

**CESIS** Electronic Working Paper Series

**Paper No. 203**

**Wages, Productivity and Industry Composition  
– agglomeration economies in Swedish regions**

**Johan Klaesson and Hanna Larsson**

CESIS and Jönköping International Business School (JIBS)

June 2009

# Wages, Productivity and Industry Composition

## - Agglomeration Economies in Swedish Regions

Johan Klaesson  
Hanna Larsson  
Jönköping International Business School  
Jönköping University  
PO Box 1026  
551 11 Jönköping  
Sweden

E-mail: [johan.klaesson@jibs.hj.se](mailto:johan.klaesson@jibs.hj.se); [hanna.larsson@jibs.hj.se](mailto:hanna.larsson@jibs.hj.se)

---

### Abstract

It is a well known fact that wages have a tendency to be higher in larger regions. The source of the regional difference in wages between larger and smaller areas can be broadly divided into two parts. The first part can be attributed to the fact that regions have different industrial compositions. The second part is due to the fact that average regional productivity differs between regions. Using a decomposition method, akin to shift-share, we are able to separate regional wage disparities into an industrial composition component and productivity component. According to theory it is expected that productivity is higher in larger regions due to different kinds of economies of agglomeration. Also, larger regions are able to host a wider array of sectors compared to smaller regions. Output from sectors demanding a large local or regional market can only locate in larger regions. Examples of such sectors are e.g. various types of advanced services with high average wages. The purpose of the paper is to explain regional differences in wages and the productivity and composition components, respectively.

The paper tests the dependence of wages, productivity and industrial composition effects on regional size (using a market potential measure). In the estimation we control for regional differences in education, employment shares, average firm size and self-employment. Swedish regional data from 2004 are used. The results verify that larger regions on average have higher wages, originating from higher productivity and more favorable industry composition.

Keywords: Agglomeration Economies, Regions, Wages, Productivity, Industrial Composition, Sweden.

JEL classification codes: C21, O18, R10, R12, J31.

# 1. Introduction

In many places in the literature it is emphasized that concentration is a prevalent and ubiquitous feature of the geographical distribution of economic activity. This is true at all geographic levels (regional, national, international) and at all levels of industrial aggregation; with the possible exception of some primary industries. This general observation is a strong indication that there are economies connected to concentration of activities, where larger regions are preferred over smaller. There are many studies that are concerned with the sources of these economies and the effects that they have on efficiency and growth. One such effect is the translation of efficiency to wage levels. Research connected to the prevailing wage differences within nations is important for a better understanding of regional development, which in turn will have an effect on national welfare.

The regional difference in wages between larger and smaller regions can be broadly attributed to two sources. The first source is the fact that average regional productivity differs between regions. The second source is that the industrial composition differs between regions.

According to economic theory it is expected that productivity is higher in larger regions due to different kinds of economies of agglomeration. The conventional macroeconomic view is that productivity growth in turn drives wage growth. Also, larger regions are able to host a larger array of sectors compared to smaller regions. Output from sectors demanding a large local or regional market can only locate in larger regions.

Given the above discussion it is clear that there exist a basic link between market-size and diversity. The extent of diversity is limited to an upper level due to the presence of fixed costs in the individual intermediate input-producing firms. This relationship implies that the size of a region will decide the degree of diversification in intermediaries. This has an impact on the productivity of all firms in the region. Thus, agglomeration provides a large market and thereby enabling the support of a wide variety of differentiated inputs. As a result of economies of scale and increasing returns, the productivity of firms in the larger agglomerations can be expected to be higher. This gives a rationale for why many firms tend to locate in agglomerations.

This leads us to the purpose of this paper where the aim is to explain the regional differences in wages, productivity and industrial composition. The paper tests the dependence of wages, productivity and industrial composition effects on regional size which in turn is measured with a market potential measure. In addition, in the estimation we control for regional differences in education, employment shares, average firm size and self-employment.

There exists a large and expanding literature concerning the regional differences in productivity and wage rates. Ciccone and Hall (1996) was not the first study that examined the relationship between productivity and economies of agglomeration, but surely one of the most influential. In this seminal study agglomeration was approximated by economic density, calculated as employment per acre.

Since then several studies have deepened and broadened the literature in these matters. Glaeser and Maré (2001) for example used wage data to investigate the wage premium paid to workers in larger cities.

More recent studies by Ciccone (2002); Rice et al. (2006) and Combes et al. (2008) argue that spatial differences in income can be ascribed to productivity differences. Higher or lower productivity is in turn assumed to be reflected in value added per unit of labor, or in earnings. Their analyses have focused on three main explanations for spatial differences in wages and productivity: (i) agglomeration economies (ii) skill or industry composition and (iii) exogenous regional characteristics. This paper adheres to this literature, but extends it by

introducing accessibility to Gross Regional Product (GRP) as a measure for regional market potential.

The paper is structured as follows. It continues in section two with a more in-depth discussion on the relationship between agglomeration, productivity and industry composition. The relationship of the paper to the existing literature is discussed. Also, the selection of the control variables is discussed. In section three the data used is presented in conjunction with a descriptive analysis. Section four present the empirical results from the estimated models. Section five offers some additional tests of the robustness of the results. The closing section summarizes the conclusions and gives suggestions for interesting future research.

## 2. Regional Differences in Wages

As has been briefly discussed in the introduction, regions with larger market size on average can be expected to have a higher average wage level, due to different forms of scale economies. This section will continue the arguments as well as motivate the various control variables used in the present study. Furthermore, a short description of the features of the Swedish labor market is also offered.

### *Agglomeration economies*

Economies of agglomeration means that cost reductions occur because economic activities are located in proximity to each other. The concept itself goes back to Weber ([1909] 1929) and Marshall (1920). Economies of agglomeration are also sometimes referred to as external economies of scale. That is, the agglomeration economies are external to the individual firms. Consequently, over time an abundance of concepts and elaborations of the general concept of agglomeration economies have evolved. (see McDonald, 1997)

Ohlin (1933), categorized economies of agglomeration into four types: (i) economies of scale within the firm, (ii) localization economies, which are external to the individual firm and arise from the size of the local industry to which the firm belongs, (iii) urbanization economies, which are more general in that it refers to cost reductions that are external to the local industry and arise from the size of the local economy as a whole and (iv) inter-industry linkages, which arise from transportation cost savings in purchases of intermediate inputs.

Before Ohlin's categorization, Marshall (1920) had already suggested three reasons for cost reductions to occur in an agglomeration. The first reason is due to knowledge spillovers, i.e. the spread of advances/innovations in production is assumed to be influenced by distance. The second explanation is the existence of a broad market for specialized skills. Specialized employees gain from having nearby access to many jobs that match their specialty. The benefits also go in the other direction. Firms benefit from having accessibility to a large pool of specialized employees matching their requirements. The third reason is the existence of backward and forward linkages between firms that are located in proximity to each other. The proximity means that the transport of inputs and outputs will be relatively inexpensive.

Hoover (1937), later made use of Ohlin's second and third categories, localization economies and urbanization economies, as he elaborated and popularized the two concepts. Hoover's definitions are the ones most often used to this present day. Localization economies can be captured by independent and similar small businesses in the form of positive external economies by locating in proximity of each other. This proximity between similar firms creates a market for specialized services, which can be provided by other independent firms as long as the total demand is great enough. In addition, it is possible to build up a local pool of labor with specialized skills. In this way, the businesses can take advantage of scale economies in production without resorting to large plants. The actual advantage over large-

scale plants is that independence and flexibility are retained. Urbanization economies can bring about scale advantages that benefit a wider group of businesses. If, for example, general manufacturing increases in a particular area, the business services and workforce improve in size and variety. These advantages do not only relate to one sector of industry but to all. If educational standards improve, or if trucking companies expand their route network, every business will benefit.

Further, agglomeration economies can be differentiated as being predominantly static or dynamic in nature. In general a static economy of agglomeration is associated with a onetime shift in costs or productivity whereas a dynamic economy of agglomeration influences costs and productivity through time. Dynamic agglomeration economies are usually connected to the production and use of knowledge. This is at the heart of the endogenous growth theory (Romer, 1986; Lucas 1988). Knowledge spillovers are an essential ingredient in this theory and since spillovers are facilitated by proximity between people and firms concentration and agglomeration are key concepts.

The empirical literature that aims to identify the sources of economies of agglomeration is reviewed in Rosenthal and Strange (2004).

### *Wages and regional characteristics*

As hinted above some sources of regional wage differences work through scale economies. But there are several different ways for those to manifest themselves. For instance, differences in wages across regions could directly reflect spatial differences in the skill level and composition of the workforce. Theory argues that the division of labor, which leads to productivity gains, is limited by the extent of the market. Then naturally, workers in larger markets may enjoy higher wages because of greater possibilities for the division of labor.

According to Combes et al. (2008) there are reasons to believe that workers may spread across employment areas so that the measured and un-measured productive abilities of the local labor force may vary. Industries are not uniformly spread across regions and they require different combinations of labor skills. Accordingly, one should expect a higher mean wage in areas specialized in more skill-intensive industries. This brings us into the discussion on industry composition and wages.

Different industries or sectors differ in their ability to pay high wages. The difference in this ability may come from the competitive situation in that particular industry or the availability of factor inputs. Since regions differ in their industry structure, it is likely that some of the regional wage differences can be explained by these structural differences. Thereby some part of the wage differences across areas reflects spatial differences in industry structure. These differences may not be directly connected to regional productivity differences.

Combes & Overman (2004) argues that failure to control for heterogeneous skill composition between regions may significantly limit the interpretation that can be given to regional differences in labor productivity or wages. Hence, higher wages or productivity might not reflect real externalities if not taking labor skill composition into account.

Non-market interactions such as technological externalities, and in particular those originating from human capital, may also play an important role in explaining wage differences between regions (Rauch, 1993). Sianesi and van Reenen (2003) presents a detailed literature review which provide strong evidence in favor of the view that the amount of human capital is positively correlated to productivity. Rauch (1993) argues that the level of human capital is a local public good, hence, regions or cities with higher levels of education attainment should consequently have higher wages. Rauch base this argument on 1980 US

cross-sectional data on a regional level. Glaeser & Maré (2001) suggest that urban workers are more productive because cities enable workers to accumulate more human capital. This accumulation can work through more effective knowledge spillovers, proximity between skilled workers and between knowledge-intensive firms, in regions where knowledge are abundant and dense. These spillovers rely on face-to-face contacts that are more likely to occur in such knowledge-intensive regions. Knowledge can be transmitted either as a by-product that is non-intentional or in more formal types of collaborations.

Previous studies point to yet another regional characteristic that have a positive effect on productivity and wage levels. This characteristic is firm size. The size of a firm influences the possibility to exploit internal economies of scale, and as a consequence, positively influence labor productivity. Extensive research in the field confirms that larger firms tend to pay higher wages than smaller ones; see More (1911), later confirmed by Mellow (1982); Brown and Medoff (1989); Troske (1999) among several others. Average firm size in a region is, hence, an important component when explaining regional wage variation. According to Glaeser et al. (1992) firm size may also affect the level of productivity if it is related to market power, where the effect itself is ambiguous a priori. A large monopolistic firm may have higher incentive to conduct R&D due to a higher probability of retrieving the returns (Romer 1990) while at the same time small firms may experience a higher market pressure to innovate in order to stay competitive (Porter, 1990).

Wages should logically also be influenced by the level of (un-)employment in the region. If the scarcity of labor increases wages go up, while if it decreases they will fall. In 1994 Blanchflower and Oswald presented their book "The Wage Curve", where they established a negative relation between unemployment and wages across regions over time.\* Blanchflower and Oswald (1994) found that the employment elasticity of wages is, for most countries, approximately -0.1, with the US and UK being the main focus of the study. Since then several studies have been presented with mixed conclusions. A study by Albæk, Asplund et al. (2000), analyzing pooled data on regional wage formation in the Nordic countries during the 1990's, systematically deviate from the results presented by Blanchflower and Oswald (1994). Even though Albæk, Asplund et al. (2000) found a significant negative relationship between real wages and unemployment at the regional level, this relationship becomes unstable once accounting for regional fixed effects. The authors explain these results by the historically strong labor unions that negotiate wages at the national level in the Nordic countries. This negotiating model implies that there is only a small local wage drift influenced by the regional labor market conditions. According to Nijkamp & Poot (2005) the variation in findings within this field is a response to heterogeneity among the studies in terms of the type of data used and differences in model specification. In order to verify this Nijkamp & Poot performed a meta-analysis revealing that the wage curve is a robust empirical phenomenon. However, they also found clear evidence of a publication bias and when correcting for this bias the elasticity becomes less than -0.07.

Some emphasis should also be put on the rate of self-employment when explaining regional variations in wages. According to research, self-employment can work in dual ways when influencing regional wages. On the one hand high unemployment rates in a region may cause entrepreneurial activities, such as self-employment, to increase. This is referred to as a so called refugee effect, since the prospect of other forms of employment is small. The relationship between wage rates and unemployment is expected to be negative. On the other hand, higher levels of self-employment may be a sign of true entrepreneurial ventures, referred to as a entrepreneurial effect, which, in the long-run, will reduce unemployment

---

\* Blanchflower, D. G. and A. J. Oswald (1990). "The Wage Curve." *The Scandinavian Journal of Economics* 92(2): 215-235. refer to 16 previous studies in this topic between 1985 and 1990. The earliest study is by Bills 1985, who also supported an unemployment – wage elasticity of -0.1.

(Thurik, Carree et al. 2007) and hence have positive effect on the regional wage level. There exists considerable theoretical and empirical support for both of these views.

While Oxenfeldt (1943) argues that low prospects for wage-employment will cause a higher self-employment level, Audretsch and Fritsch (1994) on the contrary found that this relation is negative. In a recent study Thurik, Carree et al. (2007) tested this ambiguity by using a two-equation vector autoregression model, for a panel of 23 OECD countries over the period 1974-2002. The result reveal that changes in unemployment have a positive impact on changes in self-employment rates, while simultaneously changes in self-employment rates have a negative impact on unemployment. However, the latter effect is stronger than the first one.

In addition, due to the assumption of total labor mobility the role of amenities will according to Roback (1982) also influences regional wage differences. In Roback's model, worker utility depends on the wage level, the price of land, and a vector of potential lifestyle or amenity characteristics. Wages and rents must adjust to equalize utility in all occupied locations. Otherwise some workers would have an incentive to move. Roback (1982) and more recent studies such as Graves et al. (1999) have included variables such as: education levels, crime rates, number of people living under the poverty line, congestion levels, local unemployment, weather conditions, population density and size etc., as proxies for the role played by amenities when investigating regional wage differences.

Criticism towards the "city size wage premium" has also been raised since it is plausible to imagine that workers with above average ability and motivation have a "lifestyle" preference for living in larger regions. Glaeser and Maré (2001) however, reject the argument that such omitted ability bias explains the city wage premium. They do so because including direct measures of ability (the AFQT test), instrumenting for place of birth, and then studying wages all fails to eliminate this wage premium.

There is also ample evidence of inter-industry wage differences that are related to institutional features such as unionization and not to productivity (Wheaton & Lewis, 2002) or industry composition.

### *The functioning of the Swedish labor market*

To enable us to fully interpret the results it is of some importance to take a closer look at the Swedish labor market and its dynamics. At the heart of "the Swedish model" is a system of collective bargaining between trade unions and employers' associations. Sweden distinguishes itself from most other countries in that the level of membership unions and other labor markets organizations is extremely high. Due to historical traditions the Swedish *labor market has become centralized*. Central wage agreements are set by the labor unions and representatives from the employers to obtain low levels of income differences compared to most other countries.

Calmfors (1993a; b) argues that a highly centralized wage setting system causes wage levels to increase considerably slower compared to not so centralized systems. This leads to higher employment levels, which is at the heart of the Swedish model.

Calmfors (1993a; b) further states that the highest level of real wage increases is obtained in a system which lies in-between decentralization and centralization, i.e. wage agreements at the industry level. Up until 1983 the Swedish wage setting system was a textbook example of a highly centralized negotiation scheme (Lindgren, 2006). Since then Sweden has moved towards a more decentralized system. At present, once framework agreements are set centrally for the corresponding employment sectors, individual collective agreements can be negotiated between employers and trade unions at a local level. A standard

collective agreement is therefore not automatically applicable to all regions or to an entire industrial sector. The system is thus more flexible than before.

### 3. Data and Descriptive Analysis

As stated in the theoretical background, we assume that variation in average earnings mainly comes from two sources; (1) differences in the wage rates paid to workers in a given sector arising from productivity deviations across regions, (2) and differences in the industrial composition between regions. To be able to discriminate between these sources we use a method applied in Rice et al. (2006). Analogous to that study we decompose regional wages into two components, referred to as the productivity and industry composition indices. We do this in order to assess the impacts on these indices by agglomeration effects as well as the different control variables presented in the previous section. Swedish regional data from 2004 are used for this study, where the definition of a region is municipalities with local governments. There are 290 such regions in Sweden.

The decomposition is described in the following seven equations. We start with four definitions (eq. 3.1 to 3.4).

$$w_r^s = \text{wage in sector } s \text{ in region } r \quad (3.1)$$

$$e_r^s = \text{employment in sector } s \text{ in region } r \quad (3.2)$$

$$E_r = \sum_s e_r^s = \text{total employment in region } r \quad (3.3)$$

$$\sigma_r^s = \frac{e_r^s}{E_r} = \text{employment share in sector } s \text{ in region } r \quad (3.4)$$

Next, using the definitions in equation 3.1 and 3.2 we construct a measure for the average wage level in sector  $s$ .

$$\bar{w}^s = \frac{\sum_r e_r^s w_r^s}{\sum_r e_r^s} = \text{average wage in sector } s \quad (3.5)$$

Using 3.2 and 3.3 we form the total employment share in sector  $s$ .

$$\bar{\sigma}^s = \frac{\sum_r e_r^s}{\sum_i E_r} = \text{share of total employment in sector } s \quad (3.6)$$

Combining 3.5 and 3.6 gives us the decomposed wage identity.

$$\bar{w}_r = \sum_s w_r^s \sigma_r^s = \sum_s w_r^s \bar{\sigma}^s + \sum_s \bar{w}^s \sigma_r^s + \sum_s (w_r^s - \bar{w}^s) (\sigma_r^s - \bar{\sigma}^s) - \sum_k \bar{w}^k \bar{\sigma}^k \quad (3.7)$$

where  $\bar{w}_r$  is the average wage in region  $r$ .  $\sum_s w_r^s \bar{\sigma}^s$  denotes the productivity index, where the industrial composition is assumed fixed at the average national level across regions. Thus, the



variation in this index reflects the regional productivity differences.  $\sum_s \bar{w}^s \sigma_r^s$  is the industrial composition index in which regional differences is attributed to differences in the composition of sectors in different regions. In this index the wages across sectors are assumed to be fixed at the average national level. The variation in this index represents the variation of the regional industry composition. The remaining terms in equation 3.7 represents the covariances between the industry employment shares and wages in region  $r$ .

Table 3.1 below presents the variables that we are using in the empirical analysis.

**Table 3.1:** Description of variables used

	<i>Variable</i>	<i>Description</i>	<i>Expected sign</i>
Dependent variables	Regional wage level (Wage = $w_r$ )	The total wage sum in region $r$ divided by the number of employees.	
	Productivity index (Prod = $\sum_s w_r^s \bar{\sigma}^s$ )	Weighted sum of the average earnings for each industry group in region $r$ with the weights equal to average country share of the industry group in relation to total employment. The industry groups are divided into a 2 digit level.	
	Industrial Composition index (Comp = $\sum_s \bar{w}^s \sigma_r^s$ )	Weighted sum of the shares of each industry group in region $r$ . The weight is the average country earnings in terms of the industry groups.	
Independent variables	Market potential (AccGRP) <sup>†</sup>	Accessibility to GRP (gross regional product) in region $r$ discounted from other regions by travel time distances using a distance decay function.	+
	Education (Edu)	Population in working age (20-64) with more than 3 years of university education divided by working age population without university education in region $r$	+
	Firm size (Fsize)	Number of employees divided by number of business establishments in region $r$ .	+
	Employment Share (Empshare)	Number of employees as a share of people in working age (20-64) in region $r$	+
	Self-employment (Semp)	Number of self-employed persons as a share of the population in working age (20-64) in region $r$	+/-

In figures A1.1 to A1.3 found in appendix 1 the correlation between the three variables wage, productivity index, and industrial composition index and our measure of market potential is displayed. The three figures all reveal a weak but positive relation to market potential.

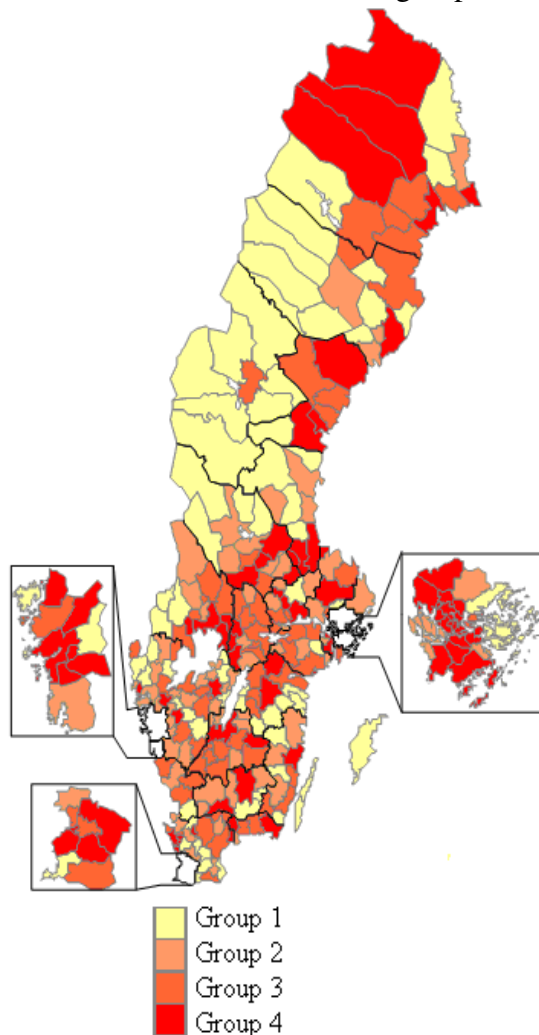
The correlation matrix presented in Table A1.2 naturally show strong correlations between the three dependent variables, since the indices are derived from average wage. The

<sup>†</sup> Agglomeration is measured as accessibility to GRP. The measure is obtained through the following formula;  $AccGRP_r = \sum_k GRP_k \exp(-\lambda d_{rk})$  were Accessibility to GRP for municipality  $r$  is equal to the sum of the GRP in all other municipalities  $k$  discounted by the distance between municipality  $r$  and all other municipalities  $k$ ,  $d_{rk}$ .  $\lambda$  is a distance decay parameter estimated using commuting patterns (see Johansson, Klaesson & Olsson, 2001).

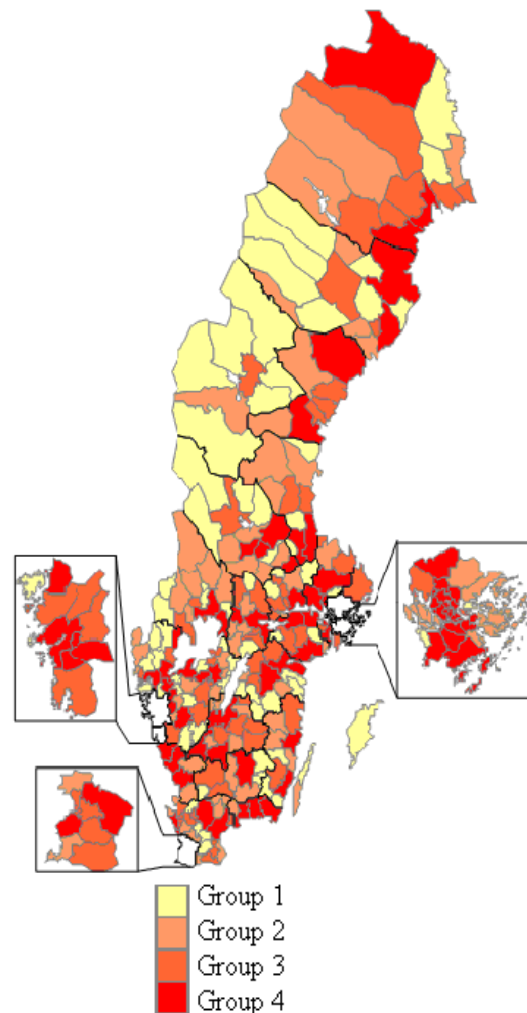
correlation coefficients between the dependent variables and the market potential variable, accessibility to GRP, vary between 0.559 and 0.331. Two of the control variables, average firm size and self employment have a strong correlation vis-à-vis the three dependent variables. The self employment variable is the only variable that shows a negative relationship. This seems to indicate a “refugee” effect in terms of entrepreneurial activities. Education also presents high correlation values.

Furthermore, the descriptive statistics presented in table A1.3 in appendix 1 show that there is a large variation in the variables across Swedish regions. This is especially true for the control variables accessibility to GRP and education, but also in terms of the average wage level itself. This supports the idea that Sweden is characterized by substantial income disparities; wages in the lowest earning region is 55 per cent lower than the level earned in the highest wage region.<sup>‡</sup>

On the following two pages four different maps are provided in order to show the geographic distribution of the dependent variables. All the 289 included Swedish regions have been divided into 4 groups depending on their performance. High performing regions are marked in darker colors, group 3 and 4 respectively. The real values and level of the intervals for the four different groups can be seen in figure A2.1 in appendix 2.



**Figure 3.4:** Average wage

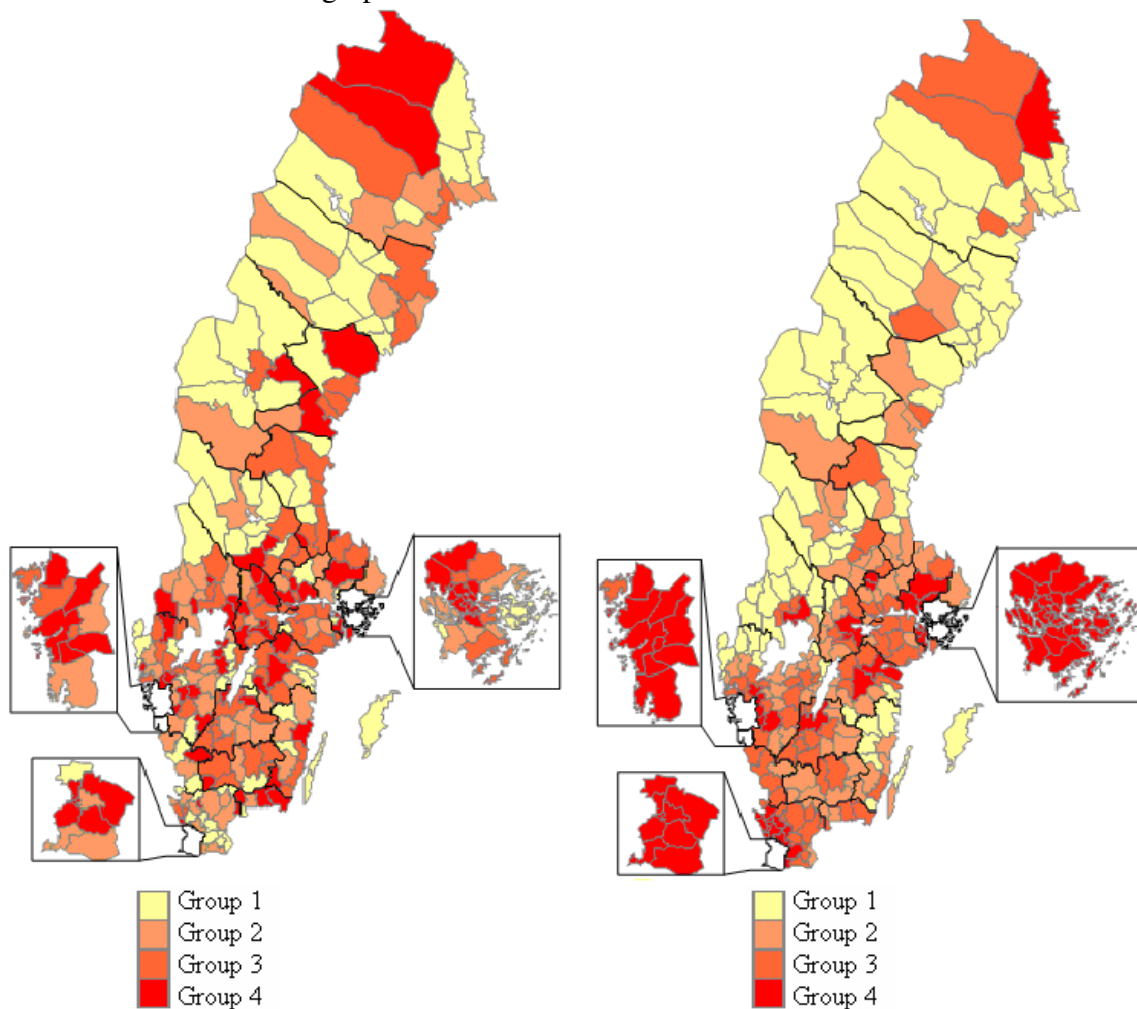


**Figure 3.5:** Productivity index

<sup>‡</sup> Calculations based on data from Statistics Sweden, 2004

Figure 3.4 displays the national variation in terms of average wages. One can see that it is especially areas in the three large city regions Stockholm, Gothenburg and Malmo which are marked as top-ranking. The map suggests that agglomeration in general has a substantial correlation with earnings since all dark areas are regions where Sweden's larger cities can be found; such as Linköping/Norrköping, Jönköping and Karlstad etc. In terms of the lowest earning regions one can say that they often are rural, low populated regions with tendency to be located in the northern parts of Sweden. The fact that some of the northern regions are placed within group 3 and 4 can be explained by the high proportion of employed within the high-technology mining industry where wages traditionally has been very high.

When it comes to productivity, figure 3.5 reveals that it is once again predominately the regions within the large city areas that on average have a higher productivity index compared to the rest of Sweden. Once more the smaller regions far from a large market show the lowest levels of productivity. One can also observe that some of the "mining-municipalities" has a lower productivity level compared to wage, which can support the idea that there is a wage-premium in these areas to enable them to attract labor.



**Figure 3.6:** *Industry Composition Index*

**Figure 3.7:** *Accessibility to GRP*

Figure 3.6 illustrates the geographical variation in terms of industry composition were more or less the same pattern is found as in terms of the wage and productivity distribution.

Finally, in figure 3.7, the accessibility to GRP clearly show that southern Sweden has a higher market potential than the north, this is most pronounced for regions

situated in proximity to either Stockholm, Gothenburg or Malmo. All of the regions within these city areas belong to the highest rated group. The further away you get from these areas you will experience a fall in accessibility and market potential. Spikes in accessibility is then found in those regions that host medium-sized cities; e.g. Linköping/Norrköping, Jönköping and Karlstad.

## 4. Empirical Results

Next let us turn to the empirical analysis of regional wages, productivity and industry composition. In the empirical models the estimated parameters are expressed as elasticities since all variables are logged.

The regression models to be estimated are:

$$\ln wage = \beta_0 + \beta_1 \ln Fsize + \beta_2 \ln Semp + \beta_3 \ln Edu + \beta_4 \ln Empshare + \beta_5 \ln AccGRP + \varepsilon \quad (4.1)$$

$$\ln prod = \beta_0 + \beta_1 \ln Fsize + \beta_2 \ln Semp + \beta_3 \ln Edu + \beta_4 \ln Empshare + \beta_5 \ln AccGRP + \varepsilon \quad (4.2)$$

$$\ln comp = \beta_0 + \beta_1 \ln Fsize + \beta_2 \ln Semp + \beta_3 \ln Edu + \beta_4 \ln Empshare + \beta_5 \ln AccGRP + \varepsilon \quad (4.3)$$

In order to check the robustness of the findings, the regression models are tested using three different methods. First a standard OLS approach is applied followed by spatial lag (SL) and spatial error (SE) models, adjusting for possible spatial autocorrelation. The SL and SE results can be found in appendix 3. Additionally, instrumental variable (IV) estimation is applied to test for possible endogeneity problems.

### Ordinary Least Squares

Table 4.1 below presents the OLS results for all the three regression models.

**Table 4.1: Estimation Results using OLS**

	Wage	Prod.	Comp.
Market potential	0.017 (4.21)***	0.018 (3.69)***	0.009 (4.09)***
Education	0.046 (5.06)***	0.105 (8.42)***	-0.004 (0.62)
Average firm size	0.057 (3.96)***	0.059 (3.01)***	0.022 (2.12)**
Employment	0.094 (4.84)***	0.112 (4.89)***	0.033 (2.83)***
Self-employment	-0.086 (6.10)***	-0.046 (2.48)**	-0.041 (3.92)***
Constant	11.962 (199.54)***	12.099 (154.99)***	12.177 (295.75)***
No. obs.	289	289	289
R <sup>2</sup>	0.74	0.71	0.44

Robust t statistics in parentheses. \*\*\* significant at 1% and \*\* significant at 5% level.

In table 4.1, the OLS regression estimates are presented. The results clearly show that the three dependent variables, average wage, productivity and industry composition indices are significantly linked to the regional attributes.

The table displays that market potential, which is measured as accessibility to GRP, affects both productivity and industrial composition, and hence, the regional variation in average wages. The estimated elasticities reveals that a twice as large market potential is expected to have approximately 1.7 percent higher wages. The productivity index is expected to be 1.8 percent higher, and the industrial composition index is expected to be 0.9 percent higher.

When comparing these results to earlier studies made in other countries, the elasticities seem to be lower in the case of Sweden. Factors that can explain this is (1) the use of accessibility to GRP as a measure of market potential, and (2) the functioning of the Swedish labor market, where wage differences traditionally has been comparatively low.

Table 4.1 also reveals how the control variables influence the regional variation in the dependent variables. As can be observed, all of the variables, with the exception of industry composition in relation to education, turn out to have a significant effect on wages, productivity and industry composition, respectively.

The results show that educational level in municipalities influence wages. It is, however, the productivity index which is most influenced by a high share of educated people in the region with an elasticity of 10.5 percent, while there is no influence on the industrial composition index.

The second explanatory variable is average firm size in each municipality, which also turns out to have a significantly positive effect on all three dependent variables. Yet as in the case of education, the effect is found to be smallest for the industrial composition index. The results points to that there are significant effects from scale economies which are internal to the firms.

When it comes to the share of employed persons within the municipality the results are similar to that of the firm size variable. Higher competition among the employers to attract labor seems to increase wages.

The final control variable is the share of self-employed persons in a municipality. For all three dependent variables the effect is negative and significant. In particular, the negative influence is found to be the largest in terms of average wages. This supports the so called “refugee effect” discussed in earlier sections. Another explanation for the negative relation can also be the structure of the data set, where one cannot separate between incomes from self-employment or regular employment. For example an individual running his/her own business can withdraw or withhold money in irregular intervals, which implies that one year the dividend might be very high while being very low in a second period. If many self-employed chose to use the firm dividend to take out a small proportion as income during our reference year, this could then affect the outcome.

The explanatory powers of the regressions are very high for average wage and productivity where the  $R^2$  is 71 percent or higher. The weaker relationship for industry composition is confirmed by a lower  $R^2$ .

#### *Testing for Endogeneity: IV results*

A problem connected to an analysis of the role played by agglomeration for productivity and income, relates to the fact that it is very hard to differentiate between two possible explanations for a positive correlation between productivity and agglomeration. Productivity and income may be high because of agglomeration effects; which is the underlying idea in this paper. However, there is also a possibility that agglomeration arises

due to the fact that wages and productivity are high due to for example a positive regional specific shock which in turn attracts labor and firms. A recent study by Fu and Ross (2007) tests the above stated endogeneity problem; if wage premiums in clusters are caused by agglomeration economies or of regional characteristics, such as labor heterogeneity. Their result show that the causality runs from agglomeration to wage, and not the other way around.

However, if any of our local characteristics is endogenous in relation to the dependent variables, our models might omit unobserved abilities that will influence the results. If such omitted variables are also correlated with any right hand side variables, then a bias can result. Reverse-causality is hence potentially still a problem which must be tested for.

Furthermore, due to the model set-up another possible bias might arise. The reason is that we do not fully control for the type of labor or firms that are located in the agglomerations, hence there may be a selection bias present. Since some people and firms are more productive than others, this can result in varying levels of productivity and wages. This could explain the wage premium found in larger regions if we assume that more high-performing firms and individuals are located there. A study by Combes et al. (2009) however, confirms that productivity differences, and thus wage differences, predominantly can be explained by the presence of agglomeration economies.

The choice of the instruments is based on the assumption that the instrumental variables represents an exogenous regional characteristic that has lasting influence on localization decisions, and thus on agglomeration, but not on the present level of productivity and income. The instruments chosen are therefore: (1) municipal 1950 population, (2) population 1950 within 1 hour driving distance, (3) municipal land area, (4) land area within 1 hour driving distance, (5) dummies for municipalities belonging to Stockholm, Gothenburg and Malmo regions. The validity of these instruments refers to that they all in some way reflect the potential market size of a region. See table 4.2 below for the IV results.

**Table 4.2: IV Estimation Results for the Wage equation**

	Wage	Prod.	Comp.
Market potential	0.021 (4.65)***	0.015 (3.14)***	0.014 (5.65)***
Education	0.042 (4.88)***	0.111 (8.94)***	-0.010 (1.77)*
Average firm size	0.049 (3.43)***	0.064 (3.12)***	0.011 (1.13)
Employment	0.096 (5.15)***	0.100 (4.45)***	0.041 (3.80)***
Self-employment	-0.098 (6.98)***	-0.049 (2.50)**	-0.052 (5.07)***
Constant	11.884 (183.94)***	12.112 (143.08)***	12.091 (268.73)***
No. obs.	289	289	289
R <sup>2</sup>	0.75	0.71	0.47

Dependent variable: wage

Robust t statistics in parentheses. \* significant at 10%, \*\* significant at 5% and \*\*\* significant at 1%,

The IV estimates corroborates the findings from the OLS regressions presented in table 4.1. However, when comparing the results to the OLS estimates there are some deviations in the elasticities for the control variables. In a majority of the cases the deviations among the elasticities are within the +/- 0.2 to 0.8 percentage bound. Especially it is for the self-employment, employment share and average firm size were the largest variations in

comparison to the OLS results occur. The negative effect of self employment on average wages is 1.2 percent higher than when estimating without instruments.

As stated earlier we have also used different spatial estimation techniques for both the OLS and IV regressions. Even though tests have indicated that there may be some problems with spatial regimes the results from these regressions do not differ in any significant way from the OLS estimations. (see appendix 3, table A3.1 to A3.6). This confirms the robustness of the results.

In summary we confirm earlier research by concluding that denser, larger firm dominating, more educated, and high employment areas are characterized by a higher wage on average.

## 5. Conclusion

The focus of this study has been to estimate regional agglomeration effects in Sweden. We have investigated how concentration and agglomeration influence regional wage levels. In addition, the regional average wage level has been decomposed into a regional productivity index and a regional industrial composition index.

In the productivity index the industrial composition are held constant and the average wage in a region is only influenced by the regional wage level per sector. The industrial composition index on the other hand, is calculated holding industrial sector wages constant across regions. The index is thus only influenced by industry composition. Using these indices we control for the fact that regions have different industry compositions and that the average regional productivity differs between regions.

In the empirical estimations we use the average regional wage level and the two indices as dependent variables. The major explanatory variable is our measure of agglomeration. As a proxy for agglomeration we use accessibility to gross regional product (GRP). This accessibility is calculated using municipal GRP for all regions and discounting them in space according to the driving time distance. Therefore, the measure takes into account both size effects and the spatial layout of the municipalities.

Also, in the regression analysis we control for other factors that may influence wage, productivity and industrial composition effects. These control variables are the education level in each municipality, the share of the working age population that holds a job, average firm size and the degree of self-employment.

The general result of the analysis is that economies of agglomeration are a prevalent feature across regions in Sweden. The results indicate that regional size (as measured by accessibility to GRP) influences productivity as well as industrial composition, and hence, the regional variation in average wages. The estimated elasticities show that a twice as large region is expected to have approximately 1.7 percent higher wages. The productivity index is expected to be 1.8 percent higher, and the industrial composition index is expected to be 0.9 percent higher.

With reference to similar studies in other countries our estimated elasticities appear small. There are at least two factors that can potentially explain this result. The first is dependent on how agglomeration itself is measured. We use an accessibility measure while in the literature it is more common to use variables measuring regional size or density. The other factor has to do with the distribution of wages in Sweden. It is a well known fact that wage disparities are relatively small in Sweden. Wages are set in a collective bargaining between different unions and employer organizations.

Turning to the control variables, the education level in municipalities influence wages and, in particular, the productivity index (about twice as large effect), while there is no influence on the industrial composition index.

The share of employed people has a significant effect on all three dependent variables, but the size of effect on the industrial composition is about one third of the size compared to the other two.

The average size of the firms in each municipality has a significantly positive effect on all three dependent variables, but once again, the effect is the smallest for the industrial composition index. This points to that there are significant effects from scale economies that are internal to the firms.

The last explanatory variable is the share of self-employed people. For all three dependent variables the effect is negative and significant. In particular, the negative influence is the largest on the average wages.

In the estimation we have used some different spatial estimation techniques. Even though tests have indicated that there may be some problems with spatial regimes the results from these regressions do not differ in any significant way from the OLS estimations.

In addition, in the empirical analysis we have also used an instrumental variable approach to investigate possible bias from reverse causality from the dependent variables to the independent ones. Results indicate that this bias is probably small.

In this study, we have estimated relationships for the aggregate economy. Of course, it is quite possible that agglomeration effects differ between industries. Therefore, one way forward is to perform this kind of analysis on industrial aggregates. Especially, it should be interesting to compare different kinds of service sectors since it can be argued that they are more dependent on proximity to larger markets. Also, sectors with a high knowledge or R&D intensity may have more to gain from agglomerations, where knowledge spillovers can be expected to be more important.

One other possible way to continue and broaden the analysis is make use of the fact that our agglomeration proxy has the form of an accessibility variable. Since this variable is calculated using the time distance of the road network between municipalities it is possible to assess effects on wages, productivity and industrial composition from changes in the quality of the road infrastructure.

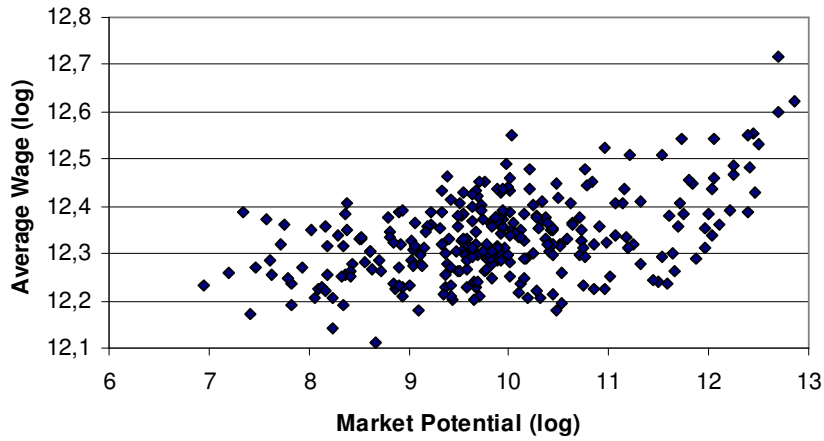


## References

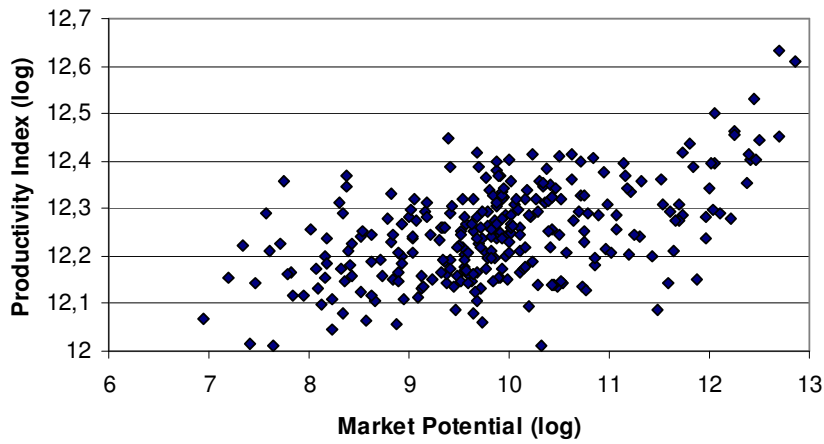
- Albæk, K., R. Asplund, et al. (2000). "Dimensions of the Wage-Unemployment Relationship in the Nordic Countries: Wage Flexibility without wage curves", Nordic Council of Ministers, Labour Market Committee.
- Asplund, R. and P. Telhado Pereira (1999). "Returns to Human Capital in Europe - A Literature Review". Helsinki, Taloustieto Oy.
- Audretsch, D. B. and M. Fritsch (1994). "The Geography of Firm Births in Germany." *Regional Studies* 28(4): 359-365.
- Becker, G. S. (1962). "Investment in Human Capital: A Theoretical Analysis." *The Journal of Political Economy* 70(5): 9-49.
- Blanchflower, D. G. and A. J. Oswald (1990). "The Wage Curve." *The Scandinavian Journal of Economics* 92(2): 215-235.
- Blanchflower, D. G. and A. J. Oswald (1994). "The wage curve", Cambridge: MIT Press.
- Brown, C. and J. Medoff (1989). "The Employer Size-Wage Effect." *Journal of Political Economy* 97: 1027-1059.
- Brunello, G. and S. Comi (2004). "Education and earnings growth: evidence from 11 European countries." *Economics of Education Review* 23: 75-83.
- Calmfors, L. (1993a). "Centralisation of Wage Bargaining and Macroeconomic performance", OECD Economic Studies No. 21.
- Calmfors, L. (1993b). "Lessons from the macroeconomic experience of Sweden". *European Journal of Political Economy*. 9, 25-72
- Ciccone, A. and Hall, R.E. (1996), "Productivity and the Density of Economic Activity", *American Economic Review*, 86(1), 54-70.
- Ciccone, A. (2002). "Agglomeration effects in Europe." *European Economic Review* 46: 213-227.
- Combes, P-P., Duranton, G. and L. Gobillon. (2008). "Spatial wage disparities: sorting matters!". *Journal of Urban Economics* 63: 723-742.
- Combes, P-P., Duranton, G., Puga, D. and S. Roux. (2009). "The Productivity advantages of large cities: Distinguishing agglomeration from firm selection". WP.
- Fingleton, B. (2003), "Increasing returns: evidence from local wage rates in Great Britain", *Oxford Economic Papers*, 55, 716-739.
- Fu, S. and S.L. Ross. (2007). "Wage Premia in Employment Clusters: Agglomeration Economies or Worker Heterogeneity?" University of Connecticut. Department of Economics Working Paper Series. 26R
- Glaeser, E.L., Kallal, H., Sheinkman, J. and Schleifer, A. (1992). "Growth in cities". *Journal of Political Economy* 100:1126-52.
- Glaeser, E.L. and Mare D.C. (2001). "Cities and Skills". *Journal of Labor Economics*. Vol 19. No 2. pp. 316-342
- Golden and Lange, (1998) "Globalization and Industrial Relations." Presented at the 1998 Congress of the International Industrial Relations Association, Bologna, September 22-26
- Graves, P.E. Arthur, M.M. and R.L. Sexton. (1999). Amenities and the Labor Earnings Function. *Journal of Labor Research*. Volume XX(3); 367-376.
- Hoover, E.M. (1937). "Spatial price discrimination". *Review of Economic Studies* 4: 182-91.
- Kenworthy, L. (2001). "Wage-Setting Measures: A Survey and Assessment", World Politics.
- Krugman, P. (1980). "Scale Economies, Product Differentiation, and the Pattern of Trade." *The American Economic Review* 70(5): 950-959.
- Lindgren, K-O. (2006). "Roads from Unemployment, Institutional Complementarities in Product and Labour Market." Statsvetenskapliga institutionen, Uppsala

- Lucas, R.E. (1988). "On the mechanics of economic development. " *Journal of Monetary Economics* 22: 3-22.
- Marshall, A. (1920). *Principles of Economics*. London: Macmillan. 8<sup>th</sup> edition.
- McDonald, J. F. (1997), "Fundamentals of Urban Economics", Prentice-Hall, New Jersey.
- Mellow, W. (1982). "Employer Size and Wages". *Review of Economics and Statistics* 64: 495-501.
- Mincer, J. (1974). "Schooling, Experience and Earnings. " New York, NBER.
- Moore, H. L. (1911). "Laws of Wages: An essay in statistical economics. " New York, Augustus M. Kelly.
- Nijkamp, P. and J. Poot (2005). "The last word on the wage curve?" *Journal of Economic Surveys* 19(3): 421-450.
- Ohlin, B. (1933). "Interregional and International Trade". Cambridge, MA: Harvard University Press.
- Oxenfeldt, A. (1943). "New Firms and Free Enterprise." Washington, D.C, American Council on Public Affairs.
- Porter, M. E. (1990). "The competitive advantage of nations". New York: The Free Press
- Rauch, J. E. (1993). "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities." *Journal of Urban Economics* 34: 380-400.
- Roback, J. (1982). "Wages, rents and the quality of life". *Journal of Political Economy*, Vol. 90 (6):1257-1278,
- Romer, P. (1986). "Increasing returns and long-run growth". *Journal of Political Economy* 94: 1002-37.
- Romer, P. (1990). "Endogenous technological change." *Journal of Political Economy* 98:S71-S101
- Rosenthal, S.S. and W.C. Strange. (2004). "Evidence on the Nature and Sources of Agglomeration Economies." In Henderson, J.V. and J.F, Thisse. *Handbook in Economics 7. Handbook of Regional and Urban Economics. Volume 4. Cities and Geography*. pp. 2119-2167. Elsevier. B.V. Amsterdam.
- Sianesi, B. and J. Van Reenen (2003). "The Returns to Education: Macroeconomics." *Journal of Economic Surveys* 17(2): 157-200.
- Thurik, A. R., M. A. Carree, et al. (2007). "Does self-employment reduce unemployment?" accepted for publication in *Journal of Business Venturing* (2008).
- Troske, K. R. (1999). "Evidence on the Employer Size-Wage Premium for Worker-Establishment Matched Data." *The Review of Economics and Statistics* 81(1): 15-26.
- Weber, A. (1909). "Über den Standort der Industrien. Tübingen": J:C.B Mohr. English translation: *The Theory of the Location Industries*. Chicago: Chicago University Press, 1929.

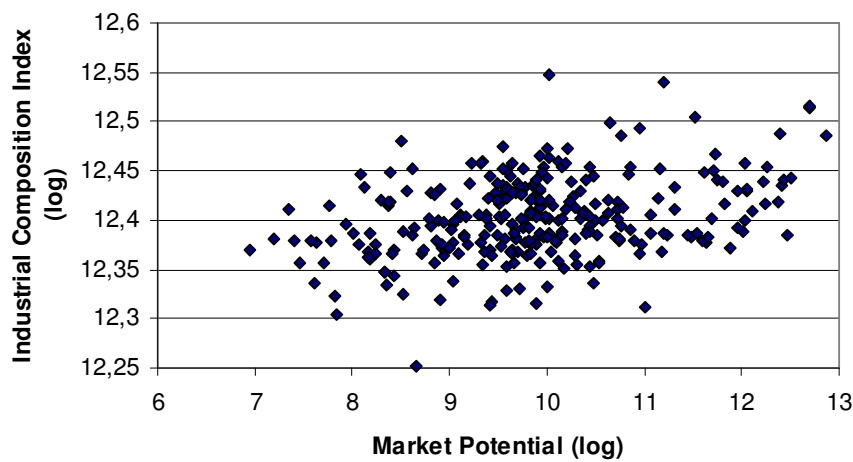
## Appendix 1. Descriptive Statistics



**Figure A1.1:** Relationship between Market Potential and Regional Average Wage



**Figure A1.2:** Relationship between Market Potential and Regional Productivity Index



**Figure A1.3:** Relationship between Market Potential and Regional Industrial Composition Index

In table A1.1 the relationships between the three variables are presented.

**Table A1.1: Relationship between Productivity Index, Industrial Composition Index and Average Wages**

Prod.index	0.680 (21.88)***		0.496 (21.36)***
Comp.index		1.582 (18.98)***	1.059 (18.48)***
Constant	4.007 (10.53)***	-7.293 (7.05)***	-6.886 (10.70)***
Observations	289	289	289
R <sup>2</sup>	0.63	0.56	0.83

**Table A1.2: Correlation matrix**

	Wage	Prod	Comp	AccGRP	Edu.	Empshare	Fsize	Selfem
Wage	1							
Prod.index	0.798	1						
Comp.index	0.747	0.433	1					
Acc.GRP	0.559	0.568	0.331	1				
Education	0.591	0.667	0.333	0.677	1			
Employ.share	0.430	0.374	0.336	-0.021	0.116	1		
Firm.size	0.750	0.635	0.613	0.246	0.318	0.542	1	
Selfemploy.	-0.678	-0.572	-0.567	-0.264	-0.358	-0.187	-0.680	1

**Table A1.3: Descriptive Statistics**

	Minimum	Maximum	Mean	Median	Std. Deviation	Skewness
Wage	182074.56	333467.57	228062.34	227168.5	20322.98	1.12
Prod.index	164539.71	306625.15	209822.38	211247.0	21535.01	0.81
Comp.index	209460.61	281583.11	243793.33	244140.6	9945.66	0.41
Acc.GRP	1036.26	386378.32	39693.93	21279.5	59561.42	2.98
Education	0.09	0.85	0.19	0.17	.10	2.72
Employ.share	0.30	1.61	0.68	0.68	.14	0.70
Firm.size	1.52	9.29	3.43	3.38	1.32	1.061
Selfemploy.	0.02	0.14	0.05	0.50	.01	0.94

## Appendix 2. Regional descriptive statistics

Group	Average wage	Group	Productivity index
1	182 074,56 - 214 885,45	1	164 539,71 - 193 174,57
2	214 885,46 - 225 487,99	2	193 174,58 - 208 072,46
3	225 488,00 - 238 047,11	3	208 072,47 - 222 612,22
4	238 047,12 - 333 467,57	4	222 612,23 - 306 625,15

Group	Composition index	Group	Accessibility to GRP
1	209 460,61 - 237 677,21	1	1 036,26 - 10 049,58
2	237 677,22 - 243 150,72	2	10 049,59 - 18 926,04
3	243 150,73 - 250 021,68	3	18 926,05 - 35 283,51
4	250 021,69 - 281 583,11	4	35 283,52 - 386 378,32

\* Values measured in number of SEK

**Figure A2.1** *Values corresponding to map groups*

## Appendix 3. Testing for Spatial Errors

SE = Spatial error model

SL= Spatial lag model

**Table A3.1: Estimation Results for the Wage equation**

	OLS	SL	SE
Ln Fsize	0.057 (3.96)***	0.047 (3.31)***	0.057 (3.60)***
Ln Semp	-0.086 (6.10)***	-0.087 (6.24)***	-0.079 (5.11)***
Ln Edu	0.046 (5.06)***	0.044 (5.57)***	0.040 (4.44)***
Ln Empshare	0.094 (4.84)***	0.118 (6.50)***	0.104 (5.28)***
Ln AccGRP	0.017 (4.21)***	0.011 (2.99)***	0.016 (3.28)***
Constant	11.962 (199.54)***	6.244 (5.38)***	11.987 (183.98)***
Obs.	289	289	289
R <sup>2</sup>	0.74	0.77*	0.74*

Dependent variable: wage

Robust t statistics in parentheses. \*\*\* significant at 1%

\* Pseudo R-squared = ratio of the variance of the predicted values over the variance of the observed values for the dependent variable

**Table A3.2: Estimation Results for the Productivity Index Equation**

	OLS	SL	SE
Ln Fsize	0.059 (3.01)***	0.056 (2.92)***	0.058 (2.99)***
Ln Semp.	-0.046 (2.48)**	-0.048 (2.60)***	-0.047 (2.53)**
Ln Edu	0.105 (8.42)***	0.105 (8.55)***	0.106 (8.64)***
Ln Empshare	0.112 (4.89)***	0.118 (5.24)***	0.113 (4.98)***
Ln AccGRP	0.018 (3.69)***	0.016 (2.80)***	0.018 (3.80)***
Constant	12.099 (154.99)***	10.791 (7.58)***	12.097 (155.81)***
Obs.	289	289	289
R <sup>2</sup>	0.71	0.72*	0.71*

Dependent variable: productivity index

Robust t statistics in parentheses. \*\* significant at 5%; \*\*\* significant at 1%

\* Pseudo R-squared = ratio of the variance of the predicted values over the variance of the observed values for the dependent variable

**Table A3.3: Estimation Results for the Industrial Composition Index Equation**

	OLS	SL	SE
Ln Fsize	0.022 (2.12)**	0.015 (1.52)	0.016 (1.35)
Ln Semp	-0.041 (3.92)***	-0.044 (4.34)***	-0.045 (3.90)***
Ln Edu	-0.004 (0.62)	-0.002 (0.31)	-0.004 (0.67)
Ln Empshare	0.033 (2.83)***	0.043 (3.86)***	0.040 (3.10)***
Ln AccGRP	0.009 (4.09)***	0.006 (2.81)***	0.007 (2.17)**
Constant	12.177 (295.75)***	6.077 (3.79)***	12.194 (272.40)***
Obs.	289	289	289
R <sup>2</sup>	0.44	0.47*	0.43*

Dependent variable: industry composition index

Robust t statistics in parentheses. \*\* significant at 5%; \*\*\* significant at 1%

\* Pseudo R-squared = ratio of the variance of the predicted values over the variance of the observed values for the dependent variable

**Table A3.4: IV Estimation Results for the Wage equation**

	OLS	SL	SE
Ln Acc.GRP	0.021 (4.65)***	0.014 (3.03)***	0.023 (4.23)***
Ln Education	0.042 (4.88)***	0.043 (5.43)***	0.037 (4.35)***
Ln Employ.share	0.096 (5.15)***	0.116 (6.51)***	0.103 (5.65)***
Ln Firm.size	0.049 (3.43)***	0.044 (3.10)***	0.051 (3.38)***
Ln Selfemploy.	-0.098 (6.98)***	-0.094 (6.67)***	-0.088 (5.80)***
Constant	11.884 (183.94)***	6.662 (5.42)***	11.893 (174.67)***
Obs.	289	289	289
R <sup>2</sup>	0.75	0.77*	0.75*

Dependent variable: wage

Robust t statistics in parentheses. \*\*\* significant at 1%

\* Pseudo R-squared = ratio of the variance of the predicted values over the variance of the observed values for the dependent variable

Instruments: (1) municipal 1950 population, (2) population 1950 within 1 hours driving distance, (3) municipal land area, (4) land area within 1 hours driving distance, (5) dummies for municipalities belonging to Stockholm, Gothenburg and Malmo regions

**Table A3.5: IV Estimation Results for the Productivity Index Equation**

	OLS	SL	SE
Ln Acc.GRP	0.015 (3.14)***	0.012 (2.07)**	0.015 (3.03)***
Ln Education	0.111 (8.94)***	0.111 (9.07)***	0.111 (9.01)***
Ln Employ.share	0.100 (4.45)***	0.107 (4.75)***	0.100 (4.46)***
Ln Firm.size	0.064 (3.12)***	0.062 (3.07)***	0.064 (3.10)***
Ln Selfemploy.	-0.049 (2.50)**	-0.049 (2.54)**	-0.049 (2.47)**
Constant	12.112 (143.08)***	10.388 (7.49)***	12.112 (140.95)***
Obs.	289	289	289
R <sup>2</sup>	0.71	0.71*	0.71*

Dependent variable: productivity index

Robust t statistics in parentheses. \*\* significant at 5%; \*\*\* significant at 1%

\* Pseudo R-squared = ratio of the variance of the predicted values over the variance of the observed values for the dependent variable

Instruments: (1) municipal 1950 population, (2) population 1950 within 1 hours driving distance, (3) municipal land area, (4) land area within 1 hours driving distance, (5) dummies for municipalities belonging to Stockholm, Gothenburg and Malmo regions

**Table A3.6: IV Estimation Results for the Industrial Composition Index Equation**

	OLS	SL	SE
Ln Acc.GRP	0.014 (5.65)***	0.011 (3.98)***	0.014 (4.69)***
Ln Education	-0.010 (1.77)*	-0.008 (1.35)	-0.010 (1.72)*
Ln Employ.share	0.041 (3.80)***	0.048 (4.39)***	0.044 (3.92)***
Ln Firm.size	0.011 (1.13)	0.008 (0.81)	0.010 (0.92)
Ln Selfemploy.	-0.052 (5.07)***	-0.052 (5.21)***	-0.051 (4.96)***
Constant	12.091 (268.73)***	7.522 (4.10)***	12.097 (249.08)***
Obs.	289	289	289
R <sup>2</sup>	0.47	0.48*	0.47*

Dependent variable: Industry composition index

Robust t statistics in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

\* Pseudo R-squared = ratio of the variance of the predicted values over the variance of the observed values for the dependent variable

Instruments: (1) municipal 1950 population, (2) population 1950 within 1 hours driving distance, (3) municipal land area, (4) land area within 1 hours driving distance, (5) dummies for municipalities belonging to Stockholm, Gothenburg and Malmo regions



## Appendix 4. Regression Results omitting the control variables

**Table A4.1** *Regression results for regional market size*

	Wage	Prod.index	Comp.index
Acc.GRP	0.036 (7.72)***	0.046 (9.13)***	0.012 (6.23)***
Constant	11.977 (265.57)***	11.796 (240.47)***	12.283 (636.85)***
Obs.	289	289	289
R <sup>2</sup>	0.24	0.28	0.12

Robust t statistics in parentheses. \*\*\* significant at 1%