration externalities

So far we have considered the bias from endogeneity resulting from nonzero correlation between the model covariates and the individual-specific effects, in particular due to workers' spatial self-selection into denser areas. The estimation of agglomeration externalities is also potentially troubled with endogeneity bias from reverse causality between wages and agglomeration, as well as reverse causality between wages and home-to-work distances.

To correct for the simultaneity between agglomeration and workers' wages we implement instrumental variables (IV) techniques, using long-lagged values of the endogenous variables as instruments. Such instruments are widely used in the literature and were first employed by Ciccone and Hall (1996). The case for instrument exogeneity, as argued by Ciccone and Hall (1996) and Combes et al. (2008a,b), consists on the idea that deep time lags of urban densities are correlated with current levels of urban densities, but uncorrelated with current levels of productivity. In other words, the urban system in the nineteenth century can explain to some extent the distribution of present urban densities, but does not explain the distribution of current productivity levels.

It is more difficult to find a valid instrument for home-to-work commuting distances, particularly one available at the worker level. The approach we adopt is to instrument home-to-work distance of a given worker *i* with the average commuting distance (excluding home-to-work distance of worker *i*) of all individuals working is same workplace area as worker *i,* regardless of where they live. The average commuting distance to a given workplace can reflect at least partially the level of accessibility to a given area, which itself can determine commuting patterns to the area (e.g. well connected areas should have people commuting from longer distances because accessibility is better). On the other hand, the average commute length of other individuals working in same area should not determine the wage level of worker *i*, other than for the fact that workers may work in the same industry, have similar occupations and skills, etc.

In what follows, we discuss the validity of our instruments and their ability to provide unbiased instrumental variables estimates.[11] There are two conditions that need to be satisfied in order for the instruments to be valid. The first condition demands that the instruments be uncorrelated with the error term: $E[\varepsilon|Z] = 0$, where Z contains the instruments, $X_{end}$ contains the endogenous regressor(s) in the model, and $\varepsilon$ is the error term. The second condition requires that the instruments and the endogenous variable be correlated with each other: $E[X_{end}|\ Z] \neq 0$. The stronger the correlation, the better the relevance of the instruments and the less biased the parameter estimates will be. The set of instruments

[11] We use Stata's unofficial command xtivreg2, provided by Schaffer, (2007).

chosen will only be valid when these two conditions are met (see Bound et al., 1995; Staiger and Stock, 1997; Murray, 2006).

Table 5 presents the results for the whole economy obtained from the IV regressions, while Table 6 reports estimates for each of the 13 economic sectors. The difference between model (1) and model (2) is that model (1) assumes that home-to-work distance is exogenous, while model (2) treats home-to-work distance as endogenous together with employment density and market potential. Where possible, we use the two-step feasible efficient generalised method of moments (2SFEGMM) estimation because it offers efficiency advantages compared to the traditional two-stage least squares (2SLS), due to the use of an optimal weighting scheme in the GMM-IV objective function.

### *Instrument exogeneity*

We complement the discussion of the arguments as to why the instruments we use are good with evidence from instrument diagnostic tests for instrument exogeneity. We follow Ciccone and Hall (1996), Mion and Naticchioni (2005), and Combes et al. (2008a,b) and use long lags of the agglomeration measures to correct for the endogeneity bias from reverse causality between productivity and agglomeration. We use data for population as gathered in the Census for 1801, 1831 and 1851. Our instruments cover a time period post industrial revolution that is already characterised by a strong urban presence in the form of industrial towns. On the other hand, the fact that our instruments are older than many of the technological innovations (e.g. the advent of the information technology and communications "revolution") and other critical events (e.g. World Wars I and II) that impacted on the structure of present economies should, in principle, allow for the presence of structural breaks that are important for instrument exogeneity to hold. These claims for exogeneity, however, can prove insufficient and require formal testing of instrument exogeneity.

Unfortunately, there are no data for population or employment in the nineteenth century Census at the level of the TTWAs. To obtain long-lagged values of the agglomeration measure for the TTWAs, we apportioned the population data in each Registration County (RC) to each TTWA, using as weights the share of the total land area of a given TTWA inside a given Registration County.[12]

$$POP_j = \sum_c POP_c * \frac{area_{jc}}{area_c}, \tag{14}$$

[12] The historic boundary shape files for the English and Welsh Registration Counties and Scottish Administrative Counties are provided by the Great Britain Historic GIS Project (Portsmouth University) and are available online from the EDINA UKBORDERS website: http://edina.ac.uk/ukborders/.

where $j$ and $c$ identify the TTWA and Registration County respectively. After obtaining population values for each TTWA we compute the long-lagged values for the two market potential measures.

We inspect the Hansen's J (1982)/Sargan's (1958) tests for the whole set of instruments (Census 1801, 1831, and 1851) and the difference-in-Hansen J/Sargan tests for sub-sets of instruments, to test the null hypothesis of exogeneity of the long-lagged instruments. The results for instrument exogeneity for the models pooling all industry groups are in agreement with previous studies using similar instruments, which generally fail to reject the null hypothesis of instrument exogeneity. The null hypothesis of instrument exogeneity for the sectoral models (Table 6) is also generally not rejected, with exception for the primary sector, manufacturing, transport, storage and communication, and renting, IT, and R&D industries.

***Instrument relevance***

The first-stage regressions indicate that the instruments for employment density and market potential have considerable explanatory power, whereas the instrument for home-to-work distance is rather poor. The Shea (1997) partial R-squared (simple partial R-squared) scores a value of 0.30 (0.31) for the first stage regression of employment density for both model (1) and model (2). As for market potential, the Shea partial R-squared (simple partial R-squared) for model (1) is 0.75 (0.77) and for model (2) it is 0.67 (0.77). On the contrary, the Shea partial R-squared and simple partial R-squared of the first-stage regression of home-to-work distance scores only a small 0.02 value. The implications of a weak instrument, even when the instruments are exogenous, are considerably large losses of efficiency. To assess the possible implications, we compare the coefficients obtained from the two IV estimations, model (1) and model (2) respectively.

With regard to the industry level estimations, the Shea partial R-squared (simple partial R-squared) measure of instrumental relevance ranges between 0.20 and 0.49 (0.21 and 0.90) for the first stage regression of employment density, and between 0.25 and 0.82 (0.73 and 0.89) for the first stage regression of market potential. The Shea partial R-squared (simple partial R-squared) for home-to-work distance ranges between 0.01 and 0.08 (0.01 and 0.23).

To further inspect the relevance of the instruments we perform tests of under identification that test whether the model is identified, where identification requires that the excluded instruments be correlated with the endogenous regressors. If the instruments are uncorrelated with the endogenous regressor, the matrix of reduced form coefficients is not of full rank and the model will be unidentified (see Baum et al., 2007). The usual tests are the Anderson (1951) Lagrange Multiplier (LM) and the Cragg and Donald (1993) Wald statistics. However, these statistics will not be valid when the independent and identically distributed (i.i.d.) error structure assumption is relaxed to allow for heteroskedasticity and intra-group correlation. In this case, the relevant tests are the LM and Wald Kleibergen and Paap (2006) rank statistics. The null hypothesis is that the matrix of reduced form coefficients is under identified and a failure to reject the null hypothesis means that the instrumental variables bias of the parameter estimates will be increased. The values presented in Table 5 and Table 6 show that both tests reject the null of under-identification, implying that the instruments are relevant.

In addition, we test for the presence of weak instruments. Weak identification is present when the excluded instruments are correlated with the endogenous regressor, but only poorly. In this case the estimators can perform badly and one is advised to use more robust estimators such as the limited information maximum likelihood (LIML) estimator (see Hahn and Hausman, 2003; Baum et al., 2007). Under i.i.d. errors the relevant test is the F-statistic version of the Cragg-Donald Wald test statistic. Where possible, we use the Kleibergen-Paap (2006) Wald rank F statistic that is valid for estimation that is robust to heteroskedastic and clustered error structures. The critical values for this test are from Stock and Yogo (2005) and are shown in the notes section of Table 5 and Table 6 The Kleibergen-Paap (2006) Wald rank F statistic for weak identification reported in Table 5 is higher than the Stock and Yogo (2005) critical values, suggesting that our instruments are not weak. The Cragg-Donald F-statistic shown in Table 6 for the sectoral regressions also tends to exceed the critical values for the weak identification tests. The only exception is model (2) for the real estate sector.

Since the critical values are for the case of i.i.d error structures, we also perform tests that are robust to the presence of weak instruments: the Anderson-Rubin (1949) Wald and F statistics tests and the Stock-Wright (2000) S statistic test. These are tests of the significance of the endogenous regressors in the structural model with the null hypothesis and are specified in an equivalent way as to testing for the joint significance of the excluded instruments in the reduced form of he model. If we reject the null that the coefficients of the excluded instruments in the reduced form equation are equal to zero, we also reject the null hypothesis that the coefficients of the endogenous regressors in the structural model are equal to zero (Baum et al., 2007, p. 25). The results reported in Table 5 show that the null hypothesis is rejected and so we conclude that the instruments are relevant. As for the sectoral regressions, the test statistics shown in Table 6 sometimes fail to reject the null hypothesis that the coefficients of the instruments and the coefficients of the endogenous regressors are jointly equal to zero (construction, hotels and restaurants, financial intermediation, other business activities, and public services).

Taken together, the results for the IV estimations generally indicate that our instruments can effectively be treated as exogenous. The estimate for the wage elasticity of employment density obtained from model (1) (model (2)) is 0.8% (0.7%). This is about 20% smaller than the elasticity value obtained with the FE estimator (1.0%), and about 64% smaller than the elasticity value obtained with the pooled OLS estimate (2.2%). As for the effect of market potential, the instrumental variables estimator for model (1) (model (2)) proposes that doubling the size of the market potential increases hourly wages by 5.8% (4.3%), in contrast to the 5.4% (10.0%) obtained with the FE (POLS) estimator. Finally, the compensation for additional commuting also differs considerably across estimators. Treating home-to-work distance as endogenous (exogenous) reduces the size of the coefficient from 5.8% (POLS)/1.4% (FE) to 3.3% (1.4%).[13]

Comparing the pooled OLS estimates with the instrumental variables estimates we observe that the elasticities of employment density and market potential are revised downwards. This is intuitive because of the positive association between the measures of agglomeration and workers' unobservable skills, which in our models account for differences in years of education and other unobserved productive abilities.

When we compare the instrumental variables estimates for agglomeration externalities with those obtained from the FE estimator, we find that correcting for endogeneity of agglomeration revises the size of the elasticity of employment density (market potential) downwards (upwards) if model (1) is taken as the reference case; both elasticity values are revised downwards is model (2) is taken as the reference case. It is not straight forward to propose a theory for why the direction of bias of the FE estimates is positive for employment density and negative for market potential. Both the FE model and the IV model are in theory cleaned from workers' unobserved heterogeneity, so the bias in the FE estimations should result purely from nonzero correlation between the idiosyncratic error term and the model covariates.

Considering the results for the sectoral models with valid instruments, we observe that labour market density has a positive impact of wages for the other business activities sector (1.52%) and a negative effect for the wholesale and retail sector. The effect from access to product markets is significant and positive only for the wholesale and retail sector.

The following paragraphs summarise the results up to this point. Our findings are consistent with previous evidence from worker-level longitudinal wage functions that also deal with workers' unobservable heterogeneity and reverse causality between productivity and agglomeration. We improve on these models by also estimating the wage compensation of commute distances and attempting to correct for its endogenous nature, together with agglomeration economies.

Combes et al. (2008a,b) estimate elasticities of wages with respect to employment density (market potential) of about 3% (2%) for French employment areas. Mion and Naticchioni (2005) estimate somewhat weaker elasticity of wages with respect to employment density (0.22%), but similar size elasticity of wages with respect to market potential (3.19%), for the Italian Provinces. Looking at Travel-to-Work Areas (best available approximations to labour markets) in Great Britain,

[13] Fitting the instrumental variables estimator to the worker-firm two-way error components model makes very little difference to the parameter estimates obtained from the worker fixed-effects instrumental variables estimation; the wage elasticity of employment density (market potential) is 0.71% (5.1%) compared to 0.79% (5.78%) respectively.

we obtain elasticities of wages with respect to employment density (market potential) of about 0.8% (5.8%).

We conclude that controlling for worker-specific unobservable heterogeneity and correcting for reverse causality between earnings and agglomeration are both important to reduce bias in the estimates of the effects from labour market size and market potential. Controlling for worker-specific unobservable heterogeneity halves the size of the elasticity of employment density, while correcting for simultaneity endogeneity further reduces the effect by more than one-fifth. As regards the effect from market potential, controlling for worker-specific unobservable heterogeneity nearly halves the elasticity value (45%), while correcting for simultaneity endogeneity produces a slight increase of about 7%.

**Table 5: Instrumental variables regressions for the whole economy**

| Dependent variable: log of real hourly wage | model (1) | model (2) |
|---|---|---|
| Log of home-to-work distance | 0.0143** | 0.0330** |
| Age | 0.0429** | 0.0426** |
| Age2 | -0.0006** | -0.0006** |
| Full-timer | -0.0633** | -0.0638** |
| Log of firm size | 0.0060** | 0.0054** |
| Log of employment density | 0.0079** | 0.0067* |
| Log of market potential | 0.0578** | 0.0431** |
| Log of index of industrial specialisation | -0.0049 | -0.0067 |
| Log of index of economic diversity | -0.0074+ | -0.0080+ |
| Log of area | 0.0150** | 0.0117** |
| **Instrument Exogeneity** | | |
| Hansen J stat (overidentification test of all instruments) | 4.9384 | 4.9142 |
| Hansen J stat (p-value) | 0.2937 | 0.2962 |
| **Instruments Relevance** | | |
| F stat (test of excluded instruments): log of density | 2078*** | 1775.04*** |
| Shea partial R2: log of density | 0.3046 | 0.3025 |
| partial R2: log of density | 0.3141 | 0.3171 |
| F stat (test of excluded instruments): log of MP | 11921.04*** | 10271.84*** |
| Shea partial R2: log of MP | 0.7466 | 0.667 |
| partial R2: log of MP | 0.7698 | 0.7725 |
| F stat (test of excluded instruments): log of home-work distance | | 376.29*** |
| Shea partial R2: log of home-work distance | | 0.0181 |
| partial R2: log of home-work distance | | 0.0211 |
| (1) Under-identification tests | | |
| Kleibergen-Paap rk LM statistic | 4226.97*** | 2315.03*** |
| Kleibergen-Paap rk Wald statistic | 10528.91*** | 2504.56*** |
| (2) Weak identification tests | | |
| Kleibergen-Paap Wald rk F statistic [a] | 1754.67 | 357.76 |
| (3) Weak-instrument-robust inference | | |
| Anderson-Rubin Wald test | 7.31*** | 18.56*** |
| Anderson-Rubin Wald test | 43.87*** | 129.93*** |
| Stock-Wright LM S statistic | 43.58*** | 129.08*** |
| Observations | 605359 | 605359 |
| RMSE | 0.1637 | 0.1637 |
| R2 centered | 0.0792 | 0.0754 |
| Time, occupation, industry, region and worker dummies | yes | yes |
| Firm dummies | no | no |

Standard errors are cluster-robust. Clustering is on workers. Significance levels: + p<0.10, * p<0.05, ** p<0.01
Notes: first-stage F-stat cluster-robust; Under identification, weak identification and weak-identification-robust test statistics cluster-robust; Anderson-Rubin stat cluster-robust.

a Stock-Yogo weak ID test critical values (Source: Stock and Yogo (2005)).

| | | |
|---|---|---|
| 5% maximal IV relative bias | 15.72 | 13.95 |
| 10% maximal IV size | 21.68 | |

**Table 6: Instrumental variables regressions for individual industry groups**

| Dependent variable: log of real hourly wage | Primary | | Manufacturing | | Electricity, gas & water | | Construction | | Wholesale & retail | | Hotels & restaurants | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | model (1) | model (2) | model (1) | model (2) | model (1) | model (2) | model (1) | model (2) | model (1) | model (2) | model (1) | model (2) |
| Log of home-to-work distance | 0.0086 | -0.0968 | 0.0086** | 0.0479** | 0.0076+ | -0.0008 | 0.0082** | 0.0047 | 0.0075** | 0.0150 | 0.0031 | -0.0101 |
| Age | 0.0183 | -0.0063 | 0.0353** | 0.0352** | 0.0777** | 0.0903** | 0.0552** | 0.0786** | 0.0349** | 0.0311** | 0.0541** | 0.0270 |
| Age2 | -0.0004** | -0.0006** | -0.0005** | -0.0005** | -0.0008** | -0.0009** | -0.0006** | -0.0006** | -0.0006** | -0.0006** | -0.0004** | -0.0004** |
| Full-timer | -0.1780** | -0.2442** | -0.1292** | -0.1317** | -0.1351** | -0.1215** | -0.1916** | -0.1699** | -0.0438** | -0.0420** | -0.0633** | -0.0635** |
| Log of firm size | -0.0161* | -0.0453* | 0.0210** | 0.0184** | -0.0047 | -0.0082 | 0.0089** | 0.0100* | 0.0021** | 0.0011 | -0.0012 | -0.0034 |
| Log of employment density | 0.0334 | -0.1681 | -0.0080 | -0.0009 | 0.0205 | 0.0235 | 0.0055 | 0.0077 | -0.0103+ | -0.0166* | -0.0179 | 0.0187 |
| Log of market potential | 0.0030 | 1.2444 | 0.0735** | 0.0187 | 0.0835 | 0.1947 | 0.0393 | 0.0586 | 0.0535* | 0.0704** | 0.0561 | 0.0013 |
| Log of index of industrial specialisation | -0.0184 | 0.0975 | -0.0037 | -0.0200 | -0.0208 | -0.0037 | -0.0344 | -0.0267 | 0.0386* | 0.0493** | 0.0433 | 0.0484 |
| Log of index of economic diversity | -0.0464 | -0.0699 | -0.0105 | -0.0059 | -0.0839** | -0.1305** | -0.0199 | -0.0134 | -0.0086 | -0.0011 | 0.0285 | 0.0151 |
| Log of area | -0.0497+ | -0.1914** | 0.0044 | 0.0014 | 0.0372* | 0.0770** | 0.0159+ | 0.0157 | 0.0089+ | 0.0036 | 0.0108 | -0.0143 |
| **Instrument Exogeneity** | | | | | | | | | | | | |
| Sargan test statistic | 10.502 | 9.732 | 17.39 | 22.338 | 1.122 | 6.129 | 3.323 | 6.502 | 11.153 | 3.133 | 7.524 | 4.311 |
| Sargan p-value | 0.0328 | 0.0452 | 0.0016 | 0.0002 | 0.8907 | 0.1897 | 0.5052 | 0.1646 | 0.0249 | 0.5358 | 0.1107 | 0.3655 |
| **Instrument Relevance** | | | | | | | | | | | | |
| Shea partial R2: log of density | 0.23 | 0.2718 | 0.2338 | 0.2279 | 0.4028 | 0.4908 | 0.2967 | 0.2832 | 0.3187 | 0.2859 | 0.3433 | 0.3404 |
| partial R2: log of density | 0.3528 | 0.8975 | 0.2672 | 0.2644 | 0.4317 | 0.5186 | 0.3067 | 0.2861 | 0.3207 | 0.3148 | 0.3464 | 0.3537 |
| F stat (test of excluded instruments): log of MP | 275.51 | 1244.84 | 3748.71 | 2857.51 | 255.42 | 222.83 | 979.85 | 449.06 | 4487.9 | 3248.18 | 812.92 | 366.48 |
| p-value | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.0000 |
| Shea partial R2: log of MP | 0.5395 | 0.2537 | 0.6394 | 0.5199 | 0.6919 | 0.3615 | 0.7191 | 0.7303 | 0.7636 | 0.7463 | 0.7987 | 0.6549 |
| partial R2: log of MP | 0.8276 | 0.8917 | 0.7306 | 0.7415 | 0.7416 | 0.7495 | 0.7432 | 0.747 | 0.7685 | 0.7709 | 0.8059 | 0.8037 |
| F stat (test of excluded instruments): log of MP | 2426.58 | 1170.6 | 27885.92 | 22802.7 | 964.97 | 618.86 | 6411.77 | 3308.45 | 31544.58 | 23793.07 | 6367.99 | 2741.73 |
| p-value | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.0000 |
| Shea partial R2: log of home-work distance | | 0.0775 | | 0.0185 | | 0.0508 | | 0.0297 | | 0.0203 | | 0.0298 |
| partial R2: log of home-work distance | | 0.23 | | 0.0232 | | 0.102 | | 0.0304 | | 0.0228 | | 0.0382 |
| F stat (test of excluded instruments): log of home-work distance | | 42.47 | | 188.52 | | 23.5 | | 35.15 | | 164.65 | | 26.57 |
| p-value | | 0.000 | | 0.000 | | 0.000 | | 0.000 | | 0.000 | | 0.0000 |
| (1) Under-identification tests | | | | | | | | | | | | |
| Anderson canon. corr. likelihood ratio stat. | 694.8 | 79.72 | 14071.38 | 1028.86 | 801.59 | 74.78 | 3871.02 | 234.11 | 17957.26 | 1003.51 | 3134.51 | 140.62 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| p-value | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Cragg-Donald stat. | 898.22 | 86.42 | 18226.26 | 1048.22 | 1315.47 | 78.75 | 5455.54 | 241.28 | 26203.4 | 1024.26 | 4744.28 | 144.93 |
| p-value | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| (2) Weak identification tests | | | | | | | | | | | | |
| Cragg-Donald F- tat[a] | 147.95 | 11.94 | 3035.94 | 149.65 | 215.5 | 10.98 | 906.8 | 34.31 | 4364.48 | 146.22 | 787.63 | 20.55 |
| (3) Weak-instrument-robust inference | | | | | | | | | | | | |
| Anderson-Rubin test F stat | 2.0419 | 2.00 | 4.6889 | 8.7301 | 0.9099 | 1.7323 | 0.7829 | 1.1567 | 3.3158 | 2.6505 | 1.6912 | 0.7169 |
| p-value | 0.0569 | 0.0527 | 0.0001 | 0.000 | 0.4866 | 0.0974 | 0.5831 | 0.3243 | 0.0029 | 0.0097 | 0.1187 | 0.6577 |
| Anderson-Rubin test -Chi2 | 12.3970 | 14.4565 | 28.1499 | 61.1502 | 5.5544 | 12.4189 | 4.7104 | 8.1339 | 19.9072 | 18.5671 | 10.1869 | 5.0569 |
| p-value | 0.0537 | 0.0436 | 0.0001 | 0.000 | 0.4749 | 0.0876 | 0.5815 | 0.3209 | 0.0029 | 0.0097 | 0.117 | 0.653 |
| Stock-Wright LM S statistic | 12.35 | 14.26 | 28.14 | 61.08 | 5.54 | 12.32 | 4.71 | 8.13 | 19.9 | 18.56 | 10.18 | 5.05 |
| p-value | 0.0547 | 0.0468 | 0.0001 | 0.000 | 0.4767 | 0.0906 | 0.5817 | 0.3217 | 0.0029 | 0.0097 | 0.1174 | 0.6537 |
| Observations | 4399 | 1544 | 87071 | 79049 | 2874.0000 | 2127 | 19608 | 11925 | 83726 | 73579 | 14346 | 7600 |
| R2 centered | 0.0756 | 0.1040 | 0.0504 | 0.0353 | 0.1118 | 0.1518 | 0.0846 | 0.1021 | 0.0397 | 0.0412 | 0.0568 | 0.0570 |
| RMSE | 0.1450 | 0.1633 | 0.1295 | 0.1290 | 0.1219 | 0.1187 | 0.1653 | 0.1579 | 0.1571 | 0.1551 | 0.1658 | 0.1650 |

Significance levels: + p<0.10, * p<0.05, ** p<0.01. Models include dummies for years, occupations, regions, and workers.

a Stock-Yogo weak ID test critical values (Source: Stock and Yogo (2005).

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5% maximal IV relative bias | 15.72 | 13.95 | 15.72 | 13.95 | 15.72 | 13.95 | 15.72 | 13.95 | 15.72 | 13.95 | 15.72 | 13.95 |
| 10% maximal IV relative bias | 9.48 | 8.5 | 9.48 | 8.5 | 9.48 | 8.5 | 9.48 | 8.5 | 9.48 | 8.5 | 9.48 | 8.5 |
| 20% maximal IV relative bias | 6.08 | 5.56 | 6.08 | 5.56 | 6.08 | 5.56 | 6.08 | 5.56 | 6.08 | 5.56 | 6.08 | 5.56 |
| 30% maximal IV relative bias | 4.78 | 4.44 | 4.78 | 4.44 | 4.78 | 4.44 | 4.78 | 4.44 | 4.78 | 4.44 | 4.78 | 4.44 |

## Table 6: Instrumental variables regressions for individual industry groups *(continued)*

| Dependent variable: log of real hourly wage | Transport, storage & communication | | Financial intermediation | | Real estate | | Renting, IT,R&D | | Other business activities | | Public services | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | model (1) | model (2) | model (1) | model (2) | model (1) | model (2) | model (1) | model (2) | model (1) | model (2) | model (1) | model (2) |
| Log of home-to-work distance | 0.0165** | 0.0349** | 0.0124** | -0.0111 | 0.0138** | 0.0894 | 0.0148** | 0.0646** | 0.0097** | 0.0422+ | 0.0068** | 0.0245** |
| Age | 0.0074 | 0.0099 | 0.0751** | 0.0787** | 0.0508** | 0.0543* | 0.0674** | 0.0755** | 0.0807** | 0.0764** | 0.0396** | 0.0383** |
| Age2 | -0.0002** | -0.0002** | -0.0009** | -0.0009** | -0.0005** | -0.0007** | -0.0007** | -0.0007** | -0.0008** | -0.0008** | -0.0005** | -0.0005** |
| Full-timer | -0.1213** | -0.0941** | -0.0429** | -0.0475** | -0.0848** | -0.0875** | -0.1279** | -0.1201** | -0.0822** | -0.0778** | -0.1086** | -0.1104** |
| Log of firm size | 0.0102** | 0.0107** | 0.0051** | 0.0044** | 0.0088** | 0.0065 | 0.0194** | 0.0024 | 0.0045** | 0.0022 | 0.0036** | 0.0028** |
| Log of employment density | -0.0013 | 0.0000 | -0.0083 | -0.0136 | 0.0087 | 0.1753** | 0.0360** | 0.0717** | 0.0152* | 0.0053 | 0.0034 | 0.0050 |
| Log of market potential | 0.1108** | 0.1089** | 0.0714* | 0.1298** | -0.1301 | -0.3390 | 0.0382 | -0.0046 | 0.0357 | -0.0084 | 0.0168 | -0.0031 |
| Log of index of industrial specialisation | 0.0597* | 0.0676* | 0.0965** | 0.0887* | -0.0271 | -0.1113 | -0.0366 | -0.0717 | -0.0290 | -0.0284 | -0.0091 | -0.0142 |
| Log of index of economic diversity | 0.0049 | 0.0027 | 0.0169 | 0.0019 | -0.0196 | -0.0864+ | -0.0421* | -0.0441* | -0.0025 | 0.0109 | 0.0077 | 0.0043 |
| Log of area | 0.0080 | 0.0073 | 0.0197** | 0.0276** | -0.0060 | -0.0534 | -0.0152 | -0.0165 | 0.0164* | 0.0054 | 0.0059* | 0.0032 |
| **Instrument Exogeneity** | | | | | | | | | | | | |
| Sargan test statistic | 10.733 | 12.526 | 5.306 | 1.804 | 14.85 | 2.114 | 25.716 | 17.325 | 4.621 | 2.836 | 5.76 | 10.886 |
| Sargan p-value | 0.0297 | 0.0138 | 0.2573 | 0.7717 | 0.005 | 0.7148 | 0.0000 | 0.0017 | 0.3284 | 0.5857 | 0.2178 | 0.0279 |
| **Instrument Relevance** | | | | | | | | | | | | |
| Shea partial R2: log of density | 0.3351 | 0.3456 | 0.2756 | 0.3251 | 0.2895 | 0.2001 | 0.4488 | 0.3193 | 0.4011 | 0.385 | 0.2798 | 0.2852 |
| partial R2: log of density | 0.3492 | 0.3667 | 0.2848 | 0.3372 | 0.2928 | 0.2074 | 0.4522 | 0.4902 | 0.4175 | 0.4327 | 0.2891 | 0.2942 |
| F stat (test of excluded instruments): log of MP | 2363.65 | 1780.46 | 1532.36 | 1449.6 | 316.98 | 76.83 | 1540.67 | 1093.81 | 3283.08 | 2385.89 | 10109.93 | 8330.38 |
| p-value | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Shea partial R2: log of MP | 0.7688 | 0.7795 | 0.7276 | 0.7282 | 0.7458 | 0.4857 | 0.8235 | 0.7035 | 0.7338 | 0.6079 | 0.7397 | 0.6393 |
| partial R2: log of MP | 0.8009 | 0.8182 | 0.7518 | 0.7668 | 0.7544 | 0.8221 | 0.8296 | 0.8512 | 0.7638 | 0.7674 | 0.7644 | 0.7618 |
| F stat (test of excluded instruments): log of MP | 17726.13 | 13835.5 | 11658.23 | 9364.96 | 2351.71 | 1356.82 | 9088.01 | 6505.47 | 14812.28 | 10320.21 | 80665.83 | 63894.98 |
| p-value | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Shea partial R2: log of home-work distance | | 0.0287 | | 0.0558 | | 0.0111 | | 0.0197 | | 0.0068 | | 0.0165 |
| partial R2: log of home-work distance | | 0.0296 | | 0.058 | | 0.0191 | | 0.0323 | | 0.0092 | | 0.0191 |
| F stat (test of excluded instruments): log of home-work distance | | 93.78 | | 175.53 | | 5.73 | | 37.98 | | 29.16 | | 389.87 |
| p-value | | 0.0000 | | 0.0000 | | 0.0000 | | 0.0000 | | 0.0000 | | 0.0000 |

(1) Under-identification tests

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Anderson canon. corr. likelihood ratio stat. | 8679.76 | 618.05 | 6133.17 | 1115.57 | 1325.07 | 23.13 | 5008.26 | 155.29 | 10761.56 | 148.89 | 40355.34 | 2300.63 |
| p-value | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0003 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Cragg-Donald stat. | 12914.32 | 636.29 | 8346.41 | 1181.55 | 1856.67 | 23.38 | 9035.83 | 158.36 | 17674.23 | 149.9 | 55319.3 | 2339.09 |
| p-value | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| (2) Weak identification tests | | | | | | | | | | | | |
| Cragg-Donald F stat[a] | 2149.46 | 90.75 | 1388.9 | 168.49 | 307.1 | 3.29 | 1501.15 | 22.52 | 2941.85 | 21.38 | 9217.66 | 334.07 |
| (3) Weak-instrument-robust inference | | | | | | | | | | | | |
| Anderson-Rubin test –F stat | 4.4895 | 5.1636 | 1.7107 | 2.1577 | 2.7871 | 1.8055 | 6.4623 | 6.2196 | 1.8883 | 1.0760 | 1.4717 | 4.0283 |
| p-value | 0.0001 | 0.000 | 0.114 | 0.0347 | 0.0105 | 0.0821 | 0.000 | 0.000 | 0.0787 | 0.3757 | 0.1834 | 0.0002 |
| Anderson-Rubin test -Chi2 | 26.9735 | 36.2059 | 10.2802 | 15.1308 | 16.8499 | 12.8476 | 38.8984 | 43.7344 | 11.3449 | 7.5441 | 8.8325 | 28.2053 |
| p-value | 0.0001 | 0.000 | 0.1133 | 0.0344 | 0.0099 | 0.0759 | 0.000 | 0.000 | 0.0783 | 0.3745 | 0.1832 | 0.0002 |
| Stock-Wright LM S statistic | 26.95 | 36.15 | 10.28 | 15.12 | 16.79 | 12.77 | 38.76 | 43.5 | 11.34 | 7.54 | 8.83 | 28.2 |
| p-value | 0.0001 | 0.000 | 0.1135 | 0.0345 | 0.0101 | 0.0779 | 0.000 | 0.000 | 0.0784 | 0.3747 | 0.1833 | 0.0002 |
| Observations | 37341 | 30752 | 32401 | 28393 | 6858 | 3224 | 16244 | 11727 | 40801 | 32970 | 206767 | 195030 |
| R2 centered | 0.0489 | 0.0396 | 0.0637 | 0.0589 | 0.0840 | 0.0564 | 0.0465 | 0.0032 | 0.0737 | 0.0720 | 0.0939 | 0.0911 |
| RMSE | 0.1454 | 0.1404 | 0.1572 | 0.1548 | 0.1542 | 0.1333 | 0.1732 | 0.1648 | 0.1684 | 0.1620 | 0.1503 | 0.1493 |

Significance levels: + p<0.10, * p<0.05, ** p<0.01. Models include dummies for years, occupations, regions, and workers.

a Stock-Yogo weak ID test critical values (Source: Stock and Yogo (2005).

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5% maximal IV relative bias | 15.72 | 13.95 | 15.72 | 13.95 | 15.72 | 13.95 | 15.72 | 13.95 | 15.72 | 13.95 | 15.72 | 13.95 |
| 10% maximal IV relative bias | 9.48 | 8.5 | 9.48 | 8.5 | 9.48 | 8.5 | 9.48 | 8.5 | 9.48 | 8.5 | 9.48 | 8.5 |
| 20% maximal IV relative bias | 6.08 | 5.56 | 6.08 | 5.56 | 6.08 | 5.56 | 6.08 | 5.56 | 6.08 | 5.56 | 6.08 | 5.56 |
| 30% maximal IV relative bias | 4.78 | 4.44 | 4.78 | 4.44 | 4.78 | 4.44 | 4.78 | 4.44 | 4.78 | 4.44 | 4.78 | 4.44 |

## 5.3. The spatial decay of agglomeration externalities

One of the key points emerging from previous research (and our own research so far) is the insufficient understanding of the pattern of spatial decay of agglomeration economies. Generally, the measurement strategies used make two important assumptions for the geographic scope of agglomeration externalities. First, they tend to assume that all agents in one given jurisdiction enjoy the same level of benefits from concentration of economic activity, that is, there is no decay in the intensity of the benefits with increasing distance from its source. Second, the spatial reach of these effects is constrained to the borders of the geographic units used by researchers and so, by definition, there is no scope for spatial spillovers. In this chapter we undertake further analysis with the goal of improving the current understanding of the geographic scope of agglomeration externalities.

In this section we adopt a strategy similar to that used by Rosenthal and Strange (2008) to estimate the spatial attenuation of agglomeration externalities. This method allows for the identification of both the geographic scope and the pattern of spatial decay of the effects from agglomeration externalities. The pattern of spatial decay can be inferred from the statistical significance and size of the coefficients for the effect of the proximity to jobs in the various concentric distance bands around the centroid of each worker's workplace ward. The model to be estimated can be written as shown below:

$$lnw_{it} = \alpha_0 + X_{it}\beta + \sum_k \alpha_k .emp_{k,it} + \delta_t + \lambda_{o(i,t)} + \sigma_{s(i,t)} + \eta_i + \omega_{\mathrm{r(i,t)}} + \varepsilon_{it} \, , \tag{15}$$

where the variables are as before and the $k$ identifies the successive concentric distance ring, and $emp_{k,it}$ measures the number of jobs within each concentric ring $k$, which is defined as follows:

$$emp_{k,it} = \sum_i w_{ij}(k).emp_{q(j,t)}, \qquad i \neq j \, , \tag{16}$$

with $emp_{q(j,t)}$ the number of jobs in worker's $j$ ward at time $t$. $w_{ij}(k)$ is a binary weight that takes value zero or one, depending on whether the bilateral distance between any two wards is within certain lower and upper thresholds, defined below.

$$w_{ij}(k) = \begin{cases} 1 & \text{if } d_{ij}(k-1) < d_{ij} \leq d_{ij}(k) \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

with $d_{ij}(k)$ the distance in kilometres between the wards of workers $i$ and $j$, where $d_{ij}(0)=0$. The general definition of $d_{ij}(k)$ we chose is defined below.

$$d_{ij}(k) = \begin{cases} 5\ (5)\ 25 \text{ kilometres} & \text{for} \quad k = 1\ (1)\ 4 \\ 25\ (25)\ 100 \text{ kilometres} & \text{for} \quad k = 5\ (1)\ 8 \end{cases} \tag{18}$$

There is no definite guideline on what a unique spatial regime should be. In choosing our concentric rings we try to be parsimonious in the number of rings, while providing a comprehensive definition of the overall spatial coverage and the individual distance intervals. In particular we wish to include distance bands that go beyond the average radius of the TTWAs (about 16 kilometres), so that we can test for the presence of spatial spillovers outside labour markets.

To build the measures of economic mass, we use data on yearly employment for the 2001 CAS wards, available from the Annual Business Inquiry (ABI) employee analysis. Data for employment at the Census Area Statistics ward level is only consistently available from 2003 onward, which limits the current analysis to the years between 2003 and 2006 instead of 2002 and 2006. For each worker's workplace Census Area Statistic (CAS) wards (referred to as wards throughout the text for simplicity) we construct concentric rings of various radii away from that ward that measure the number of jobs inside each respective ring.

It is worth noting some comments on the differences to previous measurement strategies. Di Addario and Patacchini (2007) base the concentric rings around the centroid of the worker's area of residence instead of the workplace area. Moreover, they use local labour markets (LLM) as spatial units, which are considerably larger than the areas used by Rosenthal and Strange (2008) and the areas used here. Considering that the LLM can be approximated by circles, the average radius is 16.8 kilometres, which in turn corresponds to an average area of 282.24 square kilometres.

There are 10,071 wards in Great Britain compared to the 297 TTWAs. Compared to the TTWAs the wards are therefore much smaller areas. Assuming that both geographies can be approximated by a circle, the mean and median area size (radius) of the wards is 23 square kilometres (2.7 kilometres) and 5 square kilometres (1.3 kilometres) respectively, against the 772.6 square kilometres (15.7 kilometres) and 600.5 square kilometres (13.8 kilometres) of the TTWA. The geographic units used by Di Addario and Patacchini (2007) are more in line with our TTWAs, which we believe are too large to provide a in-depth understanding of the process of spatial decay of agglomeration externalities.

Table 7 shows the results for the FE estimator using the various distance rings. The Table shows that a 100,000 increase in the number of jobs within 5 kilometres raises wages by approximately 1.19% and the effect falls sharply by to 0.38% 5 kilometres away (5-10 kilometres) and to 0.15% another 5 kilometres away (10-15 kilometres). The size of the effects remains about 0.15%

for the 15-20 kilometres band but becomes insignificant between 20-25 kilometres. This suggests that the effects from agglomeration externalities are likely to take place within labour markets boundaries (the average radius is 16 kilometres). Further analysis of the results reported in Table 8 shows that between-labour market area effects are also at place. In fact, while the effects from agglomeration externalities appear not to be significant for the 20 to 25 kilometres band, they reappear as significant for the 25 to 50 and 50 to 75 kilometres bands. An increase of 100,000 jobs within 25-50 km and 50-75 km raises wages by 0.11% (0.05%). The estimates are not statistically significant beyond 75 kilometres.

This is a very interesting finding, in particular if related to the spatial scale of the traditional *Marshallian* drivers of agglomeration economies mentioned in section 3.1. The effects from the economic mechanisms outlined therein operate over different spatial ranges (see Rosenthal and Strange, 2001; Forni and Paba, 2002; Combes et al., 2008b; Rosenthal and Strange, 2006; Combes et al., 2008a). Labour market pooling externalities are determined by the size of interactions at the labour market level, which is defined for the most part by the extent of daily commuting trips. By contrast, productivity effects from access to markets for final and intermediate products are likely to occur over longer distance ranges, and therefore are likely to display a spatial scale larger than a local labour market. Knowledge spillovers tend to be much localised in very short distance ranges and decay rapidly with increased distance.

**Table 7: Regression results for the spatial decay of agglomeration externalities**

| %change in wage from +100,000 jobs | |
|---|---|
| Employment 0-5 km | 1.19** |
| Employment 5-10 km | 0.38** |
| Employment 10-15 km | 0.15* |
| Employment 15-20 km | 0.16* |
| Employment 20-25 km | 0.07 |
| Employment 25-50 km | 0.11** |
| Employment 50-75 km | 0.05** |
| Employment 75-100 km | 0.01 |
| Observations | 542153 |
| F-statistic | 232.41 |
| R2 overall | 0.42 |
| R2 within | 0.07 |
| R2 between | 0.44 |
| Hausman (FE vs RE) | 12795** |
| Time dummies | yes |
| Occupation dummies | yes |
| Industry dummies | yes |
| Region dummies | yes |
| Worker dummies | yes |
| Firm dummies | no |

Standard errors are cluster-robust. Clustering is on workers. Significance levels: * p<0.05, ** p<0.01.

Compared to our findings, Di Addario and Patacchini (2007) find smaller effects for an additional 100,000 inhabitants. The increase in wages within 4 kilometres is about 0.1%, but falls to 0.04% (0.02%) in the 4-8 (8-12) kilometres proximity band, and becomes insignificant beyond 12 kilometres. Rosenthal and Strange (2008) estimate that the effect of employment concentration falls rapidly with distance: adding 100,000 full-time workers within 5 miles increase wages by about 2%, while the effect falls to 0.5% for the 5-25 miles ring. The main conclusion is that the great part of the spillover effects of agglomeration of employment takes place within 5 miles (about 8 kilometres), but it can reach as far as 50 miles (nearly 81 km).

In light of the above considerations, the finding that the spillover effects from agglomeration externalities are likely to occur within labour markets (the TTWAs) is believed to be a reasonable result. On the other hand, the significant effects for the 25 to 50 and 50 to 75 kilometres bands can reflect the spatial scale of input sharing externalities, while the non-significant gap between 20 and 25 kilometres may result from the presence of undeveloped land around urban areas resulting from green belt land use planning policy which limits urbanisation. Interestingly, our evidence also suggests that distance decay may be lumpy, that is, the effects from agglomeration externalities do not decay continuously with increasing distance.

Finally, it is important to think about the possible implications from our findings to transport policy. Knowledge about the pattern of spatial decay of the effects from agglomeration economies can be important for the appraisal of transport schemes. Given that transport networks alter the distribution of spatial concentration of economic activities, reflected in measures of employment density and market potential, our evidence can help the decision about the area of influence for a scheme by offering a "boundary" for the scope of the effects from agglomeration.

## 6. Conclusions

This paper investigates the role of agglomeration externalities on individual worker's hourly wages. It plays particular attention to effect from labour market scale and access to product markets as approximated by a market potential type function. The empirical analyses performed allow controlling not only for the effects from unobserved worker heterogeneity, but also for the effects from firm unobserved heterogeneity.

We find that worker unobserved heterogeneity plays a very important role in the explanation of spatial wage disparities, whereas firm unobserved heterogeneity has only a slight impact on the magnitude of the parameter estimates. Controlling for worker unobserved heterogeneity can halve the size of the wage elasticity of agglomeration externalities. Our findings also support the presence of positive sorting of workers across space: more able workers tend to work on urban areas and more agglomerated labour markets.

The paper also devotes considerable attention to the issue of reverse causality. It deals both with simultaneity between wages and agglomeration, and between wages and home-to-work commutes. Compared to worker unobserved heterogeneity, correcting for simultaneity bias appears to have only a minor impact on the size of the parameter estimates. Our best estimate for the effect of labour market density (market potential) is 0.8% (5.8%). This means that doubling labour market's employment density can raise hourly earnings by nearly 1%, while halving the distances to other markets produces an increase of hourly wages of nearly 3%.

The last piece of evidence refers to the spatial attenuation of agglomeration externalities. We test for the effect of proximity to jobs using concentric distance bands around the centroid of each worker's workplace ward. The findings propose that a 100,000 increase in the number of jobs within 5 kilometres raises hourly wages by approximately 1.19%; this effect falls sharply (0.38%) if the increase in the number of jobs occurs 10 kilometres away, and remains around 0.15% if the increase in the number of jobs occurs between 10 to 20 kilometres.

The sharp decay of agglomeration externalities suggests that the concentration of economic activity is important. Nevertheless, its importance may vary with economic sector as suggested by the differences in the size of the effects from labour market scale and access to product markets across industry groups. This is a topic for future research, which will explore the spatial attenuation of the effects from the main mechanisms determining agglomeration for the different sectors of the economy.

# Appendix A

Table A1: List of industry groups included in the analysis

| Two-digit SIC codes | Description |
|---|---|
| SIC 1-14 | Primary industries |
| SIC 15-37 | Manufacturing |
| SIC 40-41 | Electricity, gas, and water supply |
| SIC 45 | Construction |
| SIC 50-52 | Wholesale & retail trade; repair of motor vehicles, motorcycles and personal and household goods |
| SIC 55 | Hotels and restaurants |
| SIC 60-64 | Transport, storage and communication |
| SIC 65-67 | Financial intermediation |
| SIC 70 | Real estate activities |
| SIC 71-73 | Renting of machinery and equipment; computer and related activities; research and development |
| SIC 74 | Other business activities |
| SIC 75-85 | Public administration and defence; compulsory social security; education; health and social work |

Source: ONS. UK Standard Industrial Classification of Economic Activities 2003.
Available at: http://www.statistics.gov.uk/methods_quality/sic/downloads/UK_SIC_Vol1(2003).pdf

## References


ADSERA, A. (2000) Sectoral Spillovers and the Price of Land: a Cost Analysis. *Regional Science and Urban Economics,* 30**,** 565-585.

AMITI, M. & CAMERON, L. (2007) Economic Geography and Wages. *The Review of Economics and Statistics,* 89**,** 15-29.

ANDERSON, T. W. (1951) Estimating Linear Restrictions on Regression Coefficients for Multivariate Normal Distributions. *Annals of Mathematical Statistics* 2nd ed. New York, John Wiley

ANDERSON, T. W. & RUBIN, H. (1949) Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations. *Annals of Mathematical Statistics,* 20**,** 46-63.

BAUM, C. F., SCHAFFER, M. E. & STILLMAN , S. (2007) Enhanced Routines for Instrumental Variables/Generalized Method of Moments Estimation and Testing. *Stata Journal,* 7**,** 465-506.

BENITO, A. & OSWALD, A. (2000) Commuting in Great Britain in the 1990s. Department of Economics: University of Warwick.

BLASIO, G. D. & DI ADDARIO, S. (2005) Do Workers Benefit from Industrial Agglomeration? *Journal of Regional Science,* 45**,** 797.

BOUND, J., JAEGER, D. A. & BAKER, R. M. (1995) Problems with Instrumental Variables Estimation When the Correlation Between Instruments and the Endogenous Explanatory Variable is Weak. *Journal of the Ameriacan Statistical Association,* 90**,** 443-450.

CARLINO, G. A. & VOITH, R. (1992) Accounting for Differences in Aggregate State Productivity. *Regional Science and Urban Economics,* 22**,** 597-617.

CICCONE, A. (2002) Agglomeration Effects in Europe. *European Economic Review,* 46**,** 213-227.

CICCONE, A. & HALL, R. E. (1996) Productivity and the Density of Economic Activity. *American Economic Review,* 86**,** 54-70.

COMBES, P.-P., DURANTON, G. & GOBILLON, L. (2008a) Spatial Wage Disparities: Sorting Matters! *Journal of Urban Economics,* 63**,** 723-742.

COMBES, P.-P., DURANTON, G., GOBILLON, L. & ROUX, S. (2007) Estimating Agglomeration Effects: Does Playing with Different Instruments Give a Consistent Tune? *The NBER Reporter 2007 Number 4: Conferences*.

COMBES, P.-P., DURANTON, G., GOBILLON, L. & ROUX, S. (2008b) Estimating Agglomeration Economies with History, Geology, and Worker Effects. CEPR Discussion Papers 6728.

CORNELISSEN, T. (2008) The Stata Command Felsdvreg to Fit a Linear Model With Two High-Dimensional Fixed Effects. *The Stata Journal,* 8 (2).

COSTA, D. D. L. & KAHN, M. E. (2001) Power Couples: Changes in the Locational Choice of the College Educated, 1940-1990. *The Quarterly Journal of Economics,* 115**,** 1287.

CRAGG, J. G. & DONALD, S. G. (1993) Testing Identifiability and Specification in Instrumental Variable Models *Econometric Theory,* 9 (2)**,** 222-240.

DE LUCIO, J. J., HERCE, J. A. & GOICOLEA, A. (1996) Externalities and Industrial Growth: Spain 1978-1992. *FEDEA-Fundacion de Estudios de Economia Aplicada*.

DENT, A. & BOND, S. (2008) An Investigation into the Location and Commuting Patterns of Part-Time and Full-Time Workers in the United Kingdom, using Information from the 2001 Census. Office for National Statistics.

DI ADDARIO, S. & PATACCHINI, E. (2007) Wages and the City. Evidence from Italy. *Labour Economics,* In Press, Corrected Proof.

DIAMOND, C. A. & SIMON, J. C. (1990) Industrial Specialization and the Returns to Labor. *Journal of Labor Economics,* 8**,** 175.

DURANTON, G. & PUGA , D. (2000) Diversity and Specialisation in Cities: Why, Where and When Does it Matter? *Urban Studies,* 37**,** 533.

FINGLETON, B. (2003) Increasing Returns: Evidence from Local Wage Rates in Great Britain. *Oxford Economic Papers,* 55**,** 716.

FINGLETON, B. (2006) The New Economic Geography versus Urban Economics: An Evaluation Using Local Wage Rates in Great Britain. *Oxford Econonomic Papers,* 58**,** 501-530.

FOGARTY, M. S. & GAROFALO, G. A. (1978) An Exploration of the Real Productivity Effects of Cities. *The Review of Regional Studies,* 8**,** 65-82.

FORNI, M. & PABA, S. (2002) Spillovers and the Growth of Local Industries. *The Journal of Industrial Economics,* 50**,** 151.

FU, S. (2007) Smart Café Cities: Testing Human Capital Externalities in the Boston Metropolitan Area. *Journal of Urban Economics,* 61**,** 86–111.

FUJITA, M. M. & THISSE, J.-F. (2002) *Economics of Agglomeration-Cities, Industrial Location and Regional Growth*, Cambridge University.

GLAESER, E. L. & MARE, D. C. (2001) Cities and Skills. *Journal of Labor Economics,* 19**,** 316.

GRAHAM, D. J. (2000) Spatial Variation in Labour Productivity in British Manufacturing. *International Review of Applied Economics,* 14**,** 323.

GRAHAM, D. J., GIBBONS, S. & MARTIN, R. (2009a) Transport Investments and the Distance Decay of Agglomeration Benefits. Working Paper, Imperial College of London.

GRAHAM, D. J. & KIM, H. Y. (2007) An Empirical Analytical Framework for Agglomeration Economies. *Annals of Regional Science,* (in press).

GRAHAM, D. J., MELO, P., JIWATTANAKULPAISARN, P. & NOLAND, R. B. (2009b) Testing for Causality between Productivity and Agglomeration Economies. *Journal of Regional Science,* (in press).

HAHN, J. & HAUSMAN, J. A. (2003) Weak Instruments: Diagnosis and Cures in Empirical Econometrics. *The American Economic Review,* 93**,** 118.

HANSEN, L. P. (1982) Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica,* 50**,** 1029-1054.

HARRIS, C. D. (1954) The Market as a Factor in the Localization of Industry in the United States. *Annals of the Association of American Geographers,* 44**,** 315-348.

HAUSMAN, J. A. (1978) Specification Tests in Econometrics. *Econometrica,* 46**,** 1251-1271.

HAUSMAN, J. A. & TAYLOR, W. E. (1981) Panel Data and Unobservable Individual Effects. *Econometrica* 49**,** 1377-1398.

HELSLEY, R. W. & STRANGE, W. C. (1990) Matching and Agglomeration Economies in a System of Cities. *Regional Science and Urban Economics,* 20**,** 189-212.

KIM, S. (1989) Labor Specialization and the Extent of the Market. *The Journal of Political Economy,* 97**,** 692.

KLEIBERGEN, F. & PAAP, R. (2006) Generalized Reduced Rank Tests using the Singular Value Decomposition. *Journal of Econometrics,* 133**,** 97-126.

LEWIS, W. C. & PRESCOTT, J. R. (1974) Manufacturing Employment, City Size, and Inter-Urban Wage Differentials. *The Annals of Regional Science,* 8(3).

MELO, P. C., GRAHAM, D. J. & NOLAND, R. B. (2009) A Meta-Analysis of Estimates of Urban Agglomeration Economies. *Regional Science and Urban Economics,* 39**,** 332-342.

MION, G. & NATICCHIONI, P. (2005) Urbanization Externalities, Market Potential and Spatial Sorting of Skills and Firms. *CEPR Discussion Papers 5172. Centre for Economic Performance, London School of Economics and Political Science*.

MION, G. & NATICCHIONI, P. (2009) The Spatial Sorting and Matching of Skills and Firms. *Canadian Journal of Economics,* 42(1)**,** 28-55(28)

MOOMAW, R. L. (1981) Productivity and City Size: A Critique of the Evidence. *The Quarterly Journal of Economics,* 96**,** 675.

MORETTI, E. (2004) Human Capital Externalities in Cities. *Handbook of Regional and Urban Economics. Cities and Geography.* Amsterdam;London, Elsevier North Holland.

MURRAY, M. P. (2006) Avoiding Invalid Instruments and Coping with Weak Instruments. *The Journal of Economic Perspectives,* 20**,** 111.

RICE, P., VENABLES, A. J. & PATACCHINI, E. (2006) Spatial Determinants of Productivity: Analysis for the Regions of Great Britain. *Regional Science and Urban Economics,* 36**,** 727-752.

ROSENTHAL, S. S. & STRANGE, C. W. (2001) The Determinants of Agglomeration. *Journal of Urban Economics,* 50**,** 191.

ROSENTHAL, S. S. & STRANGE, W. C. (2004) Evidence on the Nature and Sources of Agglomeration Economies. IN HENDERSON, J. V. & THISSE, J. F. (Eds.) *Handbook of Urban and Regional Economics, Vol.4.* Amsterdam, Elsevier.

ROSENTHAL, S. S. & STRANGE, W. C. (2006) The Micro-Empirics of Agglomeration Economies. IN ARNOTT, R. & MCMILLEN, D. (Eds.) *A Companion to Urban Economics.* Blackwell Publishing.

ROSENTHAL, S. S. & STRANGE, W. C. (2008) The Attenuation of Human Capital Spillovers. *Journal of Urban Economics,* In Press, Corrected Proof.

ROUWENDAL, J. (1999) Spatial Job Search and Commuting Distances. *Regional Science and Urban Economics,* 29**,** 491-517.

SARGAN, J. D. (1958) The Estimation of Economic Relationships using Instrumental Variables. *Econometrica,* 26**,** 393-415.

SCHAFFER, M. E. (2007) xtivreg2: Stata Module to Perform Extended IV/2SLS, GMM and AC/HAC, LIML and k-class Regression for Panel Data Models. .

SEGAL, D. (1976) Are There Returns to Scale in City Size? *The Review of Economics and Statistics,* 58**,** 339-350.

SHEA, J. (1997) Instrument Relevance in Multivariate Linear Models: A Simple Measure. *The Review of Economics and Statistics,* 79**,** 348.

STAIGER, D. & STOCK, J. H. (1997) Instrumental Variables Regression with Weak Instruments. *Econometrica,* 65**,** 557.

STOCK, J. H. & WRIGHT, J. H. (2000) GMM with Weak Identification. *Econometrica,* 68**,** 1055–1096.

STOCK, J. H. & YOGO, M. (2005) Testing for Weak Instruments in Linear IV Regression. IN STOCK, D. W. K. A. A. J. H. (Ed.) *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg.* Cambridge: Cambridge University Press.

SVEIKAUSKAS, L. (1975) The Productivity of Cities. *The Quarterly Journal of Economics,* 89**,** 393.

TABUCHI, T. & YOSHIDA, A. (2000) Separating Urban Agglomeration Economies in Consumption and Production. *Journal of Urban Economics,* 48**,** 70.

WHEELER, C. H. (2001) Search, Sorting, and Urban Agglomeration. *Journal of Labor Economics,* 19**,** 879.

YANKOW, J. J. (2006) Why Do Cities Pay More? An Empirical Examination of some Competing Theories of the Urban Wage Premium. *Journal of Urban Economics,* 60**,** 139-161.

BERR
Department for Business
Enterprise & Regulatory Reform

Communities
and Local Government

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

Llywodraeth Cynulliad Cymru
Welsh Assembly Government

**Spatial Economics Research Centre (SERC)**
London School of Economics
Houghton Street
London WC2A 2AE

**Tel:** 020 7852 3565
**Fax:** 020 7955 6848
**Web:** www.spatialeconomics.ac.uk

SERC is an independent research centre funded by the Economic and Social Research Council (ESRC), Department for Business, Enterprise and Regulatory Reform (BERR), the Department for Communities and Local Government (CLG) and the Welsh Assembly Government.