CESIS Electronic Working Paper Series

Paper No. 170

Agglomeration and Productivity

– evidence from firm-level data

Martin Andersson* and Hans Lööf**

(*CESIS and JIBS, **CESIS and the Division of Economics, KTH)

March 2009

The Royal Institute of technology
Centre of Excellence for Science and Innovation Studies (CESIS)
http://www.cesis.se

Agglomeration and Productivity

- evidence from firm-level data

Martin Andersson[⊗] and Hans Lööf[⊕]

Abstract

Do agglomerations stimulate productivity? An extensive literature on agglomeration economies, or urban increasing returns, has analyzed this question with aggregated spatial data. This paper estimates the relationship between agglomeration and productivity at the firm level using static and dynamic models. It makes use of a rich dataset comprising register information on all manufacturing firms in Sweden with 10 or more employees over the period 1997-2004. Three things emerge. First, firms located in larger regions are more productive when controlling for size, human capital, physical capital, ownership structure, import and export, industry classification and time trend. Second, results from dynamic panel estimations suggest a learning effect in that agglomeration enhances firms' productivity. Third, the role of agglomeration phenomena does not seem to have a clear coupling to firm size.

Keywords: productivity, agglomeration economies, spatial externalities, external scale economies, urban increasing returns, spatial selection, spatial sorting

JEL: R30, R12, L25'

[⊗] Centre of Excellence for Science and Innovation Studies (CESIS), Royal Institute of Technology and Jönköping International Business School (JIBS). E-mail: martin.andersson@ihh.hj.se

[⊕] Division of Economics and Centre of Excellence for Science and Innovation Studies (CESIS), Royal Institute of Technology. E-mail: hansl@infra.kth.se

Acknowledgements: We are grateful for comments and remarks by Andreas Stephan and from seminar participants at the 11th Uddevalla Symposium (May 15-17, 2008) at the Kyoto University, Kyoto, Japan. Financial support from the Swedish Governmental Agency for Innovation Systems (VINNOVA) and from the EU sixth framework program (MICRODYN project) is gratefully acknowledged.

1 INTRODUCTION

It is a stylized fact that economic activity is concentrated in space. Some scholars maintain that the tendency of firms and individuals to cluster in space is actually one of the most prominent features of contemporary economic geography (Krugman 1991). Why does such concentration take place? A major argument is that large and dense agglomerations are associated with place-specific external scale economies that bring about productivity advantages. Such ‘agglomeration economies’ provide an economic rationale for why activities cluster spatially and why concentrations of firms and individuals tend to persist over time.[1]

The idea of agglomeration economies as forces of spatial concentration has spurred a vast amount of research on the relationship between agglomerations and productivity (cf. Rosenthal and Strange 2004). Fundamental research questions include: Do agglomerations influence productivity? If so, to what extent? What is the nature of agglomeration economies and what are the sources?

This paper asks whether agglomerations have positive effects on productivity. This is an “old” question that has been discussed over the last four decades. Examples of studies include Åberg (1973), Sveikauskas (1975), Segal (1976), Moomaw (1981), Ciccone and Hall (1996), Braunerhjelm and Borgman (2004), Rice et al. (2006), Brülhart and Mathys (2008). There is now considerable evidence that productivity is positively associated with agglomeration phenomena. The ‘urban productivity premium’ is an established term in the literature.

The contribution of this paper is that it analyzes the relationship between agglomeration and productivity with firm-level panel data and addresses heterogeneity and endogeneity issues. Three questions are analyzed. First, are firms located in larger regions more productive when controlling for attributes of individual firms? Second, is there any difference in the relationship between region size and productivity between small and large firms? Third, is there a learning effect from agglomerations, such that firms become more productive by being located in agglomerations?

The bulk of the existing evidence is based on aggregate spatial or sectoral data. A firm-level approach is warranted for several reasons. First, the theory behind agglomeration economies

---

[1] Agglomeration economies are sometimes also referred to as urban increasing returns.

is truly micro-economic in nature, making postulations about how individual firms are affected by their external local environment. Second, a firm-level approach allows us to estimate the effect of the external local environment of a firm on its productivity, while accounting for an ample set of firm attributes. Controlling for attributes of individual firms reduces, for instance, the likelihood that estimated effects of agglomeration on productivity are driven by differences in internal firm attributes across locations. This is important as the magnitude of heterogeneity in resources across firms is substantial. A key assertion in the literature adhering to the resource-based view of the firm (RBV) is, for example, that a firm's competitive advantage depends critically on its internal resources and capabilities (Penrose 1959, Barney 1991).

Our empirical approach is similar to Moretti (2004) and Henderson (2003), who estimate plant-level production functions that are extended with variables reflecting the local environment. Moretti (2004) focuses on the education-level of the employees in the region, whereas Henderson (2003) focuses on the number of other firms in the same industry in the region as a source of spillover effects. We augment a firm-level production function with the size of the region the firm is located in. Regional size is assumed to reflect the potential for agglomeration economies and is introduced as a 'shifter' of the firms' production function.[2] Its effect is assumed to be external to the individual firm but internal to the region. Compared to previous studies that apply firm- or plant-level data, we control for a richer set of firm attributes that are likely to influence productivity. In addition to capital, ordinary and knowledge labor, we control for whether the firm belongs to a multinational corporation (foreign and domestic), its export and import activities and its industry affiliation.

The analysis includes both static and dynamic panel data estimations. In the former case we test whether the level of productivity of firms is higher in larger regions while controlling for firm attributes. In the latter we condition on the level of productivity in previous periods, and test whether the current productivity is higher for firms located in larger regions. The dynamic panel model is intended to account for endogeneity that may be associated with the static panel estimations, arising, for instance, from the fact that firms with unobserved

---

[2] There is a considerable theoretical discussion on the relative importance of localization, urbanization and Jacobs externalities (see Frenken et al. 2007). Whether it is diversity or specialization that underlies agglomeration economies is beyond the scope of this paper. We focus on the general size effect and acknowledge that it may be due to external economies associated with either specialization, diversity or both (cf. Moomaw 1983). In general, a larger region permits more specialization.

attributes that influence productivity are more likely to be located in agglomerations (cf. Rosenthal and Strange 2004). Whereas a positive effect in the static models may be explained by either selection or learning, a significant effect of agglomerations in the dynamic model clearly points to a learning effect, i.e. firms become more productive by being located in agglomerations.

We find evidence of agglomeration economies with both the static and dynamic models. To be more precise, the following results emerge from the study: First, firms located in larger regions are more productive when controlling for size, human capital, physical capital, ownership structure, import and export, industry classification and time trend. Second, the marginal effect of agglomeration on firms' productivity growth can be given a positive causative interpretation. Third, the role of agglomeration phenomena does not seem to have a clear coupling to firm size.

The remainder of the paper is organized in the following fashion: Section 2 discusses the relationship between agglomerations and productivity. Section 3 contains the data and the empirical strategy. Section 4 presents the results and Section 5 concludes the paper.

2 AGGLOMERATION ECONOMIES AND PRODUCTIVITY

The argument that firms benefit from a location in an agglomeration due to place-specific external economies of scale dates back to the early work by Marshall (1920), Ohlin (1933) and Hoover (1937). Marshall (1920) maintained that concentrations of firms in a similar industry give rise to localization economies in the form of knowledge and information spillovers, labor pooling (advantages of thick markets for specialized skills) and backward and forward linkages. Ohlin (1933) and Hoover (1937) distinguished between urbanization and localization economies. Urbanization economies refer to economies that pertain to larger urban regions with a diversified economy. Localization economies are external to the firm, but internal to the industry in a given region, and are often associated with specialization phenomena. Urbanization economies are external to the individual firm, but internal to a region as a whole, and are thus assumed to bring benefits to all firms located in the region. The notion of urbanization economies corresponds closely to the ideas in Jacobs (1969), who emphasizes the role of diversity in regional economies.

We focus on productivity effects of external economies associated with the size of a region, and do not inquire into any specific type of external economy. We refer to agglomeration economies as a comprehensive term for external economies associated with the size of a region, which may be understood as urban increasing returns (cf. Fujita and Thisse 2002).

How can urban increasing returns be explained? Duranton and Puga (2004) list three types of general micro-foundations for urban agglomerations based on different mechanisms; (i) sharing, (ii) matching and (iii) learning. The first refers to benefits that arise due to larger regions being able to sustain e.g. a larger variety of input suppliers and a deeper division of labor that makes workers more productive. The second refers to the matching models of Hesley and Strange (1990) and Kim (1990), in which more agents trying to match increase the expected quality of each match. The third mechanism refers to the fact that urban agglomerations offer opportunities for the generation, the diffusion, and the accumulation of knowledge. All three mechanisms provide arguments for why productivity may be enhanced in urban agglomerations, and thus reasons why they persist.

Most of the empirical work on productivity effects focuses on the learning mechanism and emphasizes the role of spillovers and its impact on technology. The effect of the external environment is typically assumed to be such that it ‘shifts’ the production function of firms in the region (cf. Rosenthal and Strange 2004). Arguments often adhere to endogenous growth theory as formulated in Romer (1990) and Lucas (1988). Since knowledge is non-rival and tends to diffuse and spillover, there are increasing returns to knowledge through diffusion and spillovers, which are more effective in dense and populous environments (cf. Andersson and Karlsson 2007). Dense urban regions with richness in sectors and knowledge sources offer interaction opportunities with different actors embodying relevant knowledge, such as customers, knowledge-intensive business services, universities and other ‘knowledge-handlers’. Moreover, flows of labor and technical personnel between firms tend to be greater in dense locations, stimulating the diffusion of competencies and knowledge embodied in people (Almeida and Kogut 1999, Moen 2005, Fallick et al. 2006). The argument that urban agglomerations are associated with more intense knowledge flows is supported, for instance, by innovation activities being more concentrated in space than standard production activities (Audretsch and Feldman 1996), patent citations being geographically localized (Jaffe et al. 1993) and innovations tending to diffuse faster within clusters (Baptista 2000).

The various external effects associated with urban agglomerations are assumed to bring productivity advantages to firms located in larger and denser regions. There are ample empirical analyses whose results are consistent with this conjecture. Several studies find a positive relationship between the overall size of a region and productivity (Shefer 1973, Segal 1976, Sveikauskas 1975, Nakamura 1985, Rice et al. 2006) as well as between employment density and productivity (Ciccone and Hall 1996, Ciccone 2002, Brülhart and Mathys 2008). Studies at the level of individuals also show that workers are more productive in larger regions (Glaeser and Maré 2001, Wheeler 2006, Yankow 2006). At the level of individual plants, Henderson (2003) finds that the number of other plants in the same industry enhances productivity, and Moretti (2004) shows that productivity is higher in regions with higher shares of college graduates.

A number of previous studies have analyzed the relationship between productivity and agglomeration phenomena in Sweden. An early contribution is Åberg (1973), who found that labor productivity in Swedish manufacturing industries is an increasing function of regional size, when controlling for capital intensity and average firm-size in the regions. This pattern was found to be persistent over years and the metropolitan regions displayed the highest productivity. Klaesson and Larsson (2008) make use of data on Swedish municipalities for the 2000s. As in Rice et al. (2006), they employ a sector-adjusted productivity index and show that the average wage per employee of municipalities is an increasing function of their market-potential. Karlsson and Pettersson (2005) report similar findings. Braunerhjelm and Borgman (2004) analyze the productivity growth of manufacturing and service sectors across Swedish regions for the period 1975-1999. Their analysis indicate that more concentrated sectors have higher productivity growth and that regional size (in terms of population) stimulates productivity growth. Andersson and Karlsson (2007) also find a positive relationship between the change in value-added per employee and regional size. In summary, several previous studies find a positive relationship between agglomeration phenomena and productivity across Swedish regions, in terms of productivity levels and growth.[3] To the best of our knowledge, however, no previous study using Swedish data has analyzed the relationship between regional size and productivity with firm-level data and conditioned on attributes of individual firms.

---

[3] Studies on Swedish regions also show that innovations are more frequent in the larger regions (see e.g. Johansson and Lööf 2008, Andersson and Johansson 2008, Ejermo and Gråsjö 2008).

A major issue in the literature is the question of ‘learning’ versus ‘selection’. While agglomerations may enhance productivity, they may also attract productive firms with certain attributes (Rosenthal and Strange 2004). Higher productivity in agglomerations may thus not necessarily be due to firms becoming more productive because of agglomeration economies. It might instead be due to the fact that firms with advantageous internal attributes are more likely to be located in agglomerated areas. Empirical analyses indeed show that firm-level heterogeneity is substantial, even in narrowly defined industries (cf. Bernard and Jensen 1995). Recent theoretical work on agglomeration with heterogeneous firms also demonstrates precisely that estimates of agglomeration economies may be overstated. Because the most productive firms are larger and sell more, they have the strongest incentive to locate in large agglomerations to save on transportation costs, regardless of any agglomeration economy (Baldwin and Okubo 2006). The result is spatial selection of firms with different (ex ante) levels of productivity.

In individual-level studies of the ‘urban productivity premium’, the issue of ‘selection’ versus ‘learning’ is often dealt with by making a distinction between a level- and growth-effect (see e.g. Glaeser and Maré 2001, Wheeler 2006, Yankow 2006). In essence, the growth in wages of people that move to cities is compared with that of people who do not. Higher growth is consistent with a learning-effect, whereas a direct level-effect, but similar wage growth, points to a selection effect. We deal with these issues by estimating a production function at the level of individual firms, while controlling for attributes of firms which may influence their productivity, and estimate both static and dynamic panel models. Whereas a positive effect in the static models may be explained by either selection or learning, a significant effect of agglomerations in the dynamic model clearly points to a learning effect, i.e. that firms become more productive by being located in agglomerations.

3 EMPIRICAL ANALYSIS

3.1 Data

The data for this study covers the period 1997-2004 and are maintained by Statistics Sweden (SCB). In the data material, a firm is defined as a legal entity. The basic dataset consists of balance-sheet information of about 130 000 observations on all manufacturing firms in

Sweden with one or more employees. We have imposed a censoring level of 10 employees in order to guarantee quality of the data. The official trade data is somewhat poor for the very small firms. The remaining unbalanced panel has 54,279 observations and the annual number of firms varies between 6,594 and 6,898

The econometric analysis of the relationship between productivity and regional size controls for the firms' internal knowledge, possible knowledge spillovers from other units within the group, and spillovers from import and export activities. It also controls for size, capital intensity and industry categorization.

Table 1 reports the summary statistics of the data we use for this purpose. The median manufacturing firm in Sweden with 10 or more employees has 23 ordinary employees and 1 employee with a university education of three years or more. In the following we label the latter as "skilled labor". Since the mean values for ordinary and skilled labor are 83 and 8, respectively, we conclude that the distribution of firms is skewed with many small firms and few large firms.

>> TABLE 1 ABOUT HERE <<

Our additional economic variables are those commonly used in the literature. They include physical capital, measured as investments in machinery and equipment, corporate ownership structure, international trade, industry classification and time trend.

In order to investigate if there is a difference in the role of the local environment for firms in different size classes, the sample is split into three groups: firms with 10-25 25-50 and 51 or more employees. The first group is the largest and corresponds to 52% of the sample. The second group includes 22% of the firms and the last 26%. We make this split check the robustness of our results. Small firms can in general be expected to be more dependent on the local environment than larger firms (cf. Henderson 2003).

The focus is on the influence of regional size (in terms of employment) on firms' productivity. A region is defined as a *functional region*. A functional region consists of several municipalities that together form an integrated local labor market. They are delineated based on the intensity of commuting flows between municipalities. We use a definition of functional regions in which there are 72 regions in Sweden.[4]

Each firm is assigned to a given functional region through a spatial identifier. A drawback of the data is that it does not contain information on whether a given firm is a multi-plant firm or not. However, multi-plant phenomena are mostly a feature of corporations. We observe individual firms and have information on whether they are part of a corporation (uninational or multinational). The different plants of corporations like Volvo, Ericsson and SAAB in Sweden are often registered as distinct legal entities, i.e. firms. Still, it is possible for large firms to have several plants (not registered as distinct legal entities) located in different functional regions. To the extent that the multi-plant firms have their different plants in different functional regions, multi-plant phenomena may potentially affect our results. Nonetheless, we believe that we account for this by splitting the sample into different size-classes. Small individual firms with 10-25 employees, for example, are not likely to be multi-plant firms.

As stated previously, we focus on the productivity effects of external economies associated with the size of a region and do not inquire into any specific type of external economy. Instead, we use regional size as a measure of the potential for agglomeration economies. Some background information on the spatial structure of the Swedish economy is provided in the Appendix. As in many other countries, the size of the functional regions in Sweden is positively associated with several characteristics that are likely to increase the potential for agglomeration economies. For instance, larger functional regions have a higher diversity in terms of sector (Duranton and Puga 2000), greater local supply of advanced business services (Klaesson and Johansson 2008), higher average education-level of the workforce and greater richness in knowledge sources, such as university and business R&D. Regional size can be interpreted as a 'catch-all' measure of the potential for agglomeration economies. While it does not inform about any specific type of external economy, it is useful for testing agglomeration economies (or urban increasing returns) as a general size-effect, which is precisely the purpose of this paper.

---

[4] Developed by NUTEK - the Swedish Agency for Economic and Regional Growth.

3.2 Model

The general model that we use for our empirical analysis is a standard Cobb-Douglas production function augmented with regional population size. The data are repeated measurement of the same firms at different points in time. Variation in these data can be decomposed into variation between firms of different sizes and characteristics such as industry classification, and variation within firms. The former variation is usually bigger than the latter. With more informative data we can produce more reliable parameter estimates.

Hsiao (2002) lists several benefits from using panel data in micro data studies. These include (i) accounting for the observed heterogeneity among the firms in the sample, since times series and cross-section studies not controlling for this heterogeneity run the risk of obtaining biased results, (ii) more informative data, more variability, less collinearity among the variables, more degrees of freedom and more efficiency, (iii) Panel data are better able to study the dynamics of adjustments.

The basic model can be expressed as:

$$Q_{it} = AK_{it}^{\beta_K} L_{it}^{\beta_V} H_{it}^{\beta_H} \qquad (1)$$

where the subscript $i=1,2,\ldots N$ refers to a firm, subscript $t=1,2\ldots,T$ refers to a point in time, Q_{it} is the value-added of firm i at time t , K_{it} is the capital input, L_{it} is the ordinary labor input, H_{it} is skilled labor.

The distinction between ordinary labor (L) and knowledge-intense labor (H) raises the issue of how labor productivity should be measured. H, defined as employees with at least 3 years of university education, is assumed to be associated with a firm's R&D. The standard measure of labor productivity is total value added over total employment. Another is based on arguments put forward in Griliches and Mairesse (1984), and considers the results of R&D efforts as an input to the basic production process, which implies that the return to R&D is reflected by its effect on the productivity of ordinary labor, i.e., its effect on $q = Q / L$. This approach considers the distinction between the production of knowledge and the returns to its use (Geroski et al. 1993), where the latter aspect is reflected by the impact of knowledge on q.

At each point in time (H) reflects the capacity to expand future knowledge. The size of (H) will also reflect the knowledge stock of a firm and its capacity to absorb external knowledge from the local milieu, the company group, national and international suppliers and consumers, and others (cf. Bartel and Lichtenberg 1987, Cohen and Levinthal 1990).

By dividing Q with ordinary labor, our preferred productivity measure, we may express (1) as a labor productivity function:

$$q_{it} \equiv \frac{Q_{it}}{L_{it}} = AK_{it}^{\beta_K} L_{it}^{(\beta_V - 1)} H_{it}^{\beta_H} \qquad (2)$$

Let us now turn to our assumption about technology, A. We assume that agglomeration phenomena influence firms' technology such that regional population size, i.e. our measure of the potential for agglomeration economies, influences A. A similar modeling framework is applied by Moretti (2004) and Henderson (2003). We model A in the following manner:

$$\ln A = \phi \ln S_{r,t} + \mathbf{x}_{\mathbf{it}}' \boldsymbol{\lambda} + \varepsilon_{it} \qquad (3)$$

where $S_{r,t}$ denotes the size of region r at time t and is assumed to reflect the potential for agglomeration economies. In this formulation, agglomeration phenomena 'shift' the production function of the firms. A region is defined as a *functional region*, which consists of several municipalities that form an integrated local labor market.[5] Within such a region, time distances between places are small enough to allow for frequent face-to-face contacts. $\mathbf{x}_{\mathbf{it}}$ consists of control variables and includes indicators of the firm's participation in international trade, its ownership structure and industry affiliation. Recent literature on international trade and productivity at the micro-level suggests that a firm's trade status is an important control variable (see e.g. Wagner 2007, Greenaway and Kneller 2007). The ownership structure of the firm, in particular whether it is part of an MNE, may also affect its technology. By definition, MNEs have established networks to a rich set of markets and thereby a coupling to several knowledge sources and innovation systems (cf. Dachs et al. 2008). They also have strong internal capabilities pertaining to the development of proprietary information and knowledge within the corporation (Pfaffermayr and Bellak 2002). Industry affiliation controls for industry heterogeneity. ε_{it} is an error term, which may be interpreted as capturing technological shocks. The full model is given by:

---

[5] Functional regions are delineated based on the intensity of commuting flows. We use NUTEK's definition of functional regions. There are 72 functional regions in Sweden according to this definition.

$$\ln q_{it} = \beta_K \ln K_{it} + \left(\beta_L - 1\right) \ln L_{it} + \beta_H \ln H_{it} + \phi \ln S_{r,t} + \mathbf{x}'_{\mathbf{it}} \boldsymbol{\lambda} + \varepsilon_{it} \qquad (4)$$

The variable of main interest is $S_{r,t}$, and we want to estimate the parameter ϕ. The hypothesis of agglomeration economies suggests that $\hat{\phi} > 0$, which means that firms located in larger regions, i.e. regions with a larger potential for agglomeration economies, have higher productivity.

3.3 Estimation strategy

In order to find appropriate methods for estimating Equation (4) we focus on methods for a short panel, meaning data on many firms but few time periods. Moreover, we are interested in methods that can estimate the marginal effects of regional size that can be given a causative interpretation. The Arellano-Bond (1991) and Arellano-Bover (1995), Blundell-Bond (1998) dynamic panel estimators are attractive for this purpose and for solving complications with heteroskedasticity, endogeneity and autocorrelation in our “small T” and “large N” panel. Hence, this is our preferred estimator motivated by both the dataset and the research question. However, we will start by estimating some standard static models and consider alternative specifications of these models.

Using the static approach, we run separate regressions for the three different size groups 10-25 employees, 26-50 employees and firms with more than 50 employees using five different estimators. The first is a pooled OLS (with cluster robust standard errors) with the assumption that the errors are independent over firms. We then consider four different estimators that assume unobserved individual effects, u_i. In this case the error term consists of two parts; individual effects and a standard disturbance term such that $\varepsilon_{it} = u_i + e_{it}$. We could think of u_i as unobserved managerial skill. The remainder disturbance e_{it} varies with individuals and time and can be thought of as the usual disturbance in the regression. The firm’s unobserved firm-specific effects may or may not be correlated with the regressors.

The panel data model with individual effects that we employ are the fixed-effects (FE) where the unobserved firm-effect are permitted to be correlated with the regressors, the random-effects (RE) model that assumes that the firm is purely random and uncorrelated with the regressors, a three-stage panel fixed effects vector decomposition model that allows for the inclusion of time-invariant variables within a fixed effects framework (Plümper and Troeger 2007), and finally the Hausman-Taylor estimator (RE-HT) which fits a random-effects models in which some of the covariates are correlated with the unobserved firm-specific random effect.

In a second step, we estimate the baseline model in Equation (4) in a dynamic setting. Endogeneity issues are main motivations for a dynamic panel data model. Beginning with Balestra and Nerlove (1966), Anderson and Hsiao (1982) and Holtz-Eakin, Newey and Rosen (1988), various instrumental solutions to the endogeneity problem have been suggested in the literature. Currently, the state of this art is the use of fully efficient GMM-estimators that allow for heteroskedasticity across firms, and serial correlation over time (See Arellano and Bond 1991). The basic idea of the GMM approach is to take the first-difference of a dynamic model specification in order to remove unobservable and time-invariant firm-specific effects, and then instrument the right-hand-side variables in the first-differenced equations using levels of the series lagged two periods or more, under the assumption that the time-varying disturbances in the original level equations are not serially correlated.

We use a two-step system GMM estimator (Arellano and Bover 1995, Blundell and Bond 1998) as the main dynamic model. It has some attractive advantages over alternative estimators that can deal with endogeneity problems. First, it can include time-invariant regressors. Second, it can make the Windmeijer (2005) finite-sample correction of the reported standard errors in two-sample estimation, without which these standard errors tend to be severely downward biased. Third, by making the assumption that the first differences of instrumenting variables are uncorrelated with the fixed effects, more instruments may be used, which can improve efficiency. Finally, this estimator allows finer control over the instrument matrix than alternative methods. One disadvantage of the System GMM (and difference GMM as well) is that it is complicated and can easily generate invalid estimates (Roodman 2006). The test statistics for autocorrelation, overidentification and exogeneity of instruments are therefore important information. Moreover, as a first check of the validity of the employed system GMM estimator, the literature suggests a comparison of the point estimate of the lagged dependent variable with results from pooled OLS and FE. Both these estimators are likely to produce biased estimates in our dynamic setting, but in opposite directions. Bond (2002) suggests that candidate consistent estimates are be expected to lie between OLS and FE estimates.

4 RESULTS

The key results are presented in Tables 2 and 3. Table 2 summarizes the main coefficient estimates from 10 different static panel data regressions, while Table 3 shows the results of 6 dynamic estimations. The key relationship is elasticity of labor productivity with respect to regional size, ϕ, controlling for attributes of the individual firms. For concreteness, we specify and simplify the empirical specification of Equation (4) in a static and a dynamic setting:

$$q_{it} = \alpha_i + \mathbf{x}'_{\mathbf{it}}\boldsymbol{\beta} + \varepsilon_{it} \quad t = 1, 2, \ldots.8. \qquad (5)$$

where $\mathbf{x}_{\mathbf{it}}$ in the static model (5) are regressors, α_i are firm-specific effects, and ε_{it} is an idiosyncratic error. The α_i are allowed to be correlated with the regressors $\mathbf{x}_{\mathbf{it}}$ in the fixed-effects model (FE), while they are assumed to be purely random in the random-effects (RE) model. In more flexible models, such as the Hausman-Taylor estimator, it is possible to specify an equation with regressors that are uncorrelated with α_i as well as regressors that are correlated with α_i. The fixed-effects vector decomposition estimator (FEVD) makes it possible to estimate the effect of time-invariant variables in a fixed-effects framework (Plümper and Troeger, 2007). We employ the FEVD estimator since regional population size changes slowly over time.

The dynamic specification is an autoregressive model of order 2 with $q_{i,t-1}$ and $q_{i,t-2}$ included together with the regressors $\mathbf{x}_{\mathbf{it}}$. In the model:

$$q_{it} = \gamma_1 q_{i,t-1} + \gamma_2 q_{i,t-2} + \mathbf{x}'_{\mathbf{it}}\boldsymbol{\beta} + \alpha_i + \varepsilon_{it} \quad t = 3, 4, \ldots 8 \qquad (6)$$

α_i is the fixed effect (also called unobserved heterogeneity). The two-step system GMM estimator relaxes the assumption that the regressors $\mathbf{x}_{\mathbf{it}}$ should be uncorrelated with ε_{it}. We specify skilled labor, capital investments and import and export activities to be endogenous regressors.

In both models the regressors $\mathbf{x}_{it}$ consist of the following variables: log physical capital, log skilled labor, log ordinary employees, log region size, dummy for import and export activities, corporate ownership structure, industry classification and time trend. Labour productivity is measured as value added over ordinary labor.

Four things should be noted. First, labor productivity, physical capital, and skilled labor are contemporaneous and lagged dependent variables, while regional size, ordinary labor, ownership structure, industry affiliation and time trend are contemporaneous variables. Second, the number of observations are reduced to 33,826 because the two first years of data are ‘lost’ in order to construct the lags of labor productivity, physical capital and skilled labor. Third, the zero autocorrelation of the first-difference errors are tested at order 1 and order 2. Fourth, in the GMM-model there are 128 instruments used to estimate 36 parameters, so there are 92 overidentifying restrictions to consider in the specification test.

4.1 Static model

In order to save space only the two key-variables for the purpose of the study are shown in Table 2 for the five estimators and the two size groups we are considering. On the left side of the table, the OLS (column 1) provides evidence that labor productivity is an increasing function of agglomeration. Firms located in larger regions have higher productivity, all else equal. We control for physical capital skilled labor, import-and export activities, corporate ownership structure (Foreign owned MNE, domestic MNE, domestic firms belonging to a group consisting of only domestic affiliates and non-affiliate firms), industry classification, time trend and number of employees for each of the three size groups. No differences in the agglomeration effect can be found between small firms and all firms.

>> TABLE 2 ABOUT HERE <<

Looking then at the fixed effects (FE) and random effects (RE) estimates, two different results emerge. The fixed effect regression shows no correlation between region size and firm productivity. In contrast, the coefficient estimates for both samples are highly significant and similar to the OLS-estimates in the RE-estimation. The Hausman test rejects the null hypothesis that the RE estimator is consistent. The unobserved firm-specific effects do appear to be correlated with the regressors, which means that they are endogenous and the FE is a more appropriate estimator. However, as discussed by Baltagi (2008) we should not automatically interpret a rejection of the null hypothesis in a Hausman test as a rejection of the RE-model as an adoption of the FE model, since there are very strong assumptions underlying the test.

The right side of Table 2 provides the results for the two more flexible panel data models. Column 4 shows significant and positive coefficient estimates for regional size for both samples using the FEVD model. The size of the estimates is somewhat larger for the total sample. Column 5 confirms the general pattern of the FEVD results. Using the 'not too random' Hausman-Taylor estimator, however, the relationship between agglomeration and productivity is slightly stronger for small firms.

4.2 Dynamic model

We now consider the parameter estimates of applying GMM to estimation of the Cobb-Douglas production function. We use the same data as in the static specification but with fewer observations due to the lag structure of the dynamic model. Table 3 summarizes six regression results describing the dynamic relationship between productivity and agglomeration.

The first part of Table 3 reports the results when using OLS . In the middle section the fixed-effects parameters are displayed. Our results for the two-step GMM model are shown on right side of the table. The main interest is on the GMM-results and we only use the two first estimators for a robustness check. The literature (Hsiao 1986 and others) suggests that the OLS-estimate of the dependent variable lagged one period will be biased upwards in the presence of individual-specific effects. Moreover, the FE is supposed to give an estimate of the lagged dependent variable that is downward biased since the panels are short (Nickell 1981). Nerlove (1999), Bond et al. (2001) and Roodman (2006) suggest that the consistent GMM-estimate of the lagged dependent variable can be expected to lie in the interval between the OLS and the FE estimates.

The first row of the table shows that the endpoints of this interval are between -0.07 and 0.49 for small firms and between 0.03 and 0.52 for the total sample. The GMM-estimates pass the 'bound test' of the lagged productivity indicating unbiasedness. Moreover, with two lags of the dependent variable, test statistics for auto-correlation of the error term are satisfactory for both GMM-regressions. The Sargan and Hansen tests show opposite results for small firms but the same for the total sample. The null hypothesis should not be rejected in either test. It is evident from the table that this is only the case in the Hansen test for firms with 10-25 employees. However, the tests should be interpreted with care. The GMM-specification includes 128 instruments for the first difference and the level equations and this might contribute to the bad Hansen-result for the total sample. The Sargan test is not as vulnerable to instrument proliferation as the Hansen test, but it is however not consistent unless we assume homoskedasticity. Thus, technically the Sargan test does not work since the model allows for heteroskedasticity. It should also be noted that no formal test has yet been developed for weak instruments in the system GMM (cf. Brülhart and Sbergami 2009).

>> TABLE 3 ABOUT HERE <<

Looking at the estimated parameters, our first finding confirms the persistent nature of firms' productivity performance: lagged productivity is a good predictor of contemporaneous productivity. A second finding is that the GMM-results are qualitatively very close to those previously presented in Table 2. There is a significant and positive relationship between agglomeration and productivity for small and large firms as well. Quantitatively, the parameter estimates are reduced somewhat when we control for lagged productivity.

A third and important finding may be derived from the fact that the dynamic GMM-model allows for a more causative interpretation, since we condition on previous productivity levels. Controlling for previous productivity levels, contemporaneous productivity is higher for firms located in larger regions. The results thus suggest that agglomeration phenomena not only correlate positively with firms' productivity, but also suggest that agglomeration contributes to higher productivity, and that the role of agglomeration phenomena does not seem to have a clear coupling to firm size.

5 CONCLUSIONS

This paper conducts a simple test for agglomeration economies at the level of individual firms. We augment a standard production function with a variable describing the size of the region a given firm is located in. Regional size is assumed to reflect the potential for agglomeration economies in a location. To deal with endogeneity issues in the relationship between agglomeration and labor productivity we employ a dynamic panel model and apply the two-step system GMM estimator. The paper adds to relatively limited micro-level studies of the relationship between agglomeration and productivity and provides evidence that agglomeration phenomena have a positive influence on firms' productivity.

Three questions are analyzed. First, are firms located in larger regions more productive, when controlling for attributes of individual firms? Second, is there any difference in the relationship between region size and productivity between small and large firms? Third, can the marginal effect of agglomeration on firms' productivity be given a causal interpretation? Three things emerge. First, firms located in larger regions are more productive when controlling for size, human capital, physical capital, ownership structure, industry classification and time trend. The results show that there is a positive relationship between the size of a region and labor productivity at the level of individual firms. This relationship holds although we control for several attributes of individual firms that are likely to influence their labor productivity. Second, our results suggest that firms become more productive by being located in agglomerations. Controlling for previous productivity levels, contemporaneous productivity is higher for firms located in larger regions. Thus, the results suggest a positive learning effect. Third, the role of agglomeration phenomena does not seem to have a clear coupling to firm size.

The research in this paper may be extended in several ways. For example, we have used regional size as a general indicator of the potential for agglomeration economies. The analysis may be extended to comprise measures of industry concentration and assess the relative importance of localization and urbanization economies. Moreover, the type of analysis conducted here yields indirect evidence of agglomeration economies and we cannot say much about the mechanisms that drive the relationship. An avenue for future research is thus to try to untangle the various mechanisms behind agglomeration economies.

REFERENCES

Åberg, Y (1973), “Regional Productivity Differences in Swedish Manufacturing”, *Regional Science and Urban Economics*, 3, 131-156

Almeida, P. and B. Kogut, (1999), ”Localization of Knowledge and the Mobility of Engineers in Regional Networks,” *Management Science*, 45, 905–917.

Anderson, T. and Hsiao, C. (1982), “Formulation and Estimation of Dynamic Models Using Panel Data”, *Journal of Econometrics*, 18, 67-82.

Andersson M and C. Karlsson (2007), “Knowledge in Regional Economic Growth – the role of knowledge accessibility”, *Industry and Innovation*, 14, 129-149

Andersson, M and B. Johansson (2008), “Innovation Ideas and Regional Characteristics – innovations and export entrepreneurship by firms in Swedish regions”, *Growth and Change*, 39, 193-224

Arellano. M., Bond, S. (1991), “Some test of specification for panel data: Monte Carlo evidence and application to employment equations”, *The Review of Economic Studies,* 58: 277-97.

Arelleno, M., Bover, O. (1995), “Another Look at Instrumental Variable Estimation of Error Component Models”, *Journal of Econometrics*, 68: 29-51.

Audretsch, D and M. Feldman (1996), ”R&D Spillovers and the Geography of Innovation and Production”, *American Economic Review*, 86, 630-640

Baldwin, R and T. Okubo (2006), “Heterogeneous Firms, Agglomeration and Economic Geography: spatial selection and sorting”, *Journal of Economic Geography,* 6, 323-346

Balestra, P., M. Nerlove (1966). “Pooling Cross Section and Times series data in the estimation of a Dynamic Model: the Demand for Natural Gas”, *Econometrica*, 24, pp 585-612.

Baltagi, B.H. (2008) *Econometric Analysis of panel Data* forth edition, Wiley

Barney, J. (1991) “Firm Resources and Sustained Competitive Advantage”, *Journal of Management*, 17, 99-120

Bartel, A.P and Lichtenberg F.R. (1987), “The Comparative Advantage of Educated Workers in Implementing New Technology”, *Review of Economics and Statistics,* 69, 1-11

Bernard, A. B., and J. B. Jensen (1995). “Exporters, Jobs and Wages in US Manufacturing 1976–1987”, *Brookings Paper on Economic Activity*: 67–119

Blundell RS, Bond S (1998), “Initial Conditions and Moment Restrictions in Dynamic Panel data Models”, *Journal of Econometrics,* 87: 115-43.

Braunerhjelm, P. and B. Borgman, (2004), ”Geographical Concentration, Entrepreneurship and Regional Growth: Evidence from Regional Data in Sweden, 1975-99,” *Regional Studies*, 38, 929–947.

Brülhart, M. and F. Sbergarmi, (2009), “Agglomeration and Growth – cross-country evidence”, *Journal of Urban Economics*, 65, 48-63

Brülhart, M. and N. Mathys, (2008), ”Sectoral Agglomeration Economies in a Panel of European Regions,” *Regional Science and Urban Economics*, 38, 348–362.

Ciccone (2002), ”Agglomeration Effects in Europe,” *European Economic Review*, 46, 213–227.

Ciccone, A. and R.E. Hall, (1996), ”Productivity and the Density of Economic Activity,” *American Economic Review*, 86, 54–70.

Cohen, W. and D. Levinthal, (1990), ”Absorptive Capacity: A new perspective on learning and innovation,” *Administrative Science Quarterly*, 35, 128–152.

Dachs, B., B. Ebersberger and H. Lööf (2008), “The Innovative Performance of Foreign-owned Enterprises in Small Open Economies”, *Journal of Technology Transfer*, 33, 393-406

Duranton, G. and D. Puga, (2000), “Diversity and Specialization in Cities: why, where and when does it matter?”, *Urban Studies*, 37, 533-555

Duranton, G. and D. Puga, (2004), ”Micro-foundations of Urban Agglomeration Economies,” in: J. V. Henderson & J. F. Thisse (ed.), *Handbook of Regional and Urban Economics*, Elsevier. 2063–2117

Ejerno, O and U. Gråsjö (2008), “The Effects of R&D on Regional Invention and Innovation”, CIRCLE WP Series, 2008-03

Fallick, B., C.A Fleischmann and J.B. Rebitzer, (2006), ”Job Hopping in Silicon Valley: Some Evidence Concerning the Micro-Foundations of a High Technology Cluster,” *Review of Economics and Statistics*, 88, 472-481

Frenken, K., F.G Van Oort, and T, Verburg (2007), “Related Variety, Unrelated Variety and Regional Economic Growth”, *Regional Studies*, 41, 685-697

Fujita, M. and J.-F. Thisse, (2002), *Economics of Agglomeration. Cities, industrial location, and regional growth*. Cambridge, Cambridge UK: Cambridge University Press.

Fujita, M., P. Krugman and A. Venables (1999), *The Spatial Economy*, MIT Press, Cambridge

Glaeser, E. and D. Maré, (2001), ”Cities and Skills,” *Journal of Labor Economics*, 19, 316–342.

Greenaway, D and R. Kneller (2007), “Firm Heterogeneity, Exporting and Foreign Direct Investment, Economic Journal, 117, 134-161

Griliches, Z. And J. Mairesse (1984). “Productivity and R&D at the Firm Level”, in Z. Griliches (ed.), *R&D Patents and Productivity*, Chicago, University of Chicago Press: 339-375.

Hausman, J.A and W.E Taylor (1981) “Panel data and unobservable individual effects”. *Econometrica* 49, 1377-1398.

Henderson, V.J (2003), “Marshall’s Scale Economies”, *Journal of Urban Economics*, 53, 1-28

Hesley, R. and W. Strange, (1990), ”Matching and Agglomeration Economies in a System of Cities,” *Regional Science and Urban Economics*, 20, 189–212.

Holtz-Eakin, D., Newey, W., Rosen, H.S. (1988). “Estimating Vector Autoregressions with Panel Data”. *Econometrica*, 56: 1371-95.

Hoover, E (1937), *Location Theory and the Shoe and Leather Industries*, Harvard University Press, Cambridge

Jaffe, A.B., Trajtenberg, M. and R. Henderson, (1993), ”Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations,” *Quarterly Journal of Economics*, 108, 577–598.

Johansson, B and H. Lööf (2008), “Innovation Activities Explained by Firm Attributes and Location”, *Economics of Innovation and New Technology*, 17, 553-552

Karlsson, C and L. Pettersson (2005), ”Regional Productivity and Accessibility to Knowledge and Dense Markets”, CESIS WP 32

Kim, S (1990), “Labor Heterogeneity, Wage Bargaining and Agglomeration Economies”, *Journal of Urban Economics*, 28, 160-177

Klaesson, J and H. Larsson (2008), ”Wages, Productivity and Industry Composition – agglomeration economies in Swedish regions”, Mimeograph, Jönköping International Business School

Klaesson, J. and B. Johansson (2008), “Agglomeration Dynamics of Business Services”, CESIS WP 153

Krugman, P (1991), *Geography and Trade*, The MIT Press, Cambridge

Lucas, R.E. Jr, (1988), ”On the Mechanics of Economic Development,” *Journal of Monetary Economics*, 22,3–42.

Marshall, A. (1920), *Principles of Economics*, MacMillan, London

Moen, J., (2005), ”Is Mobility of Technical Personnel a Source of R&D Spillovers?,” *Journal of Labor Economics*, 23, 81–114.

Moomaw, R.L (1981), “Productive Efficiency and the Region”, *Southern Economic Journal*, 48, 344-357

Moomaw, R.L (1983), “Is Population Scale a worthless Surrogate for Business Agglomeration Economies?”, *Regional Science and Urban Economics*, 13, 525-545

Moretti, E (2004), ”Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions,” *American Economic Review*, 94, 656–690.

Nakamura, R (1985), “Agglomeration Economies in Urban Manufacturing Industries: a case of Japanese cities”, *Journal of Urban Economics*, 17, 108-124

Ohlin, B. (1933), *Interregional and International Trade*, Harvard University Press, Cambridge

Penrose, E.T (1959), *The Theory of the Growth of the Firm*, John Wiley & Sons, New York

Phaffermayr, M and C. Bellak (2002), “Why Foreign-owned Firms are Different: a conceptual framework and empirical evidence for Austria”, in R. Jungnickel (ed.), *Foreign-owned Firms: are they different?*, Palgrave Macmillan, 13-57

Plümper, T and V. Troeger (2007), “Efficient Estimation of Time-Invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed Effects”, *Political Analysis*, 15, 124-139

Quigley, J.M., (1998), ”Urban Diversity and Economic Growth,” *Journal of Economic Perspectives*, Vol. 12(2), s. 127–138.

Rauch, J.E., (1993), ”Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities,” *Journal of Urban Economics*, 34, 380–400

Rice, P., Venables, A.J. and E. Patacchini, (2006), ”Spatial Determinants of Productivity: analysis for the regions of Great Britain,” *Regional Science and Urban Economics*, 36, 727–752.

Romer, P.M., (1990), ”Endogenous Technological Change,” *Journal of Political Economy*, 98, 71–102.

Roodman. D, (2006). “How to Do xtabond: An Introduction to “Difference” and ”System” GMM an Stata”. *CGDEV Working Paper 103*.

Rosenthal, and W. Strange (2004), “Evidence on the Nature and Sources of Agglomeration Economies”, in: J. V. Henderson & J. F. Thisse (ed.), *Handbook of Regional and Urban Economics*, Elsevier p. 2118–2147.

Segal, D (1976), “Are There Returns to Scale in City Size?”, *Review of Economics and Statistics*, 58, 339-50.

Shefer, D (1973), "Localization economies in SMSAs: a production function approach", *Journal of Regional Science*, 13, 55-64

Sveikauskas, L (1975) "The Productivity of Cities", *Quarterly Journal of Economics*, 89, 393-413

Varga, A. (1997), *University Research and Regional Innovation: a spatial econometric analysis of academic technology transfers*, Kluwer Academic Publishers, Boston

Wagner, J. (2007), "Exports and Productivity - a survey of the evidence from firm level data", *World Economy*, 30, 60-82

Wheeler, C.H., (2006), "Cities and the Growth of Wages among Young Workers: Evidence from the NLSY," *Journal of Urban Economics*, 60, 162–184.

Windmeijer, F. (2000). Efficiency Comparisons for a system GMM Estimator in Dynamic Panel data Models. In Heijmans R.D.H., S.S.G. Pollock and A. Satorra (eds.), *Innovations in Multivariate Statistical Analysis. A Festschrift for Heinz Neudecker*, Kluver Akademic Publishers: 175-84.

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

Table 1: Summary statistics

	Mean	Std. Dev	Min	Max
Year	2000	2.27	1997	2004
Industries, Nace 3	265	560	151	372
Log region size[1]	10.96	1.47	5.91	13.32
Log labor productivity	6.05	0.51	-2.83	12.28
Log physical capital	8.34	2.000	0	16.68
Employment	91	445	10	23,321
Skilled labor	8	90	0	8,111
Ordinary labor	83	375	1	23,009
Non-affiliate	0.37	0.48	0	1
Uninational	0.31	0.46	0	1
Domestic MNE	0.20	0.39	0	1
Foreign MNE	0.12	0.32	0	1
Export and import	0.58	0.49	0	1

Notes:
Number of observations is 54, 278 for the eight years considered and the total number of firms is 6,898.
All firms in the Swedish manufacturing sector with 10 or more employees and observed over the whole period 1997-2004 are included in the balanced panel data.
(1) The difference between total regional employment and number of employees in the observed firm.

Table 2 Static Models

The elasticity of log Labour productivity with respect to log region size

	Pooled OLS	FE	RE	Hausman Prob>chi2	FE-VD	HAUSMAN TAYLOR
10-25 emp	0.028*** (0.002)	0.010 (0.0131)	0.028*** (0.003)	0.000	0.033*** (0.001)	0.040*** (0.004)
10-emp	0.029*** (0.002)	0.002 (0.010)	0.029*** (0.003)	0.000	0.039*** (0.001)	0.031*** (0.003)

Notes:

Significant at the 1% level ***, 5% level ** and 10% level *. Standard errors within parentheses

The table summarizes the key coefficient estimates for five different models and 10 different regressions. All firms in the Swedish manufacturing sector with 10 or more employees. Total number of observations is 54, 278 when all firms (10 or more employees) are considered and 28, 351 in the sub-sample consisting only of small firms (10-25 employees).

Additional covariates are log physical capital, log skilled labor, log ordinary labor and import and export dummy, 4 corporate ownership structure dummies, 14 industry dummies and 8 year dummies (see Equation 4 in the main text).

The Hausman test tests the null hypothesis that the difference in coefficients between FE and RE models is not systematic. The results reject the null hypothesis that RE provides consistent estimates.

FEVD is a fixed effects vector decomposition estimator which makes it possible to include time-invariant regressors in a fixed-effects setting, and the Hausman-Taylor estimator allows for regressors both correlated and uncorrelated with unobserved firm-specific effects.

Table 3 Dynamic models

The elasticity of log Labour productivity with respect to log region size

Model	Pooled OLS		FE		2-STEP system GMM	
Firm size	10-25	All (10-)	10-25	All (10-)	10-25	All (10-)
Log lab prod, lag (1)	0.492*** (0.028)	0.515*** (0.017)	-0.072 (0.044)	0.034*** (0.006)	0.286*** (0.034)	0.328*** (0.023)
Log lab prod, lag (2)	0.246*** (0.023)	0.245*** (0.014)	-0.083*** (0.017)	-0.073*** (0.006)	0.074*** (0.020)	0.082*** (0.018)
Log region size	0.005*** (0.001)	0.005*** (0.001)	0.014 (0.018)	-0.002 (0.880)	0.017*** (0.003)	0.017*** (0.003)
AR (1)					0.000	0.000
AR (2)					0.259	0.888
Sargan					0.000	0.000
Hansen					0.403	0.000
Observations	14, 437	33,140	14, 437	33,140	14, 437	33,140

Notes: Significant at the 1% level ***, 5% level ** and 10% level *. Standard errors within the parentheses
The table summarizes the key coefficient estimates for 6 different regressions. All firms in the Swedish manufacturing sector with 10 or more employees. Since the two first years of data are 'lost' in order to construct the lag structure of the model, the total number of observations is 33,140 when all firms (10 or more employees) are considered and 14,437 in the sub-sample consisting only of small firm (10-25 employees) The regressors are:
(a) Lag 1 and lag 2 of log labor productivity
(b) Contemporaneous log region size
(c) Contemporaneous and lag 1 of log physical capital, log skilled labor and a dummy for export and import
(d) Contemporaneous log ordinary labor, 4 corporate ownership structure dummies, 14 industry dummies and 8 year dummies

The test statistics indicate no problem with first order autocorrelation (null hypothesis is autocorrelation) or second order autocorrelation (null hypothesis is *no* autocorrelation). The Sargan and Hansen tests show opposite results for small firms, but the same for the total sample. The Sargan test is not as vulnerable to instrument proliferation as the Hansen, but it is also not consistent unless we assume homoskedasticity, which is unrealistic in the used panel data. Thus, technically the Sargan test does not work. Since the GMM-specification includes 128 instruments for the first difference and the level equations and this might contribute to the bad Hansen-result for the total sample. It should also be noted that no formal test has yet been developed for weak instruments in the system GMM.
The Hansen test on the validity of the instruments is not rejected because p=0.40>0.05 when the sub-sample of only small firms is considered. However, the Hansen statistics are bad for the total sample. Although the Hansen statistic is robust, it can be weakened by many instruments.

APPENDIX

Figure A1 shows the relationship between regional employment size and average wages per employee across the 72 Swedish functional regions.[6] The dots denote the actual observations whereas the bold line is a fitted line obtained with a simple OLS estimator with which regional size is regressed on the average wage per employee. As can be seen from the figure, there is a positive relationship between the size of a region and average wage per employee at the aggregate level in Sweden. The equation in the bottom-right corner in the figure informs that the elasticity between sheer employment size and average wage per employee at the aggregate level is 0.045.

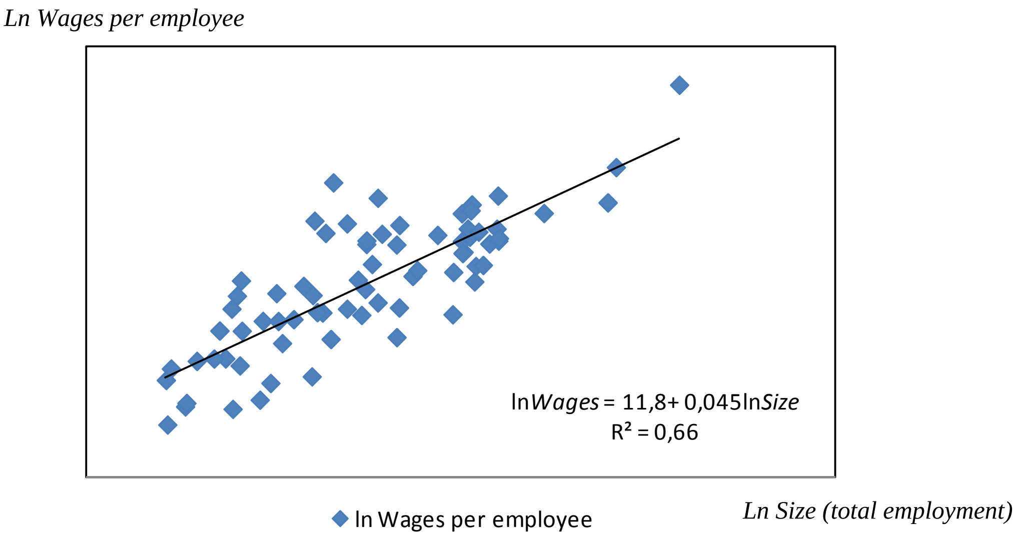

Figure A1. *The relationship between regional size (total employment) and wages per employee across Swedish functional regions (N=72) in 2004.*

Figure A2 shows that regional size is positively associated with diversity of the regional economy, as indicated by the number of sectors (5-digit according to the NACE classification) in the region. Furthermore, Figure A3 shows that regional size is positively associated with the average education level of the employees in the region. The average education level is defined as the number of employees with a long university education ($\geq$ 3 years) as a fraction of the total number of employees.

---

[6] The average wage per employee is calculated as the ratio between the total wage-sum and total employment.

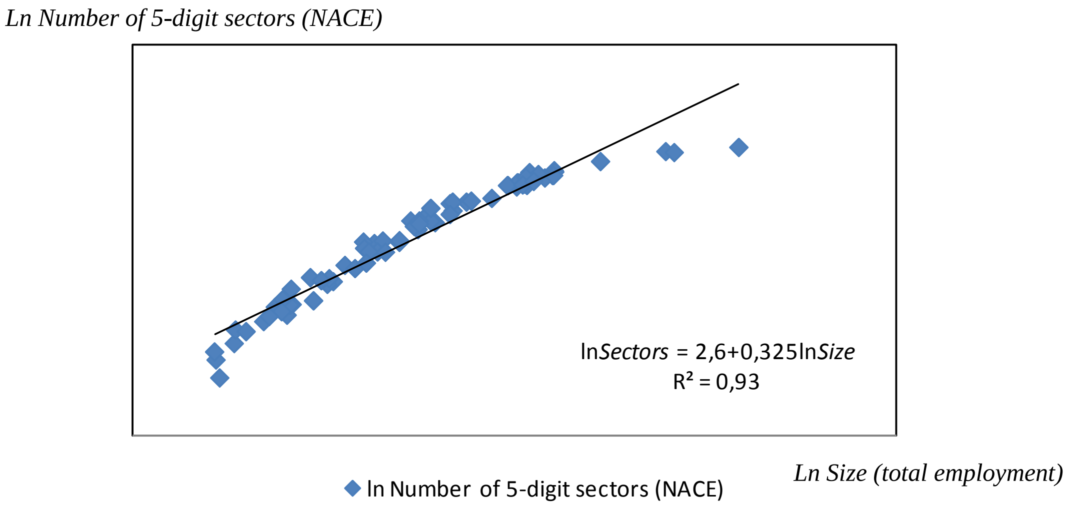

Figure A2. *The relationship between regional size (total employment) and number of 5-digit sectors according to the NACE classification across Swedish functional regions (N=72) in 2004.*

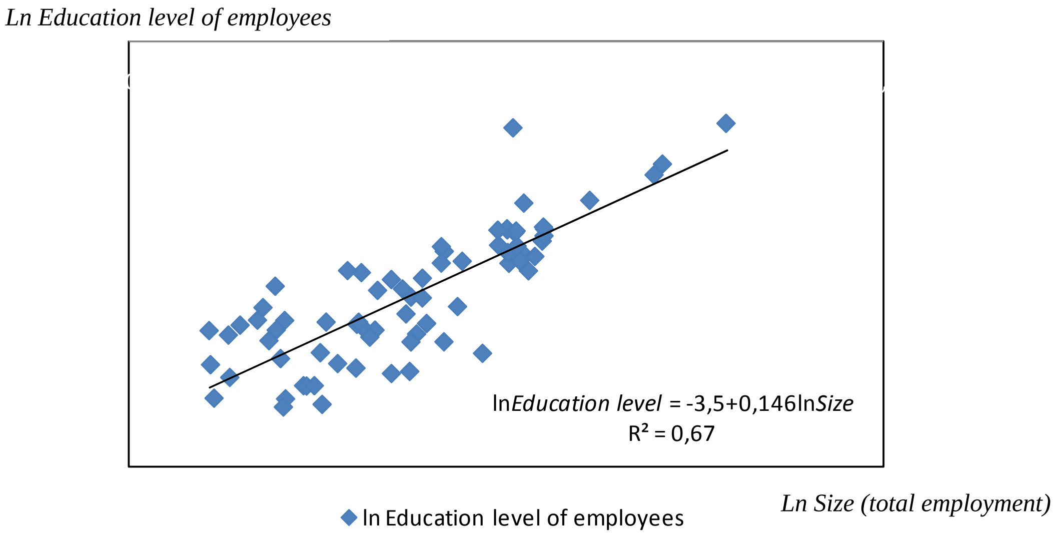

Figure A3. *The relationship between regional size (total employment) and employees with a long university education ($\geq$ 3 years) as a fraction of total employment across Swedish functional regions (N=72) in 2004.*