View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

Sectoral Agglomeration Economies in a Panel of European Regions*

Marius Brülhart†
University of Lausanne

Nicole A. Mathys‡
University of Lausanne

July 2007

Abstract

We estimate agglomeration economies, defined as the effect of density on labour productivity in European regions. The analysis of Ciccone (2002) is extended in two main ways. First, we use dynamic panel estimation techniques (system GMM), thus offering an alternative methodological treatment of the inherent endogeneity problem. Second, the sector dimension in the data allows for disaggregated estimation. Our results confirm the presence of significant agglomeration effects at the aggregate level, with an estimated long-run elasticity of 13 percent. Repeated cross-section regressions suggest that the strength of agglomeration effects has increased over time. At the sector level, the dominant pattern is of cross-sector "urbanisation" economies and own-sector congestion diseconomies. A notable exception is financial services, for which we find strong positive productivity effects from own-sector density.

JEL Classification: R10

Keywords: Employment density, Productivity, European regions, Dynamic panel GMM

*We thank Antonio Ciccone, Florian Pelgrin, and participants at the ETSG 2006 annual conference in Vienna, at the Kiel Institute's 2007 Workshop on Agglomeration and Growth, and at the 2007 Workshop on Economic Geography organised by the University of Bologna at Rimini for helpful comments; and Andrés Rodriquez-Pose and Steve Redding for the generous provision of data on regional education levels. Financial support from the Swiss National Science Foundation (grants 100012-113938/1 and 612-65970) and from the EU's 6th Framework Programme ("Micro-Dyn" project) is gratefully acknowledged.

†DEEP-HEC, University of Lausanne, Switzerland; e-mail: `marius.brulhart@unil.ch`; also affiliated with the Centre for Economic Policy Research.

‡DEEP-HEC, University of Lausanne, Switzerland; e-mail: `nicole.mathys@unil.ch`.

1 Introduction

Spatial agglomeration economies have been in intellectual fashion at least since the early 1990s. Taking their cue from Michael Porter, business leaders and policy makers continue to pursue cluster-formation strategies in a quest for performance enhancing localised synergies. Meanwhile, economic researchers have paid considerable attention to Krugman's "new economic geography" models, and developed them into a still flourishing subfield at the intersection of international, regional and urban economics.

Heightened attention to the economic benefits of co-location has stimulated much careful empirical work. The data have been found by most researchers to confirm the received wisdom: agglomeration enhances economic performance. This is true both in terms of *localisation* economies, i.e. the benefits to be derived from firms of the same sector locating in each others' proximity; and in terms of *urbanisation* economies, i.e. the benefits to be derived from firms locating in the proximity of a wide array of other, not necessarily related, firms.[1]

Researchers trying to verify empirically whether agglomeration enhances performance inevitably face the major difficulty that causality could run both ways. If a particular location offers some inherent features that improve the profitability of certain economic activities, firms will be attracted to that location. Such inherent features may be related to natural endowments or regulatory specificities, but they could also have to do with essentially unmeasurable factors such as local business cultures. How to isolate the effect that runs from agglomeration to performance (rather than the other way) thus represents a considerable challenge to the empirical researcher.

We address this issue via dynamic panel GMM estimation, using sectoral data for European regions. Dynamic panel GMM techniques have been developed precisely for the purpose at hand, that is to offer instrumental-variables estimates in settings featuring endogenous regressors, based on relatively weak assumptions on the underlying data-generating process. By drawing on sectorally disaggregated data, we can distinguish productivity effects of own-sector density from the effects of aggregate economic density.

Our results confirm the existence of positive net productivity effects of employment density, with an estimated elasticity of some 13 percent. Repeated cross-section estimation suggests that the intensity of such agglomeration effects has been increasing over our sample period 1980-2003. At the sector level, cross-sector "urbanisation" effects are generally positive. With the notable exception of financial services, own-sector "localisation" effects, however, are mostly negative, suggesting dominance of congestion diseconomies.

The paper is organised as follows. We summarize the related literature in Section 2. Section 3 sets out the empirical model and describes our approach to estimation. The results are reported in Section 4. Section 5 concludes.

[1] Ciccone and Hall (1996) and Ciccone (2002) are the seminal papers of the recent research programme that focuses on region-level productivity effects of density. Rosenthal and Strange (2004) offer an early survey. We provide a summary of this literature in Section 2 below.

2 Literature Background

This paper is essentially an extension of the work by Ciccone (2002) and Ciccone and Hall (1996). Ciccone (2002) estimates the impact of employment density on labour productivity in a cross section of NUTS-3 regions in France, Germany, Italy, Spain and the UK. Instrumenting density by regional area, he finds that a doubling of employment density increases labour productivity by some 4.5 percent - which is close to the corresponding elasticity of 6 percent estimated for the United States in Ciccone and Hall (1996).

There are a number of related papers, which together present a quite coherent body of evidence. Sveikauskas (1975), based on US data for 1967, reports that a doubling of city size raises labour productivity by 6.0 percent. Based on a panel of US plant-level data, Henderson (2003) finds that high-tech manufacturing plants are significantly more productive the more other plants of their sector are located in the same county. The estimated elasticity ranges from 1.2 to 13.5 percent. Dekle and Eaton (1999) find significant productivity effects of own-sector density in data on Japanese regions, although their estimated elasticities are a mere 1.0 percent for manufacturing and 1.2 percent for services. Comparable studies also exist for some individual European countries. Cingano and Schivardi (2004) draw on a panel of plant-level data across Italian cities and estimate a long-run elasticity of plant productivity to city employment of 6.7 percent. Rice, Venables and Patacchini (2006) estimate the effect on regional labour productivity of the proximity to economic mass across British regions, controlling for occupational composition. They obtain an elasticity of 3.5 percent. Ottaviano and Pinelli (2006), using panel data for Finnish regions, report that population density has a positive effect on regional growth of population, income and house prices, which implies a positive impact on productivity.

No paper has as yet used panel data across several European countries for an assessment of density effects on productivity. Furthermore, the literature highlights five particular issues that we address in this paper: endogeneity, the empirical modelling of agglomeration, measurement, sectoral disaggregation, and dynamics.

Endogeneity: When regressing regional productivity on a measure of regional agglomeration, one is forced to take account of the possibility that causality runs from the former to the latter.[2] Ciccone's (2002) strategy is to instrument density with regional land area. This instrument is valid in so far as the spatial distribution of regional populations in the 19th century - when most of the modern regional boundaries were drawn, largely in a way as to equalise regional populations - is uncorrelated with current productivity. In turn, any unobserved long-run persistent features that boost both regional populations and regional productivity will induce upward bias in the estimated coefficient on density. Panel estimation techniques allow us to avoid this problem by differencing out time-invariant regional effects. Specifically, this is the first study to apply the

[2] See Hanson (2001) for a discussion of this and related issues affecting estimation of agglomeration effects.

system GMM estimator developed by Arellano and Bover (1995) to the analysis of agglomeration effects.

Empirical modelling of agglomeration: Agglomeration is a rather loosely defined term. While theorists generally understand agglomeration as a process that leads to the spatial concentration of economic activity, empirical researchers often equate it with the outcome of such a process, i.e. as synonymous with spatial concentration. We employ the latter use of the term and follow the literature in taking regional employment density as the primary metric for agglomeration. However, the productivity of a region with a given density is likely to depend also on the size of that region. Hence, we follow Ciccone and Hall (1996) and Harris and Ioannides (2005) in controlling also for region size, we consider effects from the density of neighbouring regions, as in Ciccone (2002), and we control for regional market potential.

Measurement: Many related studies rely on productivity and density measures pertaining exclusively to the manufacturing sector. As discussed by Ciccone and Hall (1996, p. 60), this is likely to induce estimation bias, as denser areas (in terms of manufacturing presence) are also likely to offer better access to service inputs, such that the direct effect of manufacturing density on manufacturing productivity will be overestimated. This mitigates in favour of broad measures that account for all market-based activities. In addition, Cingano and Schivardi (2004) convincingly argue that estimation of the sources of regional growth should be based on direct measures of productivity rather than drawing on employment growth as a proxy measure for productivity changes. For these reasons, we use a dataset that reports gross value added (GVA) per worker across the full range of economic sectors.

Sectoral disaggregation: One of the most fundamental distinctions in the economics of agglomeration is between own-sector effects (also referred to as localisation economies or Marshall externalities) and cross-sector effects (also referred to as urbanisation economies or Jacobs externalities). Numerous empirical studies have addressed this issue before us. Most of these analyses are based on data for cities. By drawing on sub-national data for EU sectors, we can now explore this issue in European regional data.

Dynamics: Evidence on the dynamics of agglomeration effects remains relatively scarce.[3] Henderson (1997, 2003) estimates dynamic equations determining firm-level productivity in the United States, finding that agglomeration effects operate with significant lags. His analyses, however, are confined to a subset of manufacturing industries. Combes, Magnac and Robin (2004) and Blien, Südekum and Wolf (2006) report employment growth regressions using disaggregated regional data for France and Germany respectively. We add to this literature by estimating dynamic models of GVA at the sector-region level across multiple European countries.

[3] Hanson (2001, p. 271) identifies this as "clearly an area where more work is needed".

3 The Empirical Model

3.1 Dynamic Modelling of Agglomeration Effects

Ciccone (2002) derives a static log-linear model of regional labour productivity, based on a Cobb-Douglas production function in physical capital, labour and human capital and augmented by an exogenous term representing externalities from spatial output density. Assuming perfect capital-market integration within countries, he arrives at the following specification for log output per worker, P, in region d of country c:

$$P_{dc} = \theta D_{dc} + \sum_{s=1}^{S} \delta_{sc} F_{sdc} + \epsilon_c + u_{dc}, \quad (1)$$

where D_{dc} represents log employment density, F_{sdc} is the fraction of workers with education level s, ϵ_c is a country-specific constant term, u_{dc} is a stochastic residual capturing unexplained intra-country differences in labour productivity, and θ and δ_{sc} are coefficients to be estimated. The coefficient θ represents the net agglomeration effect (agglomeration minus congestion).

The main challenge for estimation of this model arises through the potential endogeneity of density. If there exist unmeasured region-specific features that impact positively (negatively) on regional productivity, this will attract (deter) workers and thus impact on density. Potential endogeneity of D_{dc} thus raises the possibility that OLS estimates of θ are upward biased. Ciccone's (2002) solution consists in instrumenting regional density with regional area, arguing that, while area is significantly correlated with density, regional boundaries have been set sufficiently long ago (between 1789 and 1888) as not to be correlated with any unmeasured region-specific productivity idiosyncrasies in the 1980s.

A central aim of our paper is to take advantage of the fact that we now avail of panel data on output and employment in European regions. This allows us to employ alternative strategies to identify the direct effect of density on productivity. Specifically, we employ GMM estimators for dynamic panels, which rest on relatively weak assumptions to validate instruments that can be found in dynamic transformations of the model's variables.

Mathys (2007) proposes a simple dynamic version of Ciccone's (2002) static model. Adding a Cobb-Douglas capital accumulation function to the supply-side of the model, combined with an infinitely-lived representative consumer with a constant intertemporal elasticity of substitution, leads to a dynamic productivity equation. Specifically, productivity depends on lagged productivity, present and lagged density and other control variables. This implies a first-order *a*utoregressive *d*istributed *l*ag ADL(1,1) panel model, which serves as our starting point:

$$P_{dt} = \alpha P_{d,t-1} + \beta_0 D_{dt} + \beta_1 D_{d,t-1} + \boldsymbol{\gamma}_0' \mathbf{X}_{dt} + \boldsymbol{\gamma}_1' \mathbf{X}_{d,t-1} + \epsilon_d + \varrho_t + \nu_{dt}, \quad (2)$$

where $\mathbf{X}_{dt}$ is a column vector of $k \in \{1...K\}$ control variables; α, β, and $\boldsymbol{\gamma}$ are coefficients to be estimated ($\boldsymbol{\gamma}'$ being a vector of dimension $1 \times K$); ϵ_d is a region-specific effect; ϱ_t is a period-specific effect; and ν_{dt} is a stochastic error term.[4] $\mathbf{X}_{dt}$ could contain variables representing the availability of human and physical capital in region d, or variables reflecting the time-varying component of the regional business climate or of the political environment. We shall allow for the possibility that these variables are measured with error.

The key component of this specification is ϵ_d, which represents time-invariant determinants of regional labour productivity that may or may not be correlated with density. If such effects exist and are important, any cross-section estimate of θ based on time-lagged instruments is bound to be biased. Perhaps the principal attraction of panel estimation is that it allows us to "remove" the effect of ϵ_d.

We seek to exploit the time-series dimension in the data, but theory offers little guidance on the dynamics of agglomeration economies. A considerable advantage of specification (2) therefore lies in its generality. The ADL(1,1) model nests the most widely used dynamic processes. For example, it can represent a "common factor" model with contemporaneous measured effects and autocorrelated errors. This would imply that $\beta_1 = -\frac{\beta_0}{\alpha}$ and $\gamma_1^k = -\frac{\gamma_0^k}{\alpha}$, restrictions that we will test. According to this model, the impact of changes in density on regional productivity fully materialises within the time period t, but there are persistent shocks to the stochastic component of regional productivity. In addition, (2) also nests the ADL(1,0) model, implying that $\beta_1 = \gamma_1^k = 0$. The ADL(1,0) specification in turn can derive from a number of theoretical bases, the most relevant of which is the "geometric lag" or "partial adjustment" model. In that model, the dependent variable responds sluggishly to changes in the explanatory variables, with geometrically declining lag weights. In our context this represents delayed responses in productivity to changes in density, for example because technological agglomeration externalities take time to materialise (or to unravel) as they depend on the formation of communication networks.[5]

Our main focus, however, is not on short-run dynamics, as we primarily seek to identify a long-run equilibrium relationship between density and productivity. In (2), this relationship is given by $\beta_{LR} = \frac{\beta_0+\beta_1}{1-\alpha}$. We compute this coefficient and test its statistical significance using the delta method. β_{LR} is a nonlinear combination of the estimated parameters and therefore risks being a noisy measure, especially if α is close to unity and imprecisely estimated. We therefore also test the linear restriction that $\beta_0 + \beta_1 = 0$ via the standard Wald test. If this restriction is rejected, we conclude that density has a statistically significant long-run effect on regional productivity. Conversely, if the restriction is not rejected but the parameters are individually statistically significantly different from zero, the interpretation is that changes in density have short-run effects on regional productivity without impacting on the long-run productivity

[4] The first "1" in "ADL(1,1)" stands for one autoregressive term, i.e. inclusion of $P_{d,t-1}$ on the right-hand side; while the second "1" stands for the inclusion of one lagged value for each of the regressors, e.g. $D_{d,t-1}$.

[5] For an exposition of common factor and partial adjustment models, see e.g. Davidson and MacKinnon (2004), chapters 7 and 13 respectively. Blundell and Blond (1998) apply the common factor model to the estimation of a production function.

level.

3.2 Estimation

The most widely used estimator for models such as (2) is Arellano and Bond's (1991) "DIFF-GMM". This estimator is based on first-differenced variables, thus eliminating the region-specific effects (ϵ_d), and instrumenting all potentially endogenous variables ($P_{d,t-1}, D_{dt}, D_{d,t-1},\ X^k_{dt},\ X^k_{d,t-1}$) with their own suitably lagged levels. This estimator relies on the (weak) assumption that the initial conditions are predetermined, i.e. that $E\left[P_{d1}\nu_{dt}\right] = E\left[D_{d1}\nu_{dt}\right] = E\left[X^k_{d1}\nu_{dt}\right] = 0,\ \forall k,\ t = 2, ..., T$. It is consistent as $N \longrightarrow \infty$, given T. Hence, this approach offers a solution to the challenge of instrumenting for persistent region-specific effects correlated with density.

However, the DIFF-GMM estimator has been found to behave poorly in small samples when α approaches unity and/or when the variance of ϵ_d is large compared to the variance of ν_{dt} (Blundell and Bond, 1998). Given the slow nature of changes in the spatial structure of the economy, these configurations would seem relevant in our context. In such conditions, lagged levels likely represent weak instruments for the first-differenced variables, and DIFF-GMM estimates are prone to suffer from finite-sample bias.

A related GMM dynamic panel estimator, "SYS-GMM", initially proposed by Arellano and Bover (1995), is shown by Blundell and Bond (1998) to yield potentially considerable improvements over DIFF-GMM in small samples. SYS-GMM is based on a system composed of first-differences instrumented on lagged levels, and of levels instrumented on lagged first-differences. For SYS-GMM to be valid, the following additional assumption needs to hold:

$$E\left[\Delta P_{d2}\epsilon_d\right] = E\left[\Delta D_{dt}\epsilon_d\right] = E\left[\Delta X^k_{dt}\epsilon_d\right] = 0. \tag{3}$$

The main aim of our estimation strategy is to isolate the causal effect that runs directly from density to productivity. Hence, it is important that we scrutinise the assumption underlying SYS-GMM carefully in terms of our empirical context.

Note that, as far as the first-differenced elements of the regressor and instrument matrices are concerned, the effect of area cancels out, and density becomes perfectly collinear with employment. However, since levels of the right-hand-side variables enter the estimation in both DIFF-GMM and SYS-GMM, we are able to identify the effect of regional density separately from the effect of regional employment.

A *sufficient* condition for (3) to hold is that P_{dt}, D_{dt} and X^k_{dt} be mean stationary. This might be considered a strong assumption for our study, in particular as it would rule out, without theoretical or empirical justification, the possibility of secular trends in productivity, density or other relevant variables. However, the inclusion of the time effects ϱ_t in the empirical model (2) allows for a common trend in productivity without violating (3). Furthermore, by transforming all variables into deviations from their respective country-year means, our setup also allows for idiosyncratic secular productivity trends at the country level. The stationarity assumption therefore reduces to positing that there are no secular diverging trends of relative productivity levels among regions of a given country. The assumption of mean stationarity conditional on country-year means is supported by evidence in Combes and Overman (2004, Figure 2), which shows that the within-country dispersion of regional per-capita GDPs in the EU has remained remarkably stable between 1982 and 1996 (while between-country dispersion changed substantially). Yet, the weaker stationarity assumption might still appear constraining. We therefore note that assumption (3) does not *necessarily* require mean stationarity. As shown by Blundell and Bond (1998), (3) merely implies that the initial-period-specific disturbance ν_{d1} be uncorrelated with ϵ_d. Either the underlying process has been generating the regional productivity series for long enough prior to the observed period, so that the true initial-period conditions become negligible, or the initial-period disturbances are randomly distributed across regions. In both these cases, conditional mean-stationarity of the explanatory variables is not required to satisfy (3).[6]

In addition to providing a rigorous palliative for endogeneity bias, dynamic panel GMM estimation holds two further attractions. First, it is more robust to measurement error than cross-section regressions. Time-invariant additive measurement error is absorbed into region-specific effects, and through suitably long lags delineating the instrument sets for P_{dt} (in first-differences and in levels), dynamic panel GMM remains consistent even in the presence of region-year specific (but serially uncorrelated) measurement error. Second, dynamic panel GMM remains consistent even if density (as well as other controls) is endogenous in the sense that $E\left[D_{dt}\nu_{ds}\right] \neq 0$ for $s \leq t$, if the instrumental variables are sufficiently lagged. All explanatory variables are treated as potentially endogenous (i.e. instruments are lagged at least two periods), and we allow for temporary measurement error in the dependent variable (i.e. instruments for $P_{d,t-1}$ are lagged at least three periods).

In order to maximise efficiency, we employ the two-step estimator, and we correct the standard errors for the small-sample bias of the two-step estimator applying the correction suggested by Windmeijer (2005). The maximum number of lags of the instrument sets is constrained in some specifications so as to avoid overfitting.[7]

Bun and Windmeijer (2007) show that, due to weak instruments, the SYS-GMM estimator, just like DIFF-GMM, may suffer from small-sample bias, although they confirm that SYS-GMM generally suffers from smaller bias than DIFF-GMM. Their Monte-Carlo experiments suggest that the bias for SYS-GMM coefficient estimates is particularly small when the variance of ϵ_d is equal to the variance of ν_{dt}. Inspection of our regression residuals shows that the between-variance is indeed of roughly similar magnitude to the within-variance. This offers a degree of confidence in our results despite the lack of a formal

[6] On the validity of the SYS-GMM assumption in a similar empirical setting (cross-country growth regressions), see Bond, Hoe- er and Temple (2001).

[7] All GMM estimations are carried out using David Roodman's xtabond2 command for Stata.

weak-instrument test for SYS-GMM estimation.

4 Data

We use a panel of European regional data from Cambridge Econometrics (CE).[8] This data set covers up to 245 NUTS-2 regions of 20 Western and Eastern European countries, spanning the period 1980-2003 (1990-2003 for Eastern European countries).[9] Employment and gross value added are reported for ten broad sectors: Agriculture, Construction, Manufacturing and Energy, Wholesale and Retail Services, Hotels and Restaurants, Transport and Communication Services, Financial Services, Other Market Services, and Non-Market Services. Following Ciccone (2002), we do not estimate effects on productivity in Agriculture and Non-Market Services, as the market-based model underlying our research evidently is not applicable to these sectors. In line with previous empirical work, our main focus will be on Manufacturing and on Financial Services, as these appear *a priori* as the sectors that are least dependent on the location of exogenous endowments and thus most susceptible to the agglomeration forces identified in economic geography models.

Summary statistics are provided in Table 1. Our sample regions on average contain some 500,000 employed residents and cover 16,000 square kilometres. It is not surprising that the dispersion across regions is much higher for employment density than for labour productivity (the coefficients of variation being 3.36 and 0.44 respectively). This reflects, on the one hand, the spatially concentrated nature of human settlement and economic activity, which implies large differences in regional densities, and, on the other hand, market forces that push towards equalisation of labour productivity via factor mobility and technology diffusion.

It is striking, though, that in spite of the relatively high degree of economic integration among European regions, considerable differences in labour productivity persist. Workers in the most productive sample region (Groningen, Netherlands) were more than 12 times as productive as workers in the least productive region (Eastern Slovakia). Cross-country comparisons of labour productivity may be affected by differences in national statistical practices and by the vagaries of currency conversion. Table 1 therefore summarises intra-national productivity differentials in the six largest sample countries. We observe substantial productivity differentials even within countries, the ratio of productivities between the most and the least productive region ranging from 1.3 (Poland) to 1.8 (UK).

[8] The CE data base is based on Eurostat's Regio data and then completed using national sources of information as well as some interpolations using constraints coming from aggregate nation, region and sector shares. For the variables used in this paper, bilateral correlations between published Regio data and the corresponding CE values are never below 0.96.

[9] The 20 sample countries are: Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Netherland, Norway, Poland, Portugal, Slovakia, Spain, Sweden, Switzerland, United Kingdom. Island regions are excluded from the sample.

Matching panel data on human or physical capital do not exist. Our main results therefore rely on the assumption that period-region idiosyncratic changes in density are uncorrelated with contemporaneous period-region idiosyncratic changes in human capital and prices. We can, however, draw on a cross section of human-capital data, based on the 1994 European Community Household Panel. This provides us with measures of the percentage of residents with different maximum levels of completed education (primary, secondary and tertiary). Table 1 shows that, in the average sample region, 45 percent of the population have completed their secondary education while 17 percent have attained a tertiary qualification. This variable too exhibits considerable variation across Europe. For example, the share of residents with tertiary education is 39 times higher in Brussels than in the Algarve.

All our panel variables are reported annually. In order to purge our estimation from short-run cyclical effects, we run all regressions on three-year averages of our data. This implies that we observe up to eight time periods, starting with 1980-82 and ending in 2001-03.

5 Results

5.1 Dynamic Panel Estimation: Aggregate Effects

Table 2 reports panel regressions for aggregate labour productivity. We first report pooled OLS and simple fixed-effects panel estimates in columns (a) and (b). As expected, regional productivity is a highly persistent variable, with estimated α lying between 0.73 (fixed effects) and 0.97 (OLS).[10] The two estimators imply diametrically opposite long-run effects of density: while the estimated elasticity is 11 percent with OLS, it turns strongly negative, to -57 percent, once fixed effects are included. It is well known, however, that both pooled OLS and fixed-effects estimation of an autoregressive panel model will be subject to serious biases in the estimation of all model parameters. Specifically, the OLS estimate of α will be biased upwards, while the corresponding fixed-effects estimate will be biased downwards. In large samples and given some weak assumptions, GMM estimates are free of such bias.

We first apply the DIFF-GMM estimator (column (c)). The DIFF-GMM results suggest that density has a very large and *negative* long-run effect on regional labour productivity, the estimated elasticity being -80 percent. This is a stark result that runs counter to virtually all existing comparable evidence. Should we believe it? An obvious cause for caution is the fact that this estimate is very imprecisely measured and not statistically significantly smaller than zero. One must furthermore consider Blundell and Bond's (1998) result whereby DIFF-GMM can suffer from substantial bias in autoregressive panel data with a high degree of persistence. We indeed find that the estimated α in a DIFF-GMM specification, reported in column (c), is even lower than

[10] Note that in both cases we can confidently reject the hypothesis that $\alpha \geq 1$ and thus conclude that the dynamic model is stable.

the (downwards biased) fixed-effects estimate. Our results thus confirm prior small-sample findings based on highly persistent panel data, indicating that the DIFF-GMM estimator is unreliable in such contexts (see e.g. Bond *et al.*, 2001).

The SYS-GMM results of column (d) can *a priori* be considered superior to the DIFF-GMM estimates. It is reassuring, then, to find that the estimated α from SYS-GMM, reported in column (d) lies between the fixed-effects and the OLS estimates. Furthermore, the standard diagnostic tests suggest no misspecification problems.[11] The SYS-GMM estimate of the long-run density effect is statistically significantly *positive*. The DIFF-GMM results are thus overturned. Moreover, with a point estimate of 13 percent, we find agglomeration effects that are in fact even larger than prior estimates, although the latter lie well within the 95-percent confidence interval.[12]

In column (e), we control for regions' aggregate employment, as suggested by Harris and Ioannides (2005), who point out that proximity effects are determined by both the size of a region and its average density.[13] All diagnostics again look satisfactory. Our main result is not affected: regional density retains a statistically significantly positive effect, with an estimated elasticity of 13 percent. In addition, we find that the effect of region size is significantly positive and of very similar magnitude to the density effect. Our results based on aggregate data thus confirm the qualitative findings of prior studies. Previous estimates of agglomeration effects in European regions may in fact have been rather conservative.

Our specification also offers insight on the dynamics of density effects. In terms of the nature of intertemporal variation, some of our evidence supports a partial-adjustment model (where we cannot reject the restriction that the coefficients on the lagged independent variables are zero), but support is also found for a contemporaneous model with autocorrelated stochastic shocks (where we cannot reject the common-factor restrictions). Irrespective of the estimator chosen, however, the contemporaneous impact is always negative while the coefficient on lagged density is positive. In specifications that support the partial-adjustment model (e.g. column (e) of Table 2), this suggests, quite plausibly, that positive productivity effects of density take time to materialise, the delay being three years or more. In other words, the results imply that congestion effects appear faster than agglomeration effects, such that net positive productivity effects of increases in density only materialise after some four to six years.

[11] The Hansen test for overidentifying restrictions, as well as the Difference Hansen test for the validity of the instruments used in SYS-GMM in addition to those used for DIFF-GMM, fail to reject the null hypothesis that the instruments are valid. The AR2 test fails to reject the null hypothesis of no second-order residual autocorrelation.

[12] The 95-percent confidence interval for the estimated long-run density effect in the final specification of Table 2 (column e), for instance, is [0.012, 0.248].

[13] Additive size effects are also considered by Ciccone and Hall (1996) and Henderson (1997). Note that inclusion of employment in the DIFF-GMM specification would imply perfect multicollinearity of the structural equation, and identification would come solely from the differences in instrument sets. This is why we do not report DIFF-GMM estimates of the specification that includes size effects.

5.2 Dynamic Panel Estimation: Sector-Level Effects

5.2.1 Manufacturing and Financial Services

Our data allow us to estimate agglomeration effects at the sector level. The available sectoral classification, however, is rather crude. Of the ten sectors that are distinguished in the European regional data, only two look like strong candidates for agglomeration-induced productivity effects: Manufacturing and Financial Services. The remaining sectors are either strongly tied to exogenous endowments (Hotels and Restaurants, Transport and Communication Services), "derived" activities tied to the location of other sectors (Construction, Wholesale and Retail Services, Hotels and Restaurants, Transport and Communication Services), subject to strong government interference (Transport and Communication Services), or difficult to interpret (Other Market Services). Our main focus is therefore on the two sectors Manufacturing and Financial Services.[14] We now concentrate on SYS-GMM estimation, which, in our context, is superior to DIFF-GMM due to likely weak-instrument problems affecting the latter and based on our regression results for aggregate effects. Sector-level estimations allow us to estimate the impact of own-sector density (localisation effect) separately from that of the density of other sectors (urbanisation effect).

Table 3 reports the relevant regression results for different possible permutations of right-hand-side variables. For Manufacturing, only the full specification (column (a)) yields satisfactory diagnostic statistics, the parsimonious equations (columns (b)-(d)) all failing the AR(2) test. The results suggest that own-sector density affects manufacturing productivity *negatively*, with an implied long-run elasticity of up to -20 percent (although this coefficient is not statistically significant). Congestion effects therefore would seem to dominate over own-sector agglomeration economies. In contrast, the effects of other-sector density and of region size are both estimated to be positive, with elasticities of up to 17 and 11 percent respectively. When we control for regions' aggregate employment, the positive effect of other-sector density is statistically significant. European manufacturing thus seems to exhibit negative net localisation effects and positive net urbanisation effects.

The right-side half of Table 3 reports corresponding results for Financial Services. Here, the diagnostics indicate specification problems for none of the four regressions. The implied long-run elasticities are exactly the reverse to those found for Manufacturing. Own-sector density has significantly positive effects, with an elasticity of up to 170 percent - the precisely measured coefficients, however, suggesting a more reasonable range of 23 to 26 percent. Conversely, other-sector density and region size both reduce productivity, with implied elasticities of up to -155 and -17 percent respectively.[15] Financial services therefore appear to benefit from substantial localisation economies, but not from urbanisation economies.

[14] We will refer to Manufacturing and Energy as Manufacturing, as Energy only accounts for a small fraction of the sector in employment terms.

[15] Collinearity of own-sector density with other-sector density likely explains the large magnitude of estimated effects in the specifications that include both those regressors.

The story implied by these results is that manufacturing activities, while subject to internal dispersion forces, benefit from proximity to large markets; whereas financial services exhibit strong agglomeration forces while being less reliant on proximity to large markets in general. This simple pattern tallies with casual observation as well as with previous research. Based on Japanese data, Dekle and Eaton (1999) find that agglomeration economies are larger in financial services than in manufacturing, and that the spatial decay of agglomeration effects is considerably steeper in financial services than in manufacturing. Looking at employment growth in West Germany, Blien *et al.* (2006) conclude that localisation economies relatively stronger in financial services than in manufacturing, while urbanisation economies are stronger in manufacturing than in financial services. Using a very different dataset and estimation methodology, we find strong corroboration of these prior results.

In terms of dynamics, we find that the common factor restrictions are frequently rejected, suggesting that a partial-adjustment model may be more appropriate than a model where intertemporal correlation derives solely from autocorrelated random shocks. Contemporaneous density effects again are consistently negative, while the positive effects only appear in the lagged density variables. We interpret this as evidence that the short-run dominance of congestion effects over agglomeration effects holds also at the sector level.

5.2.2 Other Sectors

Although the relevance of agglomeration forces is less evident for the remaining market-oriented service sectors identified in our database, we report SYS-GMM estimates for them too, in Table 4. For three of the five sectors, second-order residual autocorrelation casts doubt on the reliability of the estimates, with only the Wholesale and Retail and Hotel and Restaurant sectors producing satisfactory diagnostics. For these sectors, however, we find no statistically significant evidence of long-run effects from density or region size.

Where we do find statistically significant own-sector density effects, they are negative (Construction, Transport and Communication, Other Market Services). Conversely, in the two cases of statistically significant cross-sector density effects (Transport and Communication, Other Market Services), these are positive. While we must treat these results with considerable caution, they do suggest that agglomeration effects tend to be stronger across sectors than within sectors, i.e. that urbanisation economies dominate localisation economies.

5.3 Robustness

We complement our baseline GMM estimations with three extensions. First, we consider the possibility that density effects “spill over” into neighbouring regions. Second, we explore the impact of reducing lag lengths. Finally, we run period-specific cross-section regressions to test for consistency with earlier work.

5.3.1 Neighbourhood and market potential effects

In Table 5, we replicate the aggregate panel regressions reported in Table 2 with the addition of a control for output density in each region's neighbours or each region's market potential. Specifically, we add a variable representing the average output density of all contiguous neighbouring regions in columns (a) to (e). In specifications (f) and (g), we replace neighbour density by market potential, defined as $M_{d^*} = \sum_d (GDP_d/DIST_{d^*d})$, where GDP is 1998 regional GDP in purchasing-power-parity terms, and $DIST_{d^*d}$ stands for the economic distance between regions d^* and d.[16]

In a nutshell, consideration of neighbourhood effects or market potential is of no substantive consequence to our results. The effects of neighbour-region density are generally statistically insignificant (except for the fixed-effects specification), and the estimated coefficients on own-region density and size remain qualitatively unchanged, with long-run SYS-GMM elasticities estimated as 16 percent in the preferred specification (column (e) of Table 5).[17] When introducing market potential instead of neigbour-region density, the long run density effect retains its magnitude, whereas market potential is not statistically significant.

These results are not surprising, as the NUTS-2 regions of our dataset are rather large, with a median area of 10,000 square kilometres. Ciccone (2002) found no significant effects of neighbour-region density even at the NUTS-3 level, with a median area of 1,511 square kilometres. Duranton and Overman (2005), based on UK plant-level data, provide evidence that localisation effects predominantly occur at a small spatial scale, below 50 kilometres. Our regions, therefore, are large enough for cross-region agglomeration economies to be unimportant (and certainly also too large to capture all effects from proximity).

5.3.2 Reducing lag lengths

In order explore the relevance of considered lag lengths, we restrict our ADL(1,1) baseline model to an ADL(1,0) version, which features no lagged regressors bar the lagged dependent variable. These results are reported in Table 6. Our results are strikingly similar to those found for the baseline model in Table 2. Most importantly, we find a long-run effect of employment density lying in the

[16] Drawing on the data set of Schürmann and Talaat (2000), economic distances are represented by estimated road-freight travel times between regional capitals. Due to somewhat smaller regional coverage (mainly for Eastern European countries), controlling for market potential shrinks our number of observations by some 27 percent. For details, see Brülhart (2006).

[17] Appendix Table 1 replicates the sector-level results of Table 3 with inclusion of neighbourhood effects. Here too, this extension does not affect our results qualitatively. It is interesting to note, however, that the long-run effect of other-sector density (urbanisation economies) in manufacturing, while still positive, becomes statistically insignificant once neighbourhood effects are included. The long-run own-region effects in financial services, however, retain statistical significance. These results are again consistent with the findings of Dekle and Eaton (1999), whereby the spatial decay of agglomeration economies is steeper in financial services than in manufacturing.

range between 0.12 and 0.15 (the latter estimate obtaining when we do not control for employment). The 95-confidence intervals of both these estimates again include Ciccone's (2002) baseline estimate of 0.045.[18]

5.3.3 Cross-section estimation

As a check on data reliability and comparability with earlier work, we estimate the relationship between regional density and regional productivity in cross sections of time-averaged data with country fixed effects. This approach also allows us to assess the change over time of agglomeration effects. Following Ciccone (2002), we estimate the model alternatively using OLS and two-stage least squares (2SLS) with regional land area as the instrument for density.

Table 7 reports OLS and 2SLS results based on the aggregated data for four of our seven sample periods:

- 1980-82, the first period in our sample,
- 1986-88, the period that coincides with the data used by Ciccone (2002),
- 1992-94, the period for which we have human capital data, and
- 2001-03, the last period in our sample.[19]

Data on Eastern European regions are included in all post-1990 periods, but for consistency of the series we also report results based solely on Western European regions for those periods. The first-stage F statistics on the identifying instrumental variables in the 2SLS estimations are all large enough to suggest that we do not have weak instrument problems. The human capital variables have the expected positive effects on labour productivity. Yet, their inclusion has virtually no impact on the estimated coefficients on regional density, which can be taken as an indication that the absence of this control in our baseline estimations may not be critical.[20]

We retain four major findings.

1. Agglomeration effects are increasing over time. The coefficient on employment density rises over time, both in the OLS and in the 2SLS regressions. If, for comparability, we consider only the results based on Western European regions, we find the following point estimates. With OLS, the estimated elasticity rises from 3.2 percent in 1980-82 to 5.7 percent in 2001-03; and with 2SLS, the elasticity rises from 0.9 percent to 4.5 percent.

[18] We also estimated ADL(1,2) and ADL (2,2) models and found our central findings to be substantively unaffected. These results can be obtained on request.

[19] Appendix Table 2 shows corresponding results for the initial and final sample periods for our two main sectors, Manufacturing and Energy and Financial Services. The qualitative results found at the aggregate level mostly carry over to the 2SLS sector-level regressions.

[20] To see this, compare the third and fourth results columns for 1992-94 in Table 6, which are based on the same region sample, once with and once without inclusion of educational controls.

2. Instrumenting for employment density always lowers the corresponding point estimate. This is consistent with the assumption that OLS estimates are biased upwards due to simultaneity problems.

3. Inclusion of Eastern European regions systematically raises the estimated density coefficient. Agglomeration effects therefore appear to be particularly strong in the EU's new member states. This may in part reflect a concentration of productive activities in those countries' capital regions, as a legacy of central planning (Brülhart and Koenig, 2006).

4. In spite of the difference in data coverage, our results turn out to be very similar to those obtained by Ciccone (2002). In 1986-88, Ciccone's time period for which, however, we have no human-capital data, our estimated density effect is some 4 percent. In 1992-94, when we control for education attainment, we find estimates ranging from 1.0 to 5.6 percent. When we exactly replicate Ciccone's (2002, Table 1) specification by estimating country-specific coefficients on the human-human capital variables, our estimated agglomeration effects are somewhat lower than those in Ciccone. For example, our 2SLS estimated elasticity of 1.0 percent compares to an effect of 4.6 percent estimated by Ciccone. The fact that our GMM results imply even stronger density effects than those found by Ciccone is therefore unlikely to be due to the difference in data sources and coverage.

6 Conclusions

Ciccone and Hall (1996) and Ciccone (2002) have discovered, for the United States and the European Union respectively, that the elasticity of regional productivity with respect to regional density can be estimated with remarkable robustness as around 5 percent. These estimates have found widespread acceptance as evidence in favour of positive localisation economies.

The major methodological challenge faced by this type of analysis is that causality could run from productivity to density. If the impact of such reverse causality were limited in time, existing instrumentation techniques would be sufficient. One cannot, however, *a priori* exclude the possibility of permanent locational features that favour productivity and thus density. We therefore re-examine the relationship between density and productivity using a panel of sector data for European regions, employing dynamic panel estimation methods to address the simultaneity problem inherent in this type of exercise. Panel estimation allows for the inclusion of region fixed effects, which capture all unobservable and time-invariant features that may drive both productivity and density. Panel data also allow us to track intertemporal patterns. Furthermore, we extend the analysis to the sector level, which allows us to evaluate the relative importance of own-sector and cross-sector density effects.

Despite the difference in data coverage and estimation technique, our results confirm previous findings. There are significant productivity-boosting effects of aggregate economic density. Our baseline estimated elasticity is 13 percent, i.e. considerably higher still than that estimated by Ciccone (2002). Ciccone's point estimate, however, lies within the 95-percent confidence interval of ours. We also find, based on repeated cross-section regressions, that the strength of these effects has increased over time. At sector level, cross-sector effects (urbanisation economies) are mostly positive, whereas own-sector effects (localisation economies) are mostly negative, suggestion congestion diseconomies. The exception is financial services, for which we find strong positive productivity effects from own-sector density. Positive agglomeration effects take time to materialise: our estimates suggest that subsequent to an increase in density congestion diseconomies dominate in the short run (meaning at least an initial 3-year interval), but that positive productivity effects appear in the long run.

The price to pay for broad coverage across Europe is that our data are rather aggregated in regional terms. There exists strong evidence that agglomeration economies operate at relatively small spatial scale, and we must assume that most small-scale proximity effects are imperfectly captured by our analysis. In that sense, even our results probably constitute rather conservative estimates of the full economic benefits of co-location.

References

[1] Arellano, M. and S. Bond, 1991, Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies*, 58, 277-297.

[2] Arellano, M. and O. Bover, 1995, Another look at the instrumental variable estimation of error-components models, *Journal of Econometrics*, 68, 29-51.

[3] Blien, U., J. Südekum and K. Wolf, 2006, Local Employment Growth in West Germany: A Dynamic Panel Approach, *Labour Economics*, 13(4), 445-458.

[4] Blundell, R. and S. Bond, 1998, Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, *Journal of Econometrics*, 87, 115-143.

[5] Blundell, R. and S. Bond, 2000, GMM Estimation with Persistent Panel Data: An Application to Production functions. *Econometric Reviews*, 19(3), 321-340.

[6] Bond, S., A. Hoe- er and J. Temple, 2001, GMM Estimation of Empirical Growth Models, *CEPR Discussion Paper* No. 3048.

[7] Bond, S., 2002, Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice, *IFS Working Paper* 09/02, Institute for Fiscal Studies, London.

[8] Brülhart, M., 2006, The Fading Attraction of Central Regions: an Empirical Note on Core-Periphery Gradients in Western Europe, *Spatial Economic Analysis*, 1(2): 227-235.

[9] Brülhart, M. and P. Koenig, 2006, New Economic Geography Meets Comecon: Regional Wages and Industry Location in Central Europe, *Economics of Transition*, 14(2): 245-267.

[10] Bun M. J. G. and F. Windmeijer (2007), "The Weak Instrument Problem of the System GMM Estimator in Dynamic Panel Data Models", mimeo, University of Bristol.

[11] Ciccone, A., 2002, Agglomeration Effects in Europe, *European Economic Review*, 46 (2), 213-227.

[12] Ciccone, A. and R. Hall, 1996, Productivity and the Density of Economic Activity, *American Economic Review*, 86 (1), 54-70.

[13] Cingano, F. and F. Schivardi, 2004, Identifying the Sources of Local Productivity Growth. *Journal of the European Economic Association*, 2(4), 720-742.

[14] Combes, P.P., T. Magnac and J. Robin, 2004, The Dynamics of Local Employment in France, *Journal of Urban Economics*, 56(2), 217-243.

[15] Combes, P.P. and H.G. Overman, 2004, The Spatial Distribution of Economic Activities in the European Union, in: Henderson, J.V. and J.F. Thisse (eds.) *Handbook of Regional and Urban Economics, Volume 4*, Elsevier North-Holland.

[16] Davidson, R. and J.G. MacKinnon (2004) *Econometric Theory and Methods*. Oxford University Press.

[17] Dekle, R. and J. Eaton, 1999, Agglomeration and Land Rents: Evidence from the Prefectures, *Journal of Urban Economics*, 46(2), 200-214.

[18] Duranton, G. and H.G. Overman, 2005, Testing for Localisation Using Micro-Geographic Data, *Review of Economic Studies*, 72(4), 1077-1106.

[19] Hanson, G.H., 2001, Scale Economies and the Geographic Concentration of Industry, *Journal of Economic Geography*, 1(3): 255-276.

[20] Harris, T.F. and Y.M. Ioannides, 2005, Productivity and Metropolitan Density, *The Economics of Density*, American Institute of Certified Planners.

[21] Henderson, V.J., 1997, Externalities and Industrial Development, *Journal of Urban Economics*, 42(3), 449-470.

[22] Henderson,V. J., 2003, Marshall's Scale Economies, *Journal of Urban Economics*, 53(1), 1-28.

[23] Marshall, A., 1920, Principles of economics, 8th Ed. London.

[24] Mathys, N.A., 2007, A Dynamic Model of Sectoral Agglomeration Effects, *Mimeo*, University of Lausanne (www.hec.unil.ch/nmathys/).

[25] Ottaviano, G.I.P. and D. Pinelli, 2006, Market Potential and Productivity: Evidence from Finnish Regions, *Regional Science and Urban Economics*, 36 (5), 636-657.

[26] Rice P., A.J. Venables and E. Patacchini, 2006, Spatial Determinants of Productivity: Analysis for the Regions of Great Britain. *Regional Science and Urban Economics*, 36(6): 727-752.

[27] Rosenthal, S., and W. Strange, 2004, Evidence on the nature and sources of agglomeration economies, in: Henderson, J.V. and J.F. Thisse (eds.) *Handbook of Regional and Urban Economics, Volume 4*, Elsevier North-Holland.

[28] Schürmann, C. and A. Talaat, 2000, *Towards a European Peripherality Index*, Report for DG Regional Policy of the European Commission, University of Dortmund.

[29] Sveikauskas, L.A., 1975, The Productivity of Cities, *Quarterly Journal of Economics*, 89, 393-413.

[30] Windmeijer, F., 2005, A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators, *Journal of Econometrics*, 126, 25-51.

Table 1: Summary Statistics

Variables		Mean	Median	Maximum	Region with maximum	Minimum	Region with minimum	Coefficient of variation
Labour productivity (*)	**(3)/(2)**	35.65	36.16	86.05	Groningen (NL)	6.87	Vychodne Slovensko (SK)	0.44
Labour productivity by country (*)	*# regions*							
France	21	51.59	50.06	69.39	Île de France (FR)	46.54	Poitou-Charentes (FR)	0.10
Germany	41	46.02	47.06	62.35	Hamburg (DE)	31.92	Chemnitz (DE)	0.16
Italy	19	35.86	35.28	42.04	Valle d'Aosta (IT)	30.57	Puglia (IT)	0.10
Poland	16	7.83	7.70	9.24	Mazowieckie (PL)	7.10	Warminsko-Mazurskie (PL)	0.18
Spain	15	32.69	33.20	40.00	Madrid (ES)	25.75	Extremadura (ES)	0.11
United Kingdom	37	26.74	26.57	34.77	Inner London (UK)	18.93	Cornwall (UK)	0.12
Area in km^2 (*)	**(1)**	16.04	10.00	154.00	Övre Norrland (SE)	0.16	Bruxelles (BE)	1.23
Total employment (**)	**(2)**	0.50	0.38	3.43	Île de France (FR)	0.03	Valle d'Aosta (IT)	0.87
Total GVA (1995 Euro) (***)	**(3)**	18.60	12.50	237.97	Île de France (FR)	1.12	Ipeiros (GR)	1.23
Employment density	**(2)/(1)**	123.94	37.65	5220.98	Inner London (UK)	0.89	Övre Norrland (SE)	3.36
Primary education$^+$		0.45	0.47	0.89	Algarve (PT)	0.10	Brandenburg–Nordost (DE)	0.40
Tertiary education$^+$		0.17	0.17	0.39	Brussels (BE)	0.01	Algarve (PT)	0.40

Notes: (*) in 10^3, (**) in 10^6, (***) in 10^9; time averages for 245 regions; $^+$ population shares for 171 regions in Belgium, Germany, France, Portugal, Spain and the UK.

Table 2: Aggregate Panel Regressions

Dependent variable = Log labour productivity (t)		(a) OLS	(b) FE	(c) DIFF-GMM	(d) SYS-GMM	(e) SYS-GMM
	Log labour productivity (*t*-1)	0.969***	0.729***	0.724***	0.754***	0.719***
		[0.004]	[0.018]	[0.092]	[0.109]	[0.059]
	Log employment density (*t*)	-0.419***	-0.594***	-0.510*	-0.387*	-1.028
		[0.034]	[0.029]	[0.261]	[0.227]	[1.278]
	Log employment density (*t*-1)	0.422***	0.440***	0.288	0.420*	1.066
		[0.034]	[0.031]	[0.316]	[0.226]	[1.289]
	Log employment (*t*)					0.656
						[1.078]
	Log employment (*t*-1)					-0.622
						[1.082]
	Constant	0.379***	3.424***		0.001	-0.0010
		[0.045]	[0.230]		[0.002]	[0.003]
	Observations	1504	1504	1229	1504	1504
	R-squared	0.99	0.88			
	Hansen test (*p*-value)			0.38	0.63	0.99
	Difference Hansen test (*p*-value)				0.71	0.99
	AR2 test (*p*-value)			0.24	0.18	0.18
Common factor restrictions:	**Emp. density**	0.00	0.70	0.77	0.01	0.33
	Employment					0.60
LR employment density:	$(\beta_0+\beta_1)/(1-\alpha)$	0.11***	-0.57***	-0.80	0.13*	0.13**
		[0.04]	[0.08]	[1.04]	[0.07]	[0.06]
	$\beta_0+\beta_1$	0.003***	-0.15***	-0.22	0.03**	0.04***
		[0.001]	[0.02]	[0.31]	[0.01]	[0.01]
LR employment:	$(\gamma_0+\gamma_1)/(1-\alpha)$					0.12**
						[0.05]
	$\gamma_0+\gamma_1$					0.03**
						[0.02]

Notes:
Robust standard errors in brackets (for GMM: based on two-step Windmeijer finite sample adjustment). Fixed effects are not reported. GMM estimations based on country-year mean-differenced data. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: Sector-Level Panel Regressions (SYS-GMM, Main Sectors)

Dependent variable =	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Log labour productivity (t)	**Manufacturing and Energy**				**Financial Services**			
Log labour productivity (*t*-1)	0.671***	0.738***	0.695***	0.820***	0.900***	0.885***	0.780***	0.798***
	[0.059]	[0.086]	[0.089]	[0.101]	[0.073]	[0.075]	[0.064]	[0.066]
Log own-sector employment density (*t*)	-0.845*	-0.712**	-1.386**	-0.595**	-0.792***	-0.929***	-0.605***	-0.815***
	[0.444]	[0.326]	[0.583]	[0.278]	[0.198]	[0.176]	[0.187]	[0.164]
Log own-sector employment density (*t*-1)	0.781*	0.692**	1.347**	0.582**	0.963***	1.082***	0.654***	0.867***
	[0.433]	[0.309]	[0.561]	[0.252]	[0.214]	[0.190]	[0.187]	[0.160]
Log other-sector output density (*t*)	-0.556	0.006			0.552	0.676*		
	[0.793]	[0.668]			[0.366]	[0.389]		
Log other-sector output density (*t*-1)	0.612	0.019			-0.708**	-0.813**		
	[0.799]	[0.689]			[0.359]	[0.388]		
Log total employment (*t*)	0.612		1.358**		-0.289		-0.493	
	[0.705]		[0.663]		[0.404]		[0.471]	
Log total employment (*t*-1)	-0.577		-1.324**		0.272		0.500	
	[0.702]		[0.665]		[0.405]		[0.468]	
Hansen test (*p*-value)	0.99	0.99	0.99	0.97	0.71	0.70	0.72	0.52
Difference Hansen test (*p*-value)	0.99	0.99	0.99	0.91	0.77	0.86	0.79	0.63
AR2 test (*p*-value)	0.10	0.07	0.04	0.05	0.71	0.50	0.29	0.75
Common factor restrictions: Own-sector empl. density	0.14	0.14	0.08	0.20	0.00	0.00	0.00	0.00
Other-sector outp. density	0.41	0.91			0.01	0.02		
Total employment	0.50		0.15		0.84		0.36	
LR own-sector employment density: $(\beta_0^{own}+\beta_1^{own})/(1-\alpha)$	-0.20	-0.08	-0.13	-0.07	1.70	1.32	0.23***	0.26***
	[0.13]	[0.14]	[0.18]	[0.19]	[1.49]	[1.03]	[0.09]	[0.1]
$\beta_0^{own}+\beta_1^{own}$	-0.06	-0.02	-0.04	-0.01	0.17***	0.15***	0.05***	0.05***
	[0.04]	[0.04]	[0.05]	[0.03]	[0.05]	[0.04]	[0.01]	[0.01]
LR other-sector output density: $(\beta_0^{other}+\beta_1^{other})/(1-\alpha)$	0.17*	0.09			-1.55	-1.20		
	[0.09]	[0.12]			[1.50]	[1.05]		
$\beta_0^{other}+\beta_1^{other}$	0.06**	0.02			-0.16**	-0.14**		
	[0.03]	[0.03]			[0.06]	[0.05]		
LR total employment: $(\gamma_0+\gamma_1)/(1-\alpha)$	0.11		0.11		-0.17		0.03	
	[0.07]		[0.11]		[0.25]		[0.11]	
$\gamma_0+\gamma_1$	0.03		0.03		-0.02		0.01	
	[0.02]		[0.04]		[0.02]		[0.02]	

Notes:
1,504 observations. Robust two-step standard errors based on Windmeijer finite sample correction in brackets. Estimations based on country-year mean-Differenced data. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: Sector-Level Panel Regressions (SYS-GMM, Remaining Sectors)

		(a)	(b)	(c)	(d)	(e)
Dependent variable = Log labour productivity (t)		**Construction**	**Wholesale, Retail**	**Hotel, Restaurant**	**Transp., Comm.**	**Other Mkt Serv.**
Log labour productivity (*t*-1)		0.617***	0.801***	0.814***	0.855***	0.544***
		[0.153]	[0.073]	[0.174]	[0.136]	[0.071]
Log own-sector employment density (*t*)		-0.655	-0.744***	-0.548	-1.269**	-1.121***
		[0.420]	[0.154]	[0.374]	[0.507]	[0.191]
Log own-sector employment density (*t*-1)		0.221	0.712***	0.654*	1.005**	0.794***
		[0.406]	[0.171]	[0.360]	[0.433]	[0.163]
Log other-sector output density (*t*)		3.354*	0.638**	0.695	1.348**	0.489
		[1.722]	[0.277]	[0.449]	[0.532]	[0.312]
Log other-sector output density (*t*-1)		-3.012*	-0.568**	-0.715	-1.046**	-0.168
		[1.576]	[0.252]	[0.438]	[0.443]	[0.326]
Log total employment (*t*)		-1.572*	-0.07	-0.686	0.048	1.055*
		[0.892]	[0.318]	[0.548]	[0.976]	[0.569]
Log total employment (*t*-1)		1.417	0.069	0657	-0.007	-1.019*
		[0.894]	[0.314]	[0.550]	[1.066]	[0.545]
Hansen test (*p*-value)		0.79	0.70	0.26	0.70	0.18
Difference Hansen test (*p*-value)		0.79	0.56	0.51	0.62	0.01
AR2 test (*p*-value)		0.05	0.64	0.18	0.04	0.08
Common factor restrictions:	**Own-s. dens.**	0.42	0.03	0.02	0.61	0.06
	Other-s. dens.	0.39	0.24	0.31	0.52	0.57
	Total empl.	0.40	0.84	0.60	0.89	0.16
LR own-s. density:	$(\beta_0^{own} + \beta_1^{own})/(1-\alpha)$	-1.13*	-0.16	0.57	-1.82	-0.72***
		[0.66]	[0.26]	[1.00]	[1.64]	[0.16]
	$\beta_0^{own} + \beta_1^{own}$	-0.43*	-0.03	0.11	-0.26*	-0.33***
		[0.25]	[0.06]	[0.12]	[0.14]	[0.1]
LR other-s. density:	$(\beta_0^{other} + \beta_1^{other})/(1-\alpha)$	0.89	0.35	-0.11	2.08	0.70***
		[0.66]	[0.26]	[0.71]	[1.63]	[0.18]
	$\beta_0^{other} + \beta_1^{other}$	0.34	0.07	-0.02	0.30*	0.32***
		[0.24]	[0.07]	[0.12]	[0.16]	[0.11]
LR total employment:	$(\gamma_0+\gamma_1)/(1-\alpha)$	-0.41	-0.01	-0.15	0.28	0.08
		[0.43]	[0.05]	[0.41]	[1.12]	[0.07]
	$\gamma_0+\gamma_1$	-0.16	0.00	-0.03	0.04	0.04
		[0.16]	[0.01]	[0.08]	[0.14]	[0.03]

Notes:
1,504 observations. Robust two-step standard errors based on Windmeijer finite sample correction in brackets. Estimations based on country-year mean-differenced data. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: Aggregate Panel Regressions (With Neighbourhood (columns a-e) and Market Potential (columns f and g) Effects)

Dependent variable =		(a)	(b)	(c)	(d)	(e)	(f)	(g)
Log labour productivity (t)		**OLS**	**FE**	**DIFF-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**	**SYS-GMM**
Log labour productivity (*t*-1)		0.986***	0.709***	0.744***	0.768***	0.718***	0.716***	0.684***
		[0.005]	[0.017]	[0.118]	[0.119]	[0.088]	[0.094]	[0.083]
Log employment density (*t*)		-0.614***	-0.729***	-0.494*	-0.247	-0.215	-0.307	-1.039
		[0.043]	[0.026]	[0.299]	[0.238]	[2.222]	[0.214]	[1.421]
Log employment density (*t*-1)		0.617***	0.531***	0.391	0.268	0.259	0.349	1.065
		[0.043]	[0.028]	[0.301]	[0.240]	[2.233]	[0.220]	[1.426]
Log employment (*t*)						0.109		0.700
						[1.879]		[1.197]
Log employment (*t*-1)						-0.068		-0.658
						[1.867]		[1.201]
Log market potential							-0.023	-0.020
							[0.036]	[0.023]
Log neighbours' output density (*t*)		0.556***	0.577***	0.006	-0.249	-0.194		
		[0.085]	[0.028]	[0.342]	[0.329]	[0.212]		
Log neighbours' output density (*t*-1)		-0.557***	-0.429***	-0.434	0.259	0.156		
		[0.086]	[0.029]	[0.456]	[0.348]	[0.236]		
Constant		0.185***	1.683***		-0.000	0.000	0.002	-0.003
		[0.045]	[0.264]		[0.002]	[0.010]	[0.004]	[0.003]
Observations		1504	1504	1229	1504	1504	1096	1096
R-squared		0.99	0.91					
Hansen test (*p*-value)				0.50	0.57	0.99	0.82	0.65
Difference Hansen test (*p*-value)					0.54	0.99		
AR2 test (*p*-value)				0.42	0.09	0.11	0.30	0.30
Common factor restrictions:	**Emp. density**	0.00	0.45	0.93	0.16	0.87	0.10	0.42
	Employment					0.98		0.63
LR employment density:	$(\beta_0+\beta_1)/(1-\alpha)$	0.19*	-0.68***	-0.40	0.09	0.16*	0.15	0.08*
		[0.11]	[0.06]	[1.21]	[0.09]	[0.09]	[0.11]	[0.05]
	$\beta_0+\beta_1$	0.003**	-0.20***	-0.10	0.02	0.04**	0.04*	0.03**
		[0.001]	[0.02]	[0.32]	[0.02]	[0.02]	[0.02]	[0.01]
LR neighbours' output density:	$(\gamma_0+\gamma_1)/(1-\alpha)$	-0.09	0.51***	1.68	0.04	-0.14		
		[0.12]	[0.06]	[1.71]	[0.15]	[0.15]		
	$\gamma_0+\gamma_1$	-0.001	0.15***	-0.10	0.01	-0.04		
		[0.002]	[0.02]	[0.32]	[0.03]	[0.04]		

Notes: Robust standard errors in brackets (for GMM: based on two-step Windmeijer finite sample adjustment). Fixed effects are not reported. GMM estimations based on country-year mean-differenced data. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Aggregate Panel Regressions, ADL(0,1)

Dependent variable = Log labour productivity (t)		(a) **OLS**	(b) **FE**	(c) **DIFF-GMM**	(d) **SYS-GMM**	(e) **SYS-GMM**
Log labour productivity (*t*-1)		0.960***	0.664***	0.656***	0.797***	0.733***
		[0.004]	[0.019]	[0.087]	[0.107]	[0.092]
Log employment density (*t*)		0.003**	-0.294***	-0.491**	0.030*	0.032***
		[0.001]	[0.022]	[0.206]	[0.016]	[0.012]
Log employment (*t*)						0.028**
						[0.014]
Constant		0.468***	3.424***		-0.000	-0.001
		[0.044]	[0.230]		[0.001]	[0.002]
Observations		1504	1504	1229	1504	1504
R-squared		0.99	0.86			
Hansen test (*p*-value)				0.64	0.63	0.99
Difference Hansen test (*p*-value)					0.52	0.99
AR2 test (*p*-value)				0.41	0.17	0.16
LR employment density:	**(β_0)/(1-α)**	0.07**	-0.88***	-1.043*	0.15**	0.12**
		[0.03]	[0.07]	[0.77]	[0.07]	[0.05]
LR employment:	**(γ_0)/(1-α)**					0.11*
						[0.06]

Notes:
Robust standard errors in brackets (for GMM: based on two-step Windmeijer finite sample adjustment). Fixed effects are not reported. GMM estimations based on country-year mean-differenced data. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: Aggregate Cross-Section OLS and 2SLS Regressions, Individual Years

Dependent variable = Log labour productivity

OLS

	1980-82	1986-88	1992-94					1998-2000	
Log empl. density	0.032**	0.038***	0.062***	0.057***	0.058***	0.056***	0.038***	0.068***	0.057***
	[0.012]	[0.011]	[0.012]	[0.013]	[0.016]	[0.015]	[0.006]	[0.017]	[0.016]
Primary education						2.328***			
						[0.618]			
Tertiary education						2.837***			
						[0.540]			
Constant	10.467***	10.341***	10.545***	10.387***	10.562***	8.968***	8.827***	10.654***	10.715***
	[0.045]	[0.024]	[0.026]	[0.027]	[0.062]	[0.242]	[0.030]	[0.066]	[0.042]
Eastern countries	NO	NO	YES	NO	NO	NO	NO	YES	NO
Country-specif. coeffs on human cap.	NO	NO	NO	NO	NO	NO	YES	NO	NO
Observations	198	199	245	210	171	171	171	245	210
R-squared	0.85	0.85	0.96	0.80	0.77	0.82	0.89	0.95	0.79

2SLS +

	1980/82	1986/88	1992/94					1998/2000	
Log empl. density	0.009	0.043***	0.039***	0.033***	0.043**	0.049**	0.010	0.050***	0.045***
	[0.017]	[0.011]	[0.012]	[0.013]	[0.018]	[0.021]	[0.015]	[0.015]	[0.015]
Primary Education						2.331***			
						[0.503]			
Tertiary Education						2.910***			
						[0.943]			
Constant	10.192***	10.331***	10.635***	10.658***	10.417***	8.801***	8.999***	10.733***	10.748***
	[0.037]	[0.033]	[0.097]	[0.100]	[0.074]	[0.377]	[0.626]	[0.092]	[0.094]
Eastern Countries	NO	NO	YES	NO	NO	NO	NO	YES	NO
Country-specif. coeffs on human cap.	NO	NO	NO	NO	NO	NO	YES	NO	NO
Observations	198	199	245	210	171	171	171	245	210
First-stage *F*-statistic	44.2	117.88	42.78	42.23	17.57	13.47	13.25	43.06	42.43

Notes:
Robust standard errors in brackets; country fixed effect are not reported. * significant at 10%; ** significant at 5%; *** significant at 1%. + employment density instrumented with regional area.

Appendix Table 1: Sector-Level Panel Regressions (SYS-GMM, Main Sectors, With Neighbourhood Effects)

Dependent variable =	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Log labour productivity (t)	**Manufacturing**				**Financial Services**			
Log labour productivity (*t*-1)	0.672***	0.707***	0.669***	0.778***	0.926***	0.918***	0.921***	0.893***
	[0.055]	[0.077]	[0.075]	[0.129]	[0.084]	[0.075]	[0.068]	[0.060]
Log own-sector employment density (*t*)	-1.111***	-0.717***	-1.361***	-0.613**	-0.735***	-0.917***	-1.322***	-0.961***
	[0.399]	[0.229]	[0.515]	[0.306]	[0.204]	[0.166]	[0.392]	[0.200]
Log own-sector employment density (*t*-1)	1.051***	0.702***	1.323***	0.585**	0.902***	1.060***	1.326***	0.983***
	[0.396]	[0.220]	[0.488]	[0.293]	[0.227]	[0.187]	[0.379]	[0.190]
Log other-sector output density (*t*)	-0.043	0.410			0.647*	0.777**		
	[0.511]	[0.403]			[0.359]	[0.370]		
Log other-sector output density (*t*-1)	0.081	-0.400			-0.798**	-0.918***		
	[0.524]	[0.417]			[0.334]	[0.354]		
Log total employment (*t*)	0.824		1.467**		-0.519		1.02	
	[0.686]		[0.681]		[0.553]		[0.797]	
Log total employment (*t*-1)	-0.785		-1.426**		0.495		-1.014	
	[0.694]		[0.681]		[0.555]		[0.795]	
Hansen test (*p*-value)	0.99	0.98	0.99	0.93	0.63	0.43	0.94	0.56
Difference Hansen test (*p*-value)	0.99	0.99	0.99	0.92	0.86	0.36	0.95	0.29
AR2 test (*p*-value)	0.05	0.02	0.04	0.04	0.54	0.28	0.84	0.63
Common factor restrictions: Own-sector empl. density	0.06	0.03	0.04	0.15	0.00	0.00	0.21	0.04
Other-sector prod. density	0.77	0.46			0.01	0.01		
LR own-sector employment density: $(\beta_0^{own}+\beta_1^{own})/(1-\alpha)$	-0.18	-0.05	-0.12	-0.13	2.25	1.74	0.05	0.21
	[0.18]	[0.12]	[0.18]	[0.16]	[2.99]	[2.08]	[0.03]	[0.22]
$\beta_0^{own}+\beta_1^{own}$	-0.06	-0.01	-0.04	-0.03	0.17***	0.14**	0.00	0.02
	[0.06]	[0.04]	[0.05]	[0.03]	[0.06]	[0.06]	[0.02]	[0.02]
LR other-sector output density: $(\beta_0^{other}+\beta_1^{other})/(1-\alpha)$	0.12	0.03			-2.04	-1.72		
	[0.13]	[0.10]			[2.96]	[2.16]		
$\beta_0^{other}+\beta_1^{other}$	0.04	0.01			-0.15**	-0.14**		
	[0.04]	[0.03]			[0.07]	[0.07]		
LR own-sector neighbour's output density: $(\gamma_0+\gamma_1)/(1-\alpha)$	0.17	0.15	0.24	0.01	0.24	0.51	2.78	1.44
	[0.29]	[0.24]	[0.21]	[0.49]	[0.99]	[0.74]	[2.87]	[1.16]
$\gamma_0+\gamma_1$	0.05	0.04	0.08	0.00	0.02	0.04	0.22***	0.15**
	[0.10]	[0.08]	[0.07]	[0.11]	[0.07]	[0.06]	[0.07]	[0.07]

Notes:
1,504 observations. Neighbouring own-sector and other-sector output density included but only own-sector LR effects are reported. Remaining effects are not significant. Robust two-step standard errors based on Windmeijer finite sample correction in brackets. Estimations based on country-year mean-differenced data.
* significant at 10%; ** significant at 5%; *** significant at 1%.

Appendix Table 2: Cross-Section OLS and 2SLS Regressions (Main Sectors, With Neighbourhood Effects)

Dependent variable = Log labour productivity

OLS

	Manufacturing						Financial Services					
	1980-82		2001-03				1980-82		2001-03			
Log own-s. empl. density	0.026	0.008	-0.196***	0.014	-0.198***	0.022	-0.403*	-0.052	-0.332**	0.045	-0.341**	0.064**
	[0.069]	[0.022]	[0.041]	[0.022]	[0.038]	[0.021]	[0.217]	[0.035]	[0.141]	[0.026]	[0.129]	[0.029]
Log other-s. empl. density	-0.019		0.212***		0.218***		0.436*		0.455**		0.486***	
	[0.077]		[0.029]		[0.026]		[0.242]		[0.161]		[0.150]	
Constant	10.337***	10.550***	8.893***	11.149***	8.257***	10.964***	4.531	10.701***	4.286*	10.620***	4.288*	11.333***
	[0.879]	[0.060]	[0.338]	[0.031]	[0.315]	[0.055]	[3.404]	[0.028]	[2.240]	[0.006]	[2.173]	[0.025]
Eastern countries	NO	NO	NO	NO	YES	YES	NO	NO	NO	NO	YES	YES
Observations	198	198	210	210	245	245	198	198	210	210	245	245
R-squared	0.59	0.59	0.76	0.71	0.93	0.91	0.88	0.86	0.77	0.70	0.93	0.91

2SLS +

	Manufacturing						Financial Services					
	1980-82		2001-03				1980-82		2001-03			
Log own-s. empl. density	-0.163	-0.012	-0.187	0.019	-0.196**	0.025	0.121	-0.036	-0.144	0.082***	0.035	0.090***
	[0.301]	[0.036]	[0.118]	[0.019]	[0.098]	[0.019]	[0.940]	[0.040]	[0.205]	[0.020]	[0.230]	[0.021]
Log other-s. empl. density	0.164		0.204*		0.216**		-0.177		0.24		0.058	
	[0.291]		[0.107]		[0.087]		[1.100]		[0.232]		[0.255]	
Constant	8.622**	10.604***	8.427***	10.971***	8.845***	10.958***	11.855	10.688***	7.867**	11.317***	8.961***	11.311***
	[3.435]	[0.128]	[1.316]	[0.138]	[1.029]	[0.136]	[15.226]	[0.182]	[3.348]	[0.060]	[3.304]	[0.057]
Eastern countries	NO	NO	NO	NO	YES	YES	NO	NO	NO	NO	YES	YES
Observations	198	198	210	210	245	245	198	198	210	210	245	245
First-Stage *F*-statistic	8.92	52.75	8.38	47.43	11.95	48.38	1.16	40.50	8.74	34.31	8.62	34.89

Notes:
Neighbouring own-sector and other-sector output density included but not reported (no neighbour effects are statistically significant). Robust standard errors in brackets. Country fixed effect are not reported. * significant at 10%; ** significant at 5%; *** significant at 1%. + employment density instrumented with regional area.