How the Transition Affected Regional Disparities – The Case of Hungary

by David-Krisztian Nagy

Submitted to

Central European University

Department of Economics

In partial fulfilment of the requirements for the degree of Master of Arts in Economics

Supervisor: Professor Miklos Koren

Budapest, Hungary

June 2010

Abstract

Based on a linked employer-employee dataset on Hungarian workers, this paper examines the evolution of spatial distribution of economic activities in the country during the transition from socialism to capitalism. Besides confirming previous findings such as increasing concentration of the overall economy in space and a shift in economic activities from the Eastern to the Western part (as well as to the capital) of the country, I find (1) a large amount of heterogeneity in changes in the distribution of different industries and (2) differentials across regions and industries which were more open or less open to trade before the transition, with the latter mentioned effect being also different between Eastern and Western Hungary as well as across urban and rural regions. The results of the thesis might provide motivation for further studies in the topic.

Contents

Acknowledgements

I am especially grateful to my supervisor, Miklós Koren, for his help with the preparation of this thesis. In particular, I thank to him for the fascinating insights he always brought into our discussions on the topic as well as for the opportunity to participate in research projects under his assistance and, in this way, learn a lot about applied research. I am also indebted to Gábor Békés whom I worked with on a project examining spatial patterns in the distribution of Hungarian gasoline prices and who made me familiar with a variety of applied techniques, as well as to Péter Harasztosi who provided me with extensive datasets on Hungarian settlements, making my work much easier than it would have been otherwise.

Last but but not least, I would like to thank to my classmate, Ildikó Magyari, for her continuous encouragement and the enormous amount of technical support she provided me at all stages of preparing this thesis.

1 Introduction

It is a well-known and well-documented fact that the economy of Hungary went through dramatic changes in the end of 80s and the beginning of 90s, that is, during the period of transition from socialism to capitalism. These changes had a large effect on the spatial distribution of economic activities and income within the country: some regions, especially those located in the West and around Budapest, managed to improve their performance while others, especially in the Eastern part of the country, fell back, resulting in an increased concentration of the Hungarian economy (Nemes-Nagy, 2004; Major and Nemes-Nagy, 1999). The literature, however, has not provided a complete picture of these changes in geographical structure. Were these changes, for instance, similar or different across sectors such as agriculture, services and different sorts of manufacturing? Did the changes primarily occur because of the shift from a planning system to a decentralized, market-based one, because of changes in technologies or preferences, or because of opening up to trade with Western markets? In my thesis, I contribute to answering two of these questions, the first one being whether different sectors have experienced different changes in their spatial structure, and the second one being whether international trade, or, more precisely, differences in openness to trade across regions and sectors prior to the transition, might have had a significant effect on the process of spatial reordering.

Economic geography has become a prosperous and increasingly popular field since the revolutionary wave of "New Economic Geography" models, originating from Krugman (1991). As a consequence, spatial questions have become interesting on their own for many economists. Analyses that highlight the channels through which and the extent to which different forces such as increasing returns, technology, preferences, market access (in particular, barriers to trade), and others shape the spatial distribution of economic activities in the light of new models are theoretically challenging. Of course, this thesis does not attempt to do a complete analysis of these forces and their effects. However, it can support prospective theoretical work by providing stylized facts on the spatial rearrangement of a country's economy.

The results of this investigation might be especially useful because of the rapid and unexpected way the Hungarian transition, and, consequently, the changes in the economy's spatial structure, occured. This is in sharp contrast with the well-documented cases of Western European countries in which Brülhart and Traeger (2005) find insignificant changes in geographical concentration of most sectors over the 1975 to 2000 period, or the US in which Dumais et al. (2002) describe similarly smooth spatial patterns. My study can supplement the existing literature by reporting stylized facts for a country in which, as opposed to the countries above, large geographical changes took place within a relatively short period.

This paper is novel in two respects. First, as already mentioned, I attempt to answer questions which have not been investigated in the literature on the Hungarian transition, or only qualitatively, as by Barta (2003). Second, as opposed to existing papers on spatial implications of the Hungarian transition like Nemes-Nagy (2004) which measure the spatial concentration of economic activities by using standard spatial indices such as the Hoover Index, I apply relatively recently developed methods (Ellison and Glaeser, 1997; Duranton and Overman, 2005) that are more reliable than standard indices of spatial concentration because they explicitly handle some problems that would otherwise cause biases in measurement.

The main findings of the thesis are as follows. First of all, I confirm the phenomena already described in the literature, i.e., the radical increase in concentration of the overall economy and its shift from the Eastern to the Western part, as well as to the capital city, of the country. Second, I show that huge intersectoral heterogeneity can be found behind these overall tendencies: concentration rose in some industries but decreased in others, with most of the sectors[1] moving to the West but a couple of them moving in the opposite direction. Finally, I find that average differentials between the performance of industries and regions which differed in the degree of trade openness before the transition are large, both within the Western and within the Eastern part of the country, but the effect of inter-regional differentials seems to be stronger in the West while that of inter-industry differentials in the East.

The paper is structured in the following way. In Section 2, I review the related literature in more detail. Next, I describe the main steps of the Hungarian transition process in general. I present the methods used in the analysis – spatial indices, the Duranton–Overman continuous approach, and the methods used for comparing regions' positions – in Section 4. In Section 5, I describe the data used, especially focusing on the Harmonized Hungarian Wage Survey, the main dataset I build on. I also discuss weaknesses of the dataset as well as attempts to mitigate these problems. The results of the analysis are presented in Section 6; within this, I first uncover general tendencies, then look at specific industries and finally at the effect of international trade. Section 7 concludes.

[1] Throughout the paper, I use the words "sector" and "industry" interchangeably.

2 Related Literature

Literature related to this thesis can be divided into three main categories. First, as already mentioned, a relatively large number of papers have investigated the changes in regional inequalities in Hungary around the transition. Based on Hoover Indices, Nemes-Nagy (2004) as well as Nemes-Nagy and Németh (2003) find increases in spatial concentration of taxable incomes, unemployment and the number of phone lines, signs of an East-West shift, increasing disparities between urban and rural regions as well as rising polarization between Budapest and the rest of the country. Nemes-Nagy and Németh (2003) also find that the correlation between counties' per capita GDP in 1994 and 2002 is as high as 0.95, that is, the new spatial structure formed until 1994 and has been relatively stable since then. Along the same way, Major and Nemes-Nagy (1999) show by calculating eight different indices that the increase in concentration is robust to what kind of standard spatial index one uses.

Quadrado et al. (2001) find further evidence on the East-West shift by applying more sophisticated methods such as factor and cluster analysis. Kertesi and Köllő (1997) analyze regional wage differentials based on the same dataset as what I use, i.e., the Hungarian Wage Survey. They come to the conclusion that the change in wage inequalities across regions is primarily driven by the increase in unemployment. Finally, Barta (2003) seems to be the only author looking at the spatial changes of individual sectors. However, her analysis is primarily qualitative. Still, she documents a definite raise in the geographical concentration of manufacturing between 1980 and 2000 (based on GDP data), as well as an increased degree of concentration in the food industry and more dispersion in the case of machinery.

Second, many recent papers which attempt to measure spatial concentration of an economy and / or specific sectors, share their methodology with this thesis. In the literature using Theil Indices of employment, Brülhart and Traeger (2005) seems to be closest to my paper. The authors characterize changes in the spatial distribution of industries in 17 Western European countries by measuring both *topographic concentration* (that is, to what extent workplaces in the specific industry are not evenly spread in physical space) and *concentration relative to total employment* (i.e., to what extent the distribution of workplaces in the specific industry differs from the distribution of total employment). As already cited in the introduction, they find insignificant changes in most sectors and also for the overall economy between 1975 and 2000, although they report a significant dispersion trend in the case of manufacturing.

Another vein of the literature builds upon the Ellison–Glaeser Index, a nonstandard spatial index developed by Ellison and Glaeser (1997) which has the appealing property that it explicitly controls for the effect of industrial concentration (i.e., to what extent employment is concentrated across plants) on spatial concentration. Papers applying the EG Index in the examination of changes in geographical concentration include Dumais et al. (2002) – for the US – and Barrios et al. (2005) – for Ireland and Portugal –. Similarly to Brülhart and Traeger (2005), they report only slight changes (over 20- and 13-year periods, respectively) for the biggest fraction of industries. Finally, Duranton and Overman (2005) develop a third method – one which I also use in this thesis – which goes even further by taking into consideration the fact that space is continuous, a phenomenon that causes biases in both standard spatial indices and the Ellison–Glaeser Index.[2]

Finally, the effects of opening up to trade on the internal spatial structure of a country's economy have also been analyzed, both theoretically and empirically, by many authors. Krugman and Livas Elizondo (1996) is the classical theory paper in this field. By using a simple New Economic Geography model with three regions (two domestic regions within a country and the rest of the world), the authors show that concentration of manufacturing in one of the domestic regions is the only equilibrium if the country is a closed economy while if the country decreases barriers to foreign trade, at some point concentration ceases to be an equilibrium; meanwhile, a symmetric equilibrium emerges in which manufacturing is evenly divided across domestic regions. In short, opening up to trade results in a *decreased* concentration of the domestic economy. This is the same result as the one Behrens et al. (2007) obtain in a New Economic Geography model with consumers' preferences of another type (quadratic instead of CES) and additive instead of iceberg transport costs.

However, not all theory papers support this story. For instance, Montfort and Nicolini (2000) as well as Paluzie (2001) derive the opposite conclusion from setups which are, somewhat surprisingly, very similar to Krugman and Livas Elizondo's: as the degree of openness to trade increases, regional agglomerations might emerge within countries. It would have been nice if empirical papers had shown which scenario prevails in reality; however, they did not manage to come to the same conclusion either. Sanguinetti and Volpe Martincus (2009), for instance, examine the effect of tariff cuts on location patterns of different industries in Argentina by using both descriptive measures and a regression setup and find that industries which experienced higher cuts in tariffs became less concentrated. On the contrary, Crozet and Koenig Soubeyran (2004) build up a New Economic Geography model in which regions' proximity to the rest of the world is different and show that opening up to trade generally results in more spatial concentration with the core region being the one closer to foreign markets, although an opposite pattern can emerge if the competitive pressure of foreign firms is too high. They find evidence on the first case by analyzing data

[2]For more on the methods these papers apply, see Section 4.

on urbanization in Romania from the 90s.

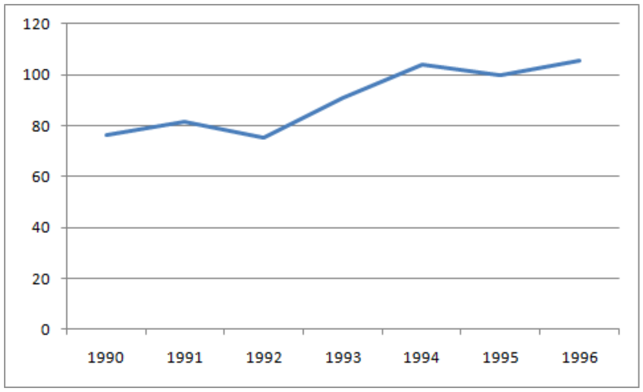


Figure 1: Volume indices of Hungarian exports. (Source: Hungarian Statistical Office)

3 The Transition Period in Hungary

The aim of this section is to give a short summary of the steps of economic transition in Hungary. The primary reason for this is to see the time interval of the process. That is, when did the Hungarian transition start and when did it finish?

Of course, it is impossible to give a precise answer to this question. Some steps toward liberalization of domestic markets were initiated as early as 1968; this year, state enterprises were given more autonomy than usual in Socialist countries, and also prices became more liberal and rational, although they remained fully under state control (Adam, 1999). From the early 80s, enterprises were allowed to choose the set of goods they produced, though the state remained able to influence these decisions through so-called "profile controls" (Hare, 2001a). However, such controls were initiated more and more rarely. Liberalization of prices took place gradually starting from 1988, and by early 1991, the price system was almost fully liberalized except the prices of energy and public services such as public transport. In 1991, the price office was transformed to the country's competition office (Hare, 2001a).

The process of privatization lagged to some extent behind the liberalization of markets. Although small private firms were allowed to operate from 1981, legislation made the entry of privately owned firms possible into all markets in the late 80s (Hare, 2001a) and the 1989 "transformation law" allowed managers to privatize their companies under some conditions (Adam, 1999), 90% of Hungarian GDP was produced by state-owned enterprises and "quasi-public" cooperatives even in 1990 (Hare, 2001a). Still, the process took place gradually and had a bigger wave again in 1994 and 1995 (Adam, 1999).

Finally, let us look at the process of international trade liberalization. Concerning exports, some industrial enterprises received the right of exporting as early as 1968 and these rights were extended in 1979; however, big changes in legislation took place only in the end of 80s (Hare, 2001b). As for import liberalization, it started in 1989 when a three-year program was adopted to decrease barriers to foreign imports. In 1991, Hungary's tariff levels were already around the world average (Hare, 2001b). As a result, the volume of trade boosted up (see the volume of exports in Figure 1; one can see a similar increase in the case of imports). This was, however, driven by trade to and from the West only; intra-CMEA trade (i.e., trade with countries in the former Soviet block) declined, for instance by more than 20 percent only in 1990 (Hare, 2001b).

4 Methodology

4.1 Indices of Spatial Concentration

Indices of spatial concentration, or simply spatial indices, attempt to summarize the degree to which an industry or an economy is concentrated in space in a single number. Many different indices can be found in the literature; for a – not comprehensive – review, see Combes, Mayer and Thisse (2008). I use two of these indices in my thesis. The first one, the *Theil Index*, is given by the formula

$$T(x) = \sum_{r=1}^{R} \frac{x_r}{x} \ln\left(R\frac{x_r}{x}\right)$$

where r is the index of regions within the country[3] that goes from 1 to R (the number of regions), x_r is the amount of variable x that belongs to region r, and $x = \sum_{r=1}^{R} x_r$ is the total amount of x in the country. x can be any variable that is observed and is divided among regions, for instance, employment, capital stock, workers' income, firms' revenues, and so on. The range of the Theil Index is between zero and $\ln(R)$, the former implying a perfectly homogenous pattern (i.e., x_r is exactly the same in all regions), and the latter implying all the activity being concentrated in one of the regions.

However, what one wants to compare the spatial distribution of a variable to is not necessarily the homogenous case. For instance, I examine in Chapter 6.2 whether the distribution of employment in specific industries is different from the distribution of total employment in Hungary. To answer this question, one can use the *Relative Theil Index*,

$$RT(x) = \sum_{r=1}^{R} \frac{x_r}{x} \ln\left(\frac{\frac{x_r}{x}}{\frac{y_r}{y}}\right).$$

Here, y_r is the benchmark variable (e.g., total employment) in region r and $y = \sum_{r=1}^{R} y_r$ is its country aggregate. The Relative Theil Index is equal to zero if the distribution of x is the same as that of y; otherwise, it takes on positive values and the bigger it is, the more concentrated the distribution of x is relative to the distribution of y.

The Theil Index is a useful tool not only because of its simplicity but also because of its *decomposability*. If one can observe more than one level of spatial aggregation, that is,

[3] By the term "country", I refer here to the spatial unit for which one calculates the index of spatial concentration. Of course, this is not necessarily a country but can be something smaller (e.g., a county) or larger (e.g., the European Union). Similarly, regions are not necessarily regions in EU terms (i.e., spatial units of the NUTS-2 level) but can be any within-country units such as settlements, counties, and so on.

smaller regions form larger ones, it is possible to write the Theil Index (and also the Relative Theil Index) as

$$T(x) = T_w(x) + T_b(x)$$

where $T_w(x)$, the *within Theil Index*, is a weighted average of the large regions' Theil Indices, while $T_b(x)$, the *between Theil Index*, is the Theil Index that we would obtain if we treated large regions instead of small regions as the spatial units of observation.[4] If $T_w(x) > T_b(x)$, that is, the within Theil Index explains a larger fraction of $T(x)$ than the between Theil Index, we can conclude that concentration of the x variable takes place primarily within the large regions, whereas the opposite tells us that concentration is mainly due to differences across the large regions while these regions themselves are more or less homogenous.

One of the main weaknesses of the Theil Index is that it is subject to the so-called *Modifiable Areal Unit Problem* (MAUP). This problem, often cited in the literature, stems from the fact that we do not observe the actual distribution of our x variable in continuous physical space but only its values associated with smaller or larger regions – in mathematical terms, a *discretization* of the actual distribution (Combes, Mayer and Thisse, 2008). As Brülhart and Traeger (2005) point out, the Theil Index suffers from three sorts of biases due to the MAUP. First of all, regions are different in size which implies that the Theil Index takes its lowest value not if our variable of interest is evenly spread in physical space but if it is the same across regions. In other words, the Theil Index falls short of measuring actual *topographic concentration* (Brülhart and Traeger, 2005). Fortunately, this type of bias can be eliminated by resorting to the Relative Theil Index, either by comparing the distribution of x to the distribution of another economic variable y (in this case, regions' size is implicitly taken into account through y), or by using the areas of regions (e.g., in square kilometres) as the benchmark. In the latter case, the Theil Index takes on its lowest value if the distribution of x is proportional to regions' areas, that is, x is not the same across regions but it is evenly distributed in physical space.

The other two types of biases are present even if one uses the Relative Theil Index. First, the Theil Index implicitly assumes that variable x is homogenously distributed *within* regions while this might not be the case. In other words, a part of the concentration of x is not accounted for. This kind of bias can be decreased only if one uses a finer spatial classification, for example, defining settlements instead of counties as regions. Of course, the finest possible classification is determined by the dataset to hand. Finally, the Theil Index (and also its relative version) is completely insensitive to the topographic location of regions. For instance, the three main centers of Hungarian metallurgy, Budapest, Dunaújváros and

[4]For a formal proof of the result, see Combes, Mayer and Thisse (2008).

Miskolc are relatively far from one another. If one were to put these three cities next to each other (while still being classified into three different regions), the degree to which Hungarian metallurgy is concentrated would definitely increase. However, the value of the Theil Index would remain the same because it does not take regions' locations into account.

Another main weakness of the Theil Index and other standard indices of spatial concentration like the Herfindahl Index or the Isard Index is pointed out by Ellison and Glaeser (1997). Their claim is that if one wants to examine spatial concentration of employment in a specific sector, one cannot forget about this sector's industrial concentration. For example, only 22 plants operated in the Hungarian coal mining industry in 1994. If we calculate the Theil Index of this sector, defining Hungarian settlements as regions, it virtually cannot achieve its lowest value because 22 plants cannot be divided equally to 3125 settlements. By contrast, the number of coal mining plants was as many as 112 in 1986. Although this is still smaller than the number of regions, the Theil Index of coal mining could have achieved much lower values in 1986 than eight years later. Therefore, it seems unfair to compare these two indices to one another. To make such comparisons meaningful, Ellison and Glaeser derive a new concentration index – called the *Ellison–Glaeser* (EG) *Index* – from a probabilistic model of firms' location choices. The EG Index of sector s can be calculated in the following way:

$$EG^s = \frac{\frac{\sum_{r=1}^{R}(\lambda_r^s - \lambda_r)^2}{1-\sum_{r=1}^{R}\lambda_r^2} - H^s}{1 - H^s},$$

where λ_r^s is sector s's employment in region r relative to the sector's total employment in the country, λ_r is total employment in region r relative to total employment in the country, and H^s is the Herfindahl Index measuring industrial concentration in sector s as follows:

$$H^s = \sum_{j=1}^{N} \left(z_j^s\right)^2,$$

where N is the number of sector s plants in the country, and z_j^s is the number of employees in plant j.

The expected value of the EG Index is 0 if industry s' employment is distributed among regions as if it had been drawn from a random process (with the probability of falling into region r proportional to λ_r). Similarly to standard measures of spatial concentration, greater values of the index imply a higher degree of concentration.

Ellison and Glaeser argue that, since the EG Index explicitly takes into account industrial concentration in the specific sector, comparisons across countries, across industries and over time remain meaningful when using this type of concentration index even if the degree of industrial concentration across countries, industries or time periods is different. It is also true that the EG Index, similarly to the Relative Theil Index, is not affected by the fact that regions do not have the same size. Yet, the other two biases stemming from the MAUP apply even for the EG Index: first, although Ellison and Glaeser (1997) claim that the index is robust to what level of spatial aggregation one uses, Feser (2000) shows that this is not true in general. Even more importantly, the EG Index is as insensitive to the location of regions as standard indices of spatial concentration.

Using the EG Index to describe spatial concentration is, of course, based on the belief that industrial concentration is a nongeographic phenomenon that needs to be taken out. Although this is true in the location choice model of Ellison and Glaeser, it may not be true in reality. If, for instance, not plants but individual workers are the agents who choose their location and plants' formation is just a consequence of this process, H^s becomes endogenous and the EG Index ceases to be a theoretically supported spatial index. It is in fact likely that to what extent workers and/or plants make location decisions is different across sectors. As a consequence, the relationship between standard indices and the EG Index should be viewed as rather complementary and, to characterize spatial concentration of employment, I always calculate both the Theil Index and the EG Index in this paper.

4.2 The Duranton–Overman Continuous Approach

Duranton and Overman (2005) propose a completely different method in order to solve one of the previously mentioned problems, that is, to take into consideration the fact that not administrative units but actual geographic locations of plants in physical space are those which determine the degree of spatial concentration. To this end, they measure all the pairwise distances between manufacturing plants in the UK and plot the density of these distances for each industry (up to the median distance). In order to control for possible measurement errors, they smooth out empirical densities by a Gaussian kernel method.

Next, they generate artificial densities for each industry by dividing plants randomly among all possible locations. Using a bootstrap method, they can then determine whether a specific industry is significantly different (i.e., more concentrated or more dispersed) from what would be implied by randomness at each and every distance. The fact, for example, that an industry is locally concentrated at distances from 0 to 50 km but locally dispersed at distances above 50 km is likely to imply that plants in this industry are clustered in one area with a diameter of 50 km (relative to overall manufacturing). However, if there are two intervals of concentration (say, between 0 and 50 km and at about 100 km), this probably means that at least two clusters are present and these clusters lie 100 km apart from one

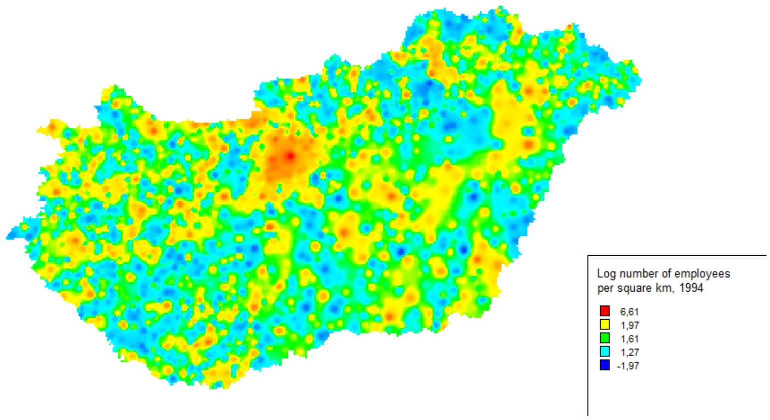


Figure 2: Log number of employees in the HWS per square km in 1994

another. Duranton and Overman (2005) also develop a test for global concentration and find that, relative to overall manufacturing, about half of UK manufacturing industries are globally concentrated.

4.3 Comparing Regions

Besides measuring changes in concentration, comparing relative positions of regions before and after the transition is another aim of this thesis. This can be done most simply by creating rankings of regions with respect to the variable of interest (e.g., employment) for both the pre- and the post-transition period and describing changes in regions' positions in the list.

Thematic maps constitute another useful tool. For instance, regions which experienced the biggest increase in employment can be shown in deep blue whereas regions on the other end of this ranking are in pale blue in the map of the country. Of course, such maps also suffer from the issue of discretization mentioned in Chapter 4.1: values of the specific variable such as employment are treated as evenly distributed within regions' boundaries even if this is not the case in reality. Consequently, the map can display large differentials at some regions' borders (e.g., if a "pale blue" region is next to a "deep blue" region), suggesting that these radical changes were real.

This problem can be solved in an elegant way. Using interpolation techniques, Geographical Information Software (GIS) are able to "smooth out" large differences and create a map which looks as if changes were continuous. Such a spatially smoothed map can be seen in Figure 2.

5 Data

In the analysis of the changes in regional disparities during the transition, I build primarily on a linked employer-employee dataset called the *Harmonized Hungarian Wage Survey* (*bértarifa-felvétel*, HWS). The first HWS took place in 1986, followed by another in 1989. Since 1992, the data have been collected each year. The target group of the survey consisted of Hungarian public firms (independently of size) and private enterprises with more than 20 employees until 1994; since 1995, private firms with 10 to 20 employees have also been included. Firms in the target group are required to provide data on a randomly selected sample of their full-time employees. For firms with multiple sites, data collection has to be carried out separately at each site. Appendix A describes the scope of the survey in more details.

First of all, I had to decide which years to include in the investigation. I chose using the data till 1994 – and, in some cases, till 1996 – for three reasons. First, as already described in Section 3, most of the changes attributed to transition had already taken place until 1994. Second, as cited in Section 2, evidence by Nemes-Nagy and Németh (2003) suggests that changes in the Hungarian economy's spatial distribution lasted until 1994 and the economy's geographical structure has been relatively stable since then. Finally, changes in the target group of the survey may cause biases in comparisons of indicators before and after 1995. For the same reasons, I decided to omit 1993 from the analysis as data on the public sector are not available for this year.

Variables observed include each worker's wage, bonuses, education level, working hours, occupation etc., as well as each firm's sectoral classification, corporate form, foreign capital share, and so on. By combining the HWS with a dataset on Hungarian firms' balance sheets for each year, even more variables such as firms' net revenues, income, exports, public capital share, etc. have become available – unfortunately, these latter variables were not recorded at the plant level, so I had to split them among sites in a way that the value of these variables at the particular site is proportional to the number of employees at the site. Most importantly for us, however, each worker is associated with a settlement ID which tells us in which of the more than 3000 Hungarian settlements the worker's site is located.

Hungarian settlements are classified into groups called *subregions* (*kistérség*), corresponding to the NUTS-4 level. The number of subregions changed in 1997 from 138 to 150, changed again in 2004 to 168, and yet again in 2007 to 174 (HCSO, 2009); I always use the classification with 150 subregions. Subregions contain 1 (in the case of Budapest) to 79 settlements. NUTS-3 spatial units are called *counties* (*megye*); there has been 19 of them since 1950. Budapest does not administratively belong to any of the counties, so I regard

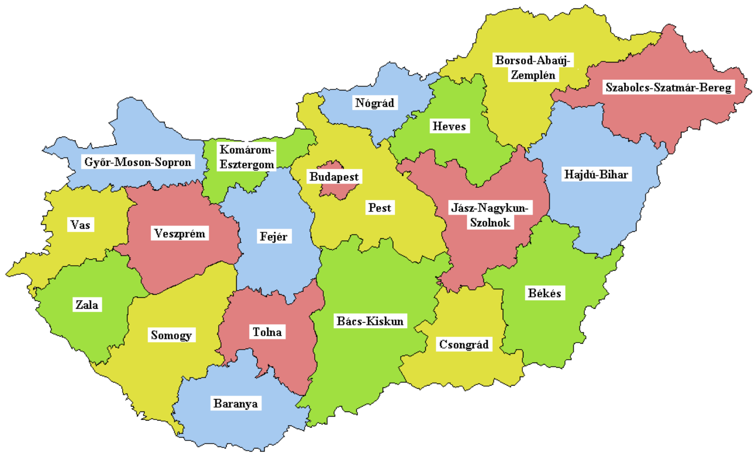


Figure 3: Counties of Hungary

it as the 20th county of Hungary for simplicity. Although the number of counties did not change during our period of interest, their borders slightly did; as a result, I use exclusively the classification as of 1997 (which is therefore consistent with the NUTS-4 classification with 150 subregions). See Figure 3 for a map of counties and Appendix B for a list of subregions.

Several weaknesses of the HWS, as well as attempts to solve these problems, are reported in HWS (2006). Besides the fact that many firms report false data, the most striking problem is that the sample of firms included in the survey – as opposed to the sample of workers *within* observed firms – is not random. Hence, nothing guarantees that observed workers and firms actually represent the total working population. To handle this problem, worker- and firm-specific weights have been created; for a detailed description of these weights, see HWS (2006). The fact, for instance, that the weight of a specific worker appearing in the survey is seven means that this person represents seven actual workers who are employed at firms of similar size *and* in the same sector. As pointed out in HWS (2006), this weighting procedure must contain errors because the sum of weights for the whole country markedly differs from the actual employment figure reported by the Hungarian Statistical Office.

Table 1

Employment shares of firms not included in the Hungarian Wage Survey

Sector	Country-level employment share (%)	Standard deviation of county-level employment shares (%)
Agriculture	3.8	1.9
Mining	1.9	35.2
Food	3.6	1.5
Textiles & paper	6.1	3.0
Chemicals	3.2	12.7
Construction materials	3.9	13.1
Metallurgy	1.2	14.3
Machinery	7.3	3.3
Energy	0.3	0.7
Construction	15.0	5.0
Trade	21.2	3.7
Transportation	4.2	4.9
Communication	0.6	35.4
Other services	15.7	5.8
Health, educ & culture	29.3	15.2

Besides these errors, however, it is even more important for us that weights are not calculated differently for different geographic regions. Therefore, if the share of those firms which are not included in the survey varies a lot across regions, concentration measures as well as regional rankings become biased. Unfortunately, there is no way to mitigate this bias[5]; however, it is possible to say something about its magnitude. Using the balance sheets of Hungarian firms, I calculated the employment of firms not included in the HWS as a share of total employment for each of the 20 Hungarian counties, each of the 15 sectors (according to the sectoral classification of the balance sheet dataset) and each year between 1992 and 1996.[6] For a specific industry, the above mentioned bias is small if (1) only a very small fraction of total employment belongs to not included firms, or (2) there are no large differentials in not included firms' employment shares across regions. To see if either (1) or (2) holds, I calculated excluded firms' country-level employment share as well as the

[5] Recalculating weights in such a way that they reflect regional characteristics is the only possible solution. This is far behind the scope of this thesis, but is a necessary condition for more reliable future research.

[6] I did not include 1986 and 1989 because balance sheet employment data are not reliable for these years.

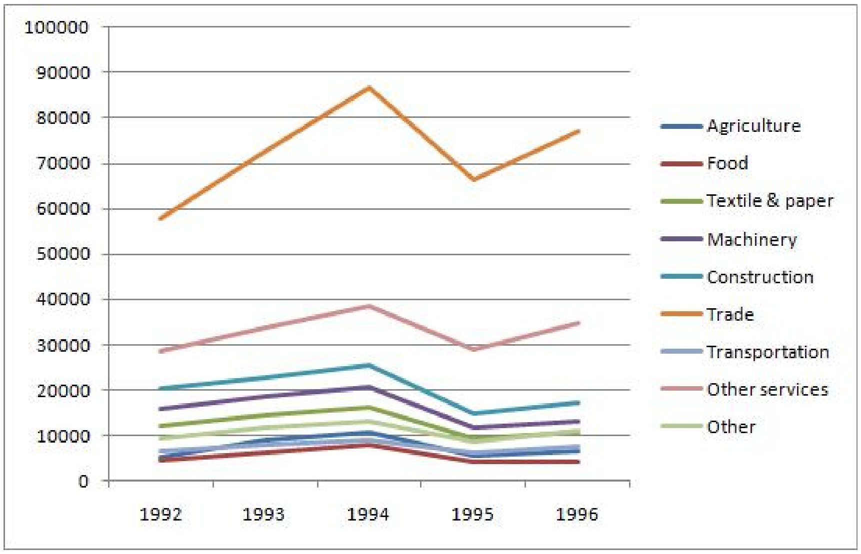


Figure 4: Number of small enterprises, by sector

standard deviation of county-level employment shares for each sector. Table 1 reports the time averages of these two measures. If one requires any of the two measures not to exceed 5%, all the sectors except the last one fulfil this requirement. Both measures are below 5% in the case of agriculture, food production, energy and transportation. Thus, one can conclude that, although this type of bias is present, it does not seem to be extremely large.

Size limits for firms applied in the HWS make our estimates biased also in another way. It is very likely that the smallest settlements have only firms with less than 20 (or even less than 10) workers – in most of the sectors, at least – implying that these settlements are associated with 0 employment. In the case of these "hidden settlements", any sort of weighting is useless as any multiple of 0 is still 0. As employment is actually positive for at least some of these settlements, calculated values of country-level concentration indices will be bigger than in reality. Fortunately, this does not affect time comparisons if the number of "hidden settlements" is stable over time. Although we do not know anything about these settlements, Figure 4 shows that the number of small enterprises (i.e., those with 20 employees till 1994 and those with 10 employees in 1995 and 1996) remained about constant for most sectors, implying that the probability of a settlement being "hidden" remained the same, too. This is, however, not true for some industries such as trade, in which the number of small firms was increasing so the expected number of "hidden settlements" probably raised as well, leading to an upward bias in the estimated change in concentration of this sector. In addition, bias can arise in all industries for the years 1986 and 1989, from which we do not know the number of small enterprises. Therefore, I decided to control for the bias by calculating concentration indices based on subregion-level instead of settlement-level data; if one does so, the effect of "hidden settlements" can be expected to be washed out by other settlements in their vicinities. At the same time, of course, the bias implied by the MAUP becomes larger in this case than if one uses the finest possible spatial classification and defines settlements as regions.

As already mentioned, the HWS assigns a settlement ID to each employee appearing in the survey. This implies another problem to be solved. Specifically, the names and areas of some settlements have changed during the 1986 to 1996 period. These changes can be classified into four groups. First, some settlements have changed their names without anything else changing; this does not call for any correction as the ID's of these settlements remained the same. Second, some settlements which had been united during the Socialist era were split into two, with the newly formed settlements receiving new ID's. In this case, to make time comparisons meaningful, I divided the employees observed in the still united settlement between the two new settlements, with their weights being modified according to population shares in the year of split. For instance, Kerepestarcsa was split into two villages (Kerepes and Kistarcsa) in 1994. Therefore, a common settlement ID was assigned to employees working in Kerepestarcsa until 1994 but one of two different ID's (the ID of Kerepes or that of Kistarcsa, both different from the previous ID) after that year. Consequently, I changed the settlement ID's of workers in Kerepestarcsa before 1994, assigning either a "Kerepes" ID or a "Kistarcsa" ID to them such that the ratio of total employments (in each sector) correspond to the ratio of population of the newly formed villages in 1994.[7]

Some settlements seceded from bigger ones. In these cases, the seceded settlement received a new ID while the big settlement continued to have the common ID. It could have been, in theory, possible to divide the employees observed before the secession between the two settlements in the same fashion as previously. However, as only 51 such settlements exist and their population is small (36 of them have less than one thousand inhabitants, and even the biggest one has below 3000), I decided to omit this step as the bias must be very small anyway.

Finally, the administrative borders of some settlements changed. This may have affected concentration indices in two ways: (1) some workplaces may have been classified into new settlements, and (2) in the case of topographic indices, values of the benchmark variable, i.e., settlements' areas, may have changed. As for (1), administrative borders of settlements are usually uninhabited, so changes of borders are unlikely to have any effect on workers' location. Concerning (2), the correlation coefficient of settlements' areas in 1990 and 1996 is

[7] Data on settlements' population and area comes from the Hungarian Statistical Office's *T-STAR* database.

0.9965, implying that these changes must have been minor ones. Therefore, I always apply the 1996 areas of settlements when calculating indices of topographic concentration.

Table 2

Sector matching indicators

Sector	K1 (%)	K2 (%)	Sector	K1 (%)	K2 (%)
Agriculture & food	96	100	International trade	98	92
Forestry & wood	98	99	Tourism	99	94
Coal mining	97	100	Road transport	96	95
Oil extraction	100	100	Railway transportation	100	100
Other mining	94	98	Local transport	100	92
Textile & clothing	97	97	Communication	100	100
Paper	97	88	Financial services	100	99
Printing & publishing	78	91	Insurance	100	100
Refined petroleum products	100	100	Informatics	90	83
Chemicals	98	97	Other services	84	89
Non-metallic minerals	97	98	Public education	100	100
Metallurgy	85	98	Higher education	100	100
Machinery	96	94	Health care	99	100
Furniture & other products	98	68	Public administration	100	100
Energy	87	100	Research & development	100	89
Water supply & management	96	91	Culture	100	90
Construction	100	96	Other	100	78
Domestic trade	99	93			

Let us turn now to the classification of sectors. In 1992, a new industrial classification system was introduced, consisting of 353, 4-digit categories instead of the previously used system with 258 categories. Of course, I would need a unified system to ensure comparability across years before and after 1992. Fortunately, this work has already been done by Kertesi and Köllő (1997). They define 35 relatively broad sectors (listed in Table 2) and assign each pre-1992 and each post-1992 category to one of them. Using a full dataset on Hungarian firms between 1990 and 1992, they also check the correctness of the procedure by calculating two indicators for each of the 35 sectors: the share of workers who belong to the sector according to their post-1992 industry in the population of workers who belong to the sector according to their pre-1992 industry (which they denote by $K1$), and the share of workers who belong to the sector according to their pre-1992 industry in the population of workers who belong to the sector according to their post-1992 industry (denoted by $K2$). In the case of a perfect match, both numbers are equal to 100%. As can be seen in Table 2, this is unfortunately not always the case. Therefore, to avoid big errors, I decided to concentrate only on those sectors for which both $K1$ and $K2$ are 85% or greater.

Due to other problems, I had to decrease the number of sectors included in the analysis even further. In 1986, no firms are associated to two sectors, namely, financial services and insurance, making time comparisons practically impossible for these industries. Finally, separate plants within a settlement cannot be identified for education, health care, administration, research and development and culture for 1986 as no firm identifiers are provided in the case of these sectors. Hence, Ellison–Glaeser Indices cannot be calculated in 1986 for these industries. As a consequence, I decided to omit them as well, and do the analysis only for the remaining 22 sectors.

6 Results

6.1 Changes in Regional Disparities in Hungary

In this section, I start presenting the results of the thesis by describing tendencies of spatial concentration and reordering in the overall economy. One can see the evolution of the Theil Index and the Ellison–Glaeser Index (both relative to area, that is, characterizing topographic concentration) as well as the within-county and between-county components of the Theil Index in Table 3. The tendencies which can be seen in the table inform us about a definite increase in spatial concentration of the Hungarian economy. Looking at the Theil Index first, it increased by about 0.228 throughout the 10-year period, and, in particular, by almost 0.3 from 1986 to 1994. If one compares this to, for instance, the results of Brülhart and Traeger (2005) who find a change of 0.161 (in absolute value) for manufacturing and even smaller changes for other sectors, as well as a decrease of 0.002 (!) for the entire economy in the case of 17 Western European countries throughout a 25-year period, one can conclude that the Hungarian economy went through radical changes in its spatial structure.

Table 3

Spatial indices of employment (all relative to area)

Year	Theil Index	Within-county	Between-county	EG Index
1986	1.499	0.838	0.660	0.045
1989	1.535	0.885	0.651	0.052
1992	1.592	0.872	0.720	0.060
1994	1.784	0.934	0.850	0.075
1995	1.699	0.904	0.795	0.069
1996	1.727	0.880	0.847	0.075

Of course, one could argue that this result is, at least partially, due to the fact that large socialist firms were split up, hence the number of plants increased in the transition period. (In particular, the number of plants in the HWS rose from 15,652 in 1986 to 18,796 in 1994.) Therefore, in order to control for such changes in industrial structure, I calculated EG Indices for each year; these can be seen in the last column of Table 3. As this type of index also raised definitely, we can conclude that the increase in the Theil Index is not caused by changes in industrial concentration. In other words, spatial concentration raised during the transition, no matter if we think of employees or plants as decisionmakers in the problem of location choice.

Turning to the within and between components of the Theil Index, one can see that both of them increased but the between-county component did to a larger extent. The share of the between component went up from 44% in 1986 to 49% in 1996. Hence, although both within- and between-county concentration were essential in the entire period, reordering *across counties* turns out to be the primary reason for the increase in concentration of the country's economy.

Some words on the negative change in all the indices between 1994 and 1995 are in order. Notice that this might be caused by the fact that even firms with 10 to 20 employees have been included in the Wage Survey since 1995. However, recalculating the indices after dropping these firms from the database only slightly increases the Theil Index of 1995 and even decreases the EG Index. This means that other factors must have been responsible for the decrease in spatial indices between 1994 and 1995. What these factors were may be a question for future research.

Although the examination of "income-type" indicators such as workers' compensation, firms' revenues, etc. are not the primary aim of this thesis, at this point it is worth looking at the changes in concentration of these variables as well. I calculated Theil Indices relative to area for two variables which are available in the HWS database: (1) workers' wagebill, and (2) firms' revenues. Concerning the former, the value of the index happens to have gone up from 1.622 (in 1986) to 2.045 (in 1994), with the within-county component raising from 0.842 to 0.936 and the between-county component increasing from 0.779 to 1.109. That is, spatial concentration of wages definitely went up during the transition, and the change is mainly due to between-county concentration, so much so that this component became larger in magnitude than the within-county part. As for (2), the Theil Index of firms' revenues also increased (from 1.999 to 2.345), with the between Theil Index raising by more than 0.4 (from 0.895 to 1.309) and the within Theil Index falling to some extent (from 1.104 to 1.037). In sum, not only employment but also income-type measures became more concentrated throughout the period of interest, and, again similarly to employment, these changes primarily originated from the between-county level. All these results are in line with previous findings such as those of Nemes-Nagy (2004) and Major and Nemes-Nagy (1999).

Finally, the above results seem to be robust to spatial classification. In particular, by calculating all the indices by using subregions instead of settlements as spatial units, the same patterns arise. Thus, the bias due to "hidden settlements" cannot be the reason for the findings.

Table 4

Rankings of counties with respect to employment

County	Employment share, 1986 (%)	County	Employment share, 1994 (%)
Budapest	22.6	Budapest	27.2
Borsod-Abaúj-Zemplén	7.6	Borsod-Abaúj-Zemplén	6.5
Pest	6.1	Pest	5.4
Szabolcs-Szatmár-Bereg	5.2	Hajdú-Bihar	4.7
Hajdú-Bihar	5.1	Győr-Moson-Sopron	4.6
Bács-Kiskun	5.0	Bács-Kiskun	4.6
Győr-Moson-Sopron	4.5	Csongrád	4.5
Csongrád	4.4	Baranya	4.4
Baranya	4.2	Szabolcs-Szatmár-Bereg	4.3
Jász-Nagykun-Szolnok	4.1	Fejér	4.0
Fejér	3.8	Jász-Nagykun-Szolnok	3.7
Veszprém	3.5	Komárom-Esztergom	3.4
Zala	3.5	Veszprém	3.4
Békés	3.5	Békés	3.4
Komárom-Esztergom	3.2	Vas	3.1
Somogy	3.1	Zala	3.0
Heves	3.0	Somogy	2.9
Vas	2.8	Heves	2.7
Tolna	2.6	Tolna	2.3
Nógrád	2.1	Nógrád	1.9

Let us turn now to relative positions of regions. Table 4 presents the rankings of the 20 Hungarian counties with respect to their employment shares (i.e., employment in the county as a fraction of total employment in Hungary) in 1986 and 1994. What is most striking about this table is that counties which gained positions are almost exclusively Transdanubian ones (i.e., they are located in the Western part of the country) whereas counties which lost positions are almost all in the East. More precisely, out of the nine Transdanubian counties, five increased their ranks, one of them (Tolna) stayed in 19th place and only three (Somogy, Veszprém, and Zala) lost one to three positions; by contrast, three of the nine Eastern counties went down in the list, four kept their ranks and only two (Csongrád and

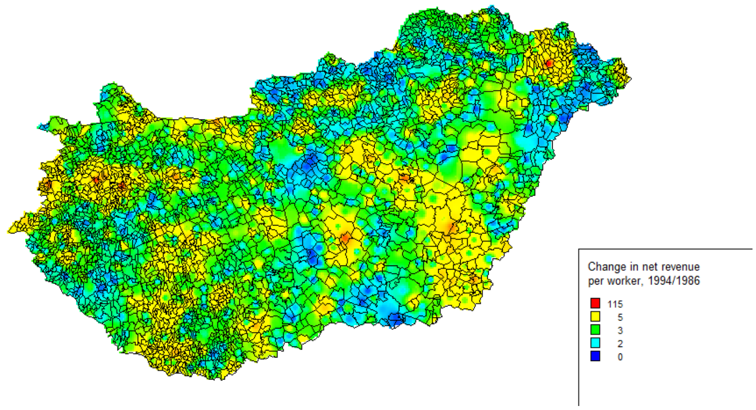


Figure 5: Changes in net revenue per worker between 1986 and 1994

Hajdú-Bihar) managed to improve by one position. (Budapest and Pest county are usually regarded as neither Western nor Eastern regions; they remained in the first and third place, respectively.) This confirms the definite shift of economic activities from the Eastern part to the Western part (as well as to the capital city) of Hungary which has been already documented in the literature (Nemes-Nagy, 2004; Quadrado et al., 2001; Barta, 2003).

The East-West shift can be even more clearly seen from the changes in some "income-type" indicators, specifically, net revenues and profits of firms. In the rankings of counties with respect to net revenues per worker, six Transdanubian but only two Eastern regions gained some positions. As for profits (before tax) per worker, the same numbers are seven and two, respectively. As a result, the top seven counties in 1994, apart from Budapest, with respect to net revenues per worker were exlusively Transdanubian and the same is true for profits per worker, with the only exception being Jász-Nagykun-Szolnok county in sixth place. (See Appendix C for these rankings.)

To tell more about the changes, let us look at a spatially smoothed map as well. Figure 5 presents the spatial distribution of changes in net revenue per worker at the settlements' level, smoothed out in the way described in Section 4.3. In line with previous results, one can see that the change in Eastern regions is generally smaller than that in Western regions, since the blue color is overwhelming in the Eastern part of the country (and especially in the North), but not in Transdanubia. However, even some Western territories seem to have experienced improvements which remained below the average, whereas some Eastern regions did a good job even relative to the entire economy. That is, in spite of the fact that the shift of economic activities and incomes from the East to the West is an existing phenomenon, one cannot completely forget about the fact that there have been very different improvements

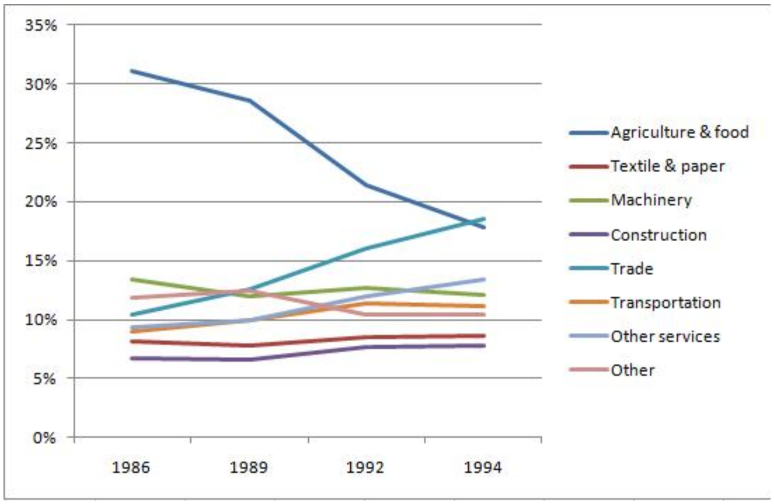


Figure 6: EMPLOYMENT SHARES OF EXAMINED SECTORS IN THE HWS

even *within* these large regions.

6.2 Changes in the Spatial Distribution of Industries

I present results related to the spatial reordering of specific industries in this subsection. First of all, it is worth looking at the country-level employment shares of industries (that is, their employment in Hungary relative to total employment in all sectors examined) in order to learn which of the sectors were more and which were less important in the entire economy. The evolution of these employment shares throughout the examined period is presented in Figure 6. It can be seen from the figure that most employees worked for agricultural, food producing, machinery and trade firms (as well as firms providing other services) between 1986 and 1994.[8] Besides this, it can also be seen that the shares of some sectors, especially those of agriculture / food and trade, saw dramatic changes during the transition. However, even if these phenomena are fascinating, they are not what this thesis attempts to investigate; what I care about is the shares of different Hungarian *regions* in specific sectors' employment and not the shares of these *sectors* in total employment of the country.

[8] In this subsection, I report all the results only until 1994; I calculated all the indices and employment shares for 1995 and 1996 but they did not seem to yield much additional information. Similarly to concentration indices of the entire economy, one can see slight decreases in most sectors' spatial indices in 1995 and slight increases in 1996.

Table 5

Employment Theil indices of sectors

	Area Theil				Employment Theil			
Sector	1986	1989	1992	1994	1986	1989	1992	1994
Agriculture & food	0.725	0.798	0.916	1.054	0.498	0.532	0.647	0.682
Forestry & wood	2.541	2.453	2.255	2.249	1.591	1.492	1.495	1.448
Coal mining	4.242	4.558	4.917	5.382	2.643	2.813	3.408	3.644
Oil extraction	4.101	4.219	4.945	4.679	2.520	2.780	3.565	2.325
Other mining	4.240	5.085	4.821	4.857	4.007	4.298	4.192	3.914
Textile & clothing	2.145	2.082	2.004	2.063	0.594	0.558	0.590	0.571
Paper	4.425	4.391	3.843	3.874	2.157	1.983	1.634	1.687
Refined petroleum products	6.439	6.412	4.005	3.475	4.929	4.879	2.264	1.452
Chemicals	4.193	4.223	4.042	3.968	1.857	1.947	1.877	1.725
Non-metallic minerals	3.031	3.120	3.232	3.460	1.684	1.900	1.797	1.921
Metallurgy	4.056	4.124	4.313	4.154	1.920	1.924	2.075	2.055
Machinery	2.454	2.313	2.267	2.487	0.514	0.466	0.417	0.514
Energy	2.605	2.592	2.818	3.097	0.725	0.692	0.733	1.070
Water supply & management	2.513	2.502	2.477	2.427	0.683	0.653	0.679	0.475
Construction	2.647	2.772	2.429	2.567	0.508	0.465	0.397	0.398
Domestic trade	1.849	2.045	2.094	2.577	0.344	0.277	0.272	0.325
International trade	4.804	4.947	4.666	4.416	1.278	1.378	1.218	1.246
Tourism	2.910	3.030	3.038	3.348	0.509	0.568	0.714	0.777
Road transport	2.858	2.810	2.892	2.785	0.441	0.429	0.556	0.420
Railway transportation	3.050	2.885	2.780	2.626	0.956	0.822	0.869	0.672
Local transport	4.846	4.709	4.641	4.660	1.297	1.245	1.216	1.025
Communication	2.435	3.893	2.264	2.363	0.470	1.116	0.339	0.304

To tell more about the concentration of regions' employment at the level of industries, I calculated Theil Indices (both relative to area and relative to total employment) for each year and industry. The results, reported in Table 5, are interesting for multiple reasons. First of all, the degree to which employment is concentrated is very different across industries both before and after the transition. In every year, agriculture and food production is the most dispersed sector topographically while mining, some manufacturing industries such as paper, chemicals and metallurgy, as well as international trade and local transport are the most concentrated ones. Regarding other "big" sectors besides agriculture, machinery is around the middle in the ranking while domestic trade is somewhat less concentrated than manufacturing, except in the last year observed.

As for concentration relative to total employment, domestic trade is found to be the most dispersed industry, i.e., this is the sector for which the distribution of sectoral employment is the closest to the distribution of total employment in Hungary. Domestic trade is followed by most of the service sectors as well as agriculture, machinery and textiles, while one can again find mining industries on the other end of the list. These findings more or less coincide with those of Brülhart and Traeger (2005) in the case of Western European countries, although they find agriculture to be the most concentrated sector relative to total employment. However, such a direct comparison cannot really be made for multiple reasons. First, Brülhart and Traeger examine more countries in lieu of only one. Second, they use NUTS-2 or NUTS-3 regions as the spatial units of observation instead of settlements. Finally, and probably most importantly, they can separate agriculture from food production as they have a somewhat different industrial classification.

The second set of conclusions which can be drawn from Table 5 concerns *changes* in concentration. Throughout the entire period, 10 industries became more concentrated whereas 12 became more dispersed in physical space. In the case of concentration relative to total employment, the same numbers are 7 and 15, respectively.[9] Some sectors underwent radical changes such as the refined petroleum products industry for which the Theil Index fell from 6.439 to 3.475 (and from 4.929 to 1.452 relative to employment), or domestic trade for which the value of the index went up from 1.849 to 2.577 (although it slightly decreased relative to employment). That is, one can observe huge intersectoral heterogeneity behind the overall trend of increasing spatial concentration.

I also calculated within- and between-county Theil Indices for each industry and year. Between-county components turn out to be generally smaller than within-county components, both in 1986 and in 1994, apart from some small service sectors. The production of non-metallic minerals, domestic trade and tourism experienced the biggest increases in the share of the between component, that is, these are the sectors in which cross-county differentials became definitely more important than before.

As a robustness check, I repeated all the calculations with subregions, similarly as in Section 6.1. The rankings of industries as well as the changes in their concentration are basically the same as what could be observed from Theil Indices based on settlement-level

[9] These numbers seem strange at first glance as one would expect industries with rising concentration to be more numerous than those with rising dispersion in order for the overall concentration of the economy to increase. However, notice that sectors do not, of course, have the same size and the "biggest" sectors, i.e., agriculture, domestic trade and machinery all became more concentrated in physical space.

data.

The problem with the Theil Index is, as also in Section 6.1, that it does not control for differences in industrial concentration while this sort of concentration definitely changed in most sectors. For instance, the number of plants operating in the refined petroleum products industry was only 10 in 1986 but increased to 74 until 1994. This resulted in a drop in the sector's Herfindahl Index from 0.249 to 0.105. The change in Herfindahl Indices from 1986 to 1994 was above 0.05 (in absolute value) in eight sectors out of 22. Therefore, it seems to be essential to calculate EG Indices in order to take out the effect of these large changes. At the same time, of course, not only comparisons across years but also comparisons across sectors with different levels of industrial concentration become more correct by using the EG Index.

Table 6

Ellison-Glaeser Indices of sectors

Sector	Area EG 1986	Area EG 1989	Area EG 1992	Area EG 1994	Employment EG 1986	Employment EG 1989	Employment EG 1992	Employment EG 1994
Agriculture & food	0.004	0.003	0.003	0.005	0.028	0.030	0.038	0.048
Forestry & wood	0.008	0.002	0.008	0.005	0.026	0.029	0.027	0.049
Coal mining	0.006	0.003	0.000	0.080	0.054	0.055	0.073	0.167
Oil extraction	0.052	0.031	−0.005	0.065	0.024	0.029	0.107	−0.012
Other mining	0.009	0.001	0.003	0.003	0.061	0.065	0.043	0.067
Textile & clothing	0.042	0.040	0.033	0.046	0.001	0.000	0.005	0.002
Paper	0.060	0.009	0.062	0.095	−0.034	−0.079	−0.019	−0.011
Refined petroleum	0.072	0.000	−0.003	−0.003	0.169	0.100	−0.028	−0.087
Chemicals	0.091	0.076	0.069	0.098	0.008	−0.006	−0.007	−0.006
Non-metallic min	0.016	0.011	0.019	0.024	0.018	0.014	0.019	0.062
Metallurgy	0.048	0.032	0.056	0.057	0.025	0.013	0.014	0.063
Machinery	0.101	0.083	0.069	0.080	0.009	0.003	−0.001	−0.001
Energy	0.031	0.035	0.039	0.030	−0.006	−0.007	−0.006	−0.003
Water supply & m	0.033	0.024	0.021	0.041	−0.007	−0.009	−0.007	−0.021
Construction	0.091	0.131	0.096	0.119	0.003	0.016	0.003	0.003
Domestic trade	0.045	0.069	0.078	0.120	0.000	0.000	0.001	0.002
International trade	0.824	0.895	0.759	0.311	0.491	0.557	0.429	0.073
Tourism	0.136	0.146	0.161	0.235	0.017	0.015	0.021	0.041
Road transport	0.136	0.120	0.134	0.100	0.008	0.004	0.006	−0.011
Railway transport	0.006	−0.004	−0.002	−0.003	−0.055	−0.051	−0.054	−0.066
Local transport	0.639	−0.036	−0.025	−0.034	−0.047	−1.235	−1.226	−1.356
Communication	0.052	−0.009	0.053	0.074	−0.089	−0.177	−0.047	−0.047

Table 6 shows the Ellison–Glaeser Indices for each sector and year, calculated either by using areas of settlements as the benchmark (i.e., in order to characterize topographic concentration), or by using total employment (in order to measure relative concentration). One can discover relatively many differences between Table 5 and Table 6. First, although agriculture was the most dispersed sector in physical space in 1986, it lost its position due to some service sectors' sharp drops in their spatial concentration. Second, while machinery was around the middle in the ranking of Theil Indices, according to the EG Index it is one of the industries most concentrated in space. Mining industries' EG Indices are also relatively low, at least compared to their top positions in the Theil Index rankings.

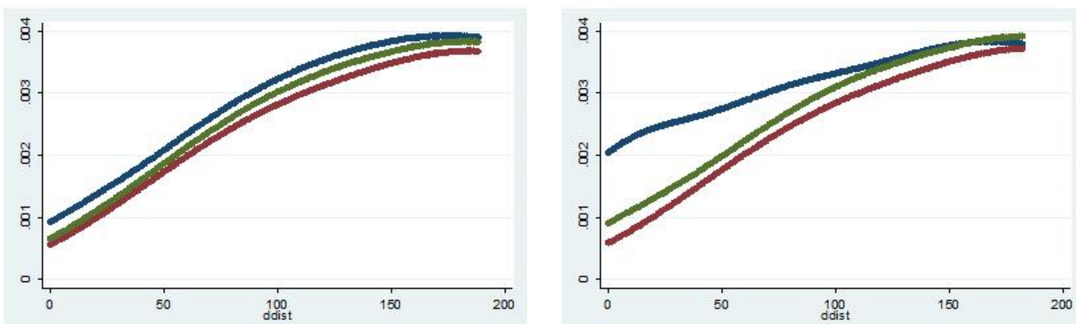
These findings can be explained by differences in industrial concentration across sectors: the Herfindahl Indices of mining sectors, especially that of coal mining, were high, boosting up the Theil Index above the actual degree of these industries' spatial concentration, while the industrial concentration of machinery was low both before and after the transition, resulting in low Theil Indices but higher EG Indices in the case of this sector. Fewer differences can be seen between EG and Theil Indices measuring concentration relative to employment. Domestic trade and some of the service sectors such as railway transportation and communication are the most dispersed industries relative to total employment, and they are followed by machinery, textiles and chemicals, similarly as with the employment Theil Index. Relative to employment, agriculture is rather concentrated than dispersed but, again similarly to Theil Indices relative to employment, mining sectors as well as – prior to the transition – the refined petroleum products industry can be found on the top of the list (together with international trade).

All in all, the degree of spatial concentration increased in 13 industries (both according to the area EG Index and the relative EG Index) which is a somewhat larger number than what we obtained by using the Theil Index. The sectors for which one can see increasing concentration include the "big" sectors of agriculture and domestic trade, but not machinery.

The correlation coefficient of industries' EG Indices between 1986 and 1994 happens to be 0.4584 for the EG Index relative to area and only 0.2043 for the index relative to total employment. What this result tells us is that the degree to which sectors were concentrated relative to each other changed dramatically. It is worth mentioning that, as opposed to these numbers, Dumais et al. (2002) find a correlation coefficient of 0.92 between the 1972 and 1992 values of EG Indices for the US.

To see if all these findings are robust to spatial classification, I applied the same check as before, i.e., I calculated EG Indices based on subregion-level in lieu of settlement-level data. As a result, the values of indices changed to some extent but not drastically. The only sector in which large differences can be found is coal mining (along with other mining and refined petroleum products in the case of the area EG index), although the tendency in the index remained the same even for this industry. This result, i.e., the robustness of Hungarian EG Indices to spatial aggregation, is especially appealing in light of Feser (2000) who shows that EG Indices calculated for Tennessee and North Carolina are very sensitive to changes in the level of spatial aggregation. He also finds that the difference between EG Indices calculated by using different spatial units is the biggest for those sectors which have (1) a high degree of industrial concentration, and (2) a low degree of spatial concentration. This can be the explanation of large differences seen in the case of coal mining which was a very concentrated sector – especially in the late transition period – but was, as can be observed especially from the EG Index relative to area, quite dispersed in space.

Finally, I also calculated EG Indices for each year and industry by using the specific sector's 1986 Herfindahl Index in 1989, 1992 and 1994 as well. One can expect to obtain tendencies which are similar to those of sectoral Theil Indices in this case, because changes in individual sectors' industrial concentration are not controlled for. The results more or less coincide with what one would expect based on the previous considerations, although the number of sectors experiencing an increase in concentration is somewhat greater than with the Theil Index.



Estimated kernel densities for manufacturing in 1986 (left) and 1994 (right)

Let us see now the results coming from the Duranton–Overman continuous method.[10] Figure 7 illustrates two empirical densities, the left one being the density of machinery for 1986 and the right one being the density of the same industry in 1994 – up to the median distance of Hungarian plants, as in Duranton and Overman (2005). The blue line represents smoothed density of the industry, while the red and green lines are the 5 and 95 percent threshold values to the confidence interval of randomness. If the density of the specific industry is between the two lines at a particular distance, one can conclude that the spatial distribution of plants is not significantly different from the random distribution at that distance. However, if the blue line is above the green one, the industry is found to be geographically more concentrated than what would be implied by randomness at the specific distance, and it is more dispersed than randomness if the blue line lies below the red one. Based on the figures, one can say that machinery was concentrated before the transition at any distance but this is no longer true for large distances after the transition.

[10] Due to computational problems, no densities could be estimated for agriculture and domestic trade.

However, there might exist other cases in which it is not clear how to compare two industries. For instance, what can we say if a sector was concentrated at small distances and dispersed at large distances in 1986, but the opposite held in 1994? Therefore, in order to find a comparable measure, I calculated the share of distances for which industries were dispersed and concentrated in both 1986 and 1994 and each industry. The share of intervals of concentration increased in the case of 10 industries, while that of intervals of dispersion increased for 8. Thus, we can conclude, that the Duranton–Overman continuous approach produces very similar results as indices of spatial concentration.

Table 7

Rankings of country parts with respect to employment, by sector

	Emp share, 1986 (%)			Emp share, 1994 (%)		
Sector	West	Central	East	West	Central	East
Agriculture & food	33.7	14.9	51.3	38.5	13.6	47.9
Forestry & wood	46.4	12.7	41.0	55.2	11.7	33.1
Coal mining	62.4	1.2	36.3	19.2	0.0	80.8
Oil extraction	23.2	18.9	57.9	32.6	26.4	41.0
Other mining	47.1	11.2	41.6	69.6	11.7	18.7
Textile & clothing	35.5	24.8	39.7	35.5	26.0	38.5
Paper	34.7	38.9	26.4	31.6	44.6	23.9
Refined petroleum products	32.7	50.3	17.1	39.9	31.0	29.1
Chemicals	32.0	33.1	34.9	32.8	38.1	29.1
Non-metallic minerals	33.2	14.6	52.2	53.9	9.6	36.5
Metallurgy	31.3	14.4	54.3	42.2	12.8	45.0
Machinery	27.1	39.8	33.1	30.3	37.7	32.0
Energy	37.4	23.1	39.5	45.0	23.0	31.9
Water supply & management	29.4	27.0	43.6	34.7	28.7	36.5
Construction	32.2	35.8	32.0	27.6	39.1	33.3
Domestic trade	30.4	26.5	43.1	24.0	38.2	37.8
International trade	2.1	92.7	5.2	40.4	55.1	4.5
Tourism	27.3	40.9	31.7	25.8	51.1	23.2
Road transport	28.8	43.4	27.8	29.4	38.9	31.7
Railway transportation	29.3	24.2	46.5	30.8	26.3	42.8
Local transport	0.0	91.4	8.6	2.3	85.2	12.5
Communication	24.6	42.7	32.7	27.4	39.9	32.7

Let us turn to the examination of regions' relative positions. If one were to look at the entire picture of changes in each industry's employment share in each region – that is, in each county, say –, one would need to analyze a three dimensional, 4 by 20 by 22 array of numbers. This seems to be, however, unnecessary as we can simplify the investigation in two possible ways. First, let us look at two points in time only, specifically, 1986 ("before transition") and 1994 ("after transition"), and let us divide Hungary into three parts instead of twenty counties. In fact, these "county parts" correspond to NUTS-1 regions. The country

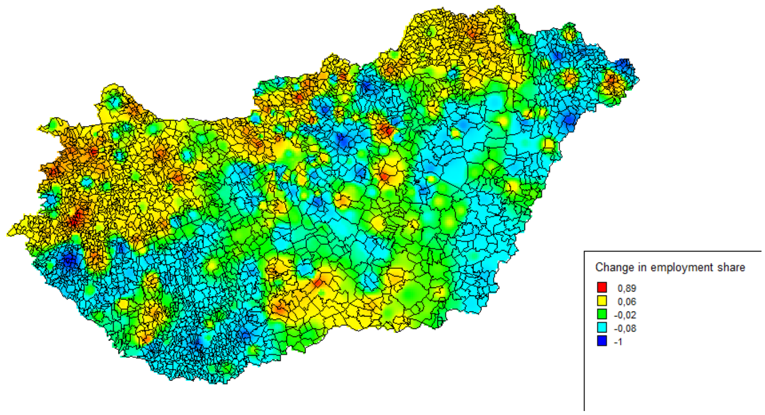


Figure 7: Changes in the share of machinery in total employment

part called *Transdanubia* consists of counties located West from the Danube; *Great Plain and North* includes counties East from the Danube; finally, *Central Hungary* includes Pest county and the capital. In this way, the 4 by 20 by 22 array has been reduced to a 2 by 3 by 22 array which is presented above (Table 7). The first observation that can be made based on the table is that, despite the fact that many industries, as seen from the investigation of spatial indices, became more dispersed, all of them except coal mining, the refined petroleum products industry, construction, road transport and local transport fell back in the East and all of them except coal mining, paper, construction, domestic trade and tourism improved in the West. As for the central region, it increased its share in ten industries, the most dramatic improvement being experienced in the case of domestic trade. To sum up, it seems that most industries followed the East-West shift of the overall economy, although there exist some exceptions. For "big" sectors, the most definite shift from East to West can be seen for domestic trade, followed by agriculture and food, while employment shares were almost constant in the case of machinery.

Using maps is the second simple way of comparing the changes in industries' regional shares. This also has the advantage that we can look at heterogeneity not only between country parts but also within them. Figures 8 and 9 present the changes in settlement-level employment shares for two industries, specifically, textiles and machinery. The East-West shift can quite clearly be identified from both maps. However, there were also other forces at work. In the case of machinery, settlements located in the Southern part of Transdanubia did not manage to improve in spite of the fact that they are located in the West. On the contrary, employment shares of machinery increased in some regions located in the Northern part of Eastern Hungary. As for textiles, the picture seems to be even more heterogenous: although

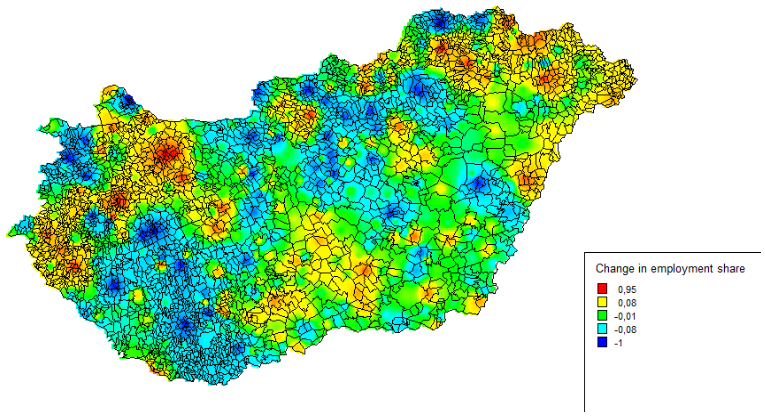


Figure 8: CHANGES IN THE SHARE OF TEXTILES IN TOTAL EMPLOYMENT

the Northwest generally experienced improvements, a large part of Transdanubia fell back in fact while the share of the textile industry rose in many settlements in the Northeast. What the determinants of these changes within country parts were is, of course, out of the scope of this paper.

6.3 The Effects of International Trade

This subsection uses simple methods to answer the second main question of this thesis, that is, whether Hungary's opening up to international trade, through differences in regions' and industries' openness before the process, might have played a role in the spatial reordering of industries documented in Section 6.2. Sectors and regions which had got used to competition by being more open might have managed to improve their positions relative to others during the transition; the question is, to put it differently, whether this effect might have been present and contributed significantly to spatial changes of industries during the transition.

I measure pre-transition openness to trade by the amount of 1986 exports per worker, coming from the balance sheet data on Hungarian firms.[11] Unfortunately, no data are available on imports; however, we have some, although not much, information on the location of trading partners of firms since the balance sheet dataset also includes the amount of rubel exports for each firm. Using this, I splitted total exports into rubel exports (i.e, those probably flowing to the Soviet Union) and non-rubel exports (i.e., those flowing to other, mainly Western, countries) and calculated the per worker amounts of these variables as well.

[11] As already mentioned in Section 5, export data were unfortunately not recorded for individual plants but only for firms. Therefore, I divided exports of firms among plants in such a way that the same amount is associated to each worker. Once we know export per worker for every plant in each settlement of the country, this allows us to calculate export per worker for settlements, subregions, counties, and so on.

I can then use the amount of rubel export per worker as a measure of openness to trade "with the East" and that of non-rubel export per worker as a measure of openness to trade "with the West" and analyze the effects of these two types of openness separately.

Table 8

Top 15 sectors with respect to export per worker in 1986

Sector	Export per worker, thousand HUF	Share of rubel exports, %
Refined petroleum products	1481	7.9
International trade	983	73.9
Informatics	412	71.0
Chemicals	370	47.4
Metallurgy	190	23.4
Machinery	116	67.7
Textile & clothing	103	42.5
Paper	97	42.3
Communication	91	14.0
Railway transportation	74	48.2
Road transport	73	13.7
Furniture & other products	60	25.4
Non-metallic minerals	49	26.6
Forestry & wood	45	4.1
Agriculture & food	43	40.5

It is worth looking at country-level openness of industries as a first step. One can see those 15 sectors (out of 37 which can be found in the Hungarian Wage Survey) in Table 8 which had the biggest amount of export per worker in 1986. Besides the amount of export per worker, I also report the share of rubel exports for each industry. One can already realize the signs of industrial heterogeneity by looking at Table 8: not only sectors' degree of openness was markedly different but also the extent to which these sectors were tied to the Soviet market. Whereas, for instance, two thirds of machinery exports went to the East, this share was below 10% in the case of wood and refined petroleum products.

The first idea is to use smoothed maps to see if "open" industries experienced different changes in their spatial patterns than "non-open" ones. Figure 10 shows the changes in spatial distribution of an "open" industry, chemicals (in which export per worker was 370.000

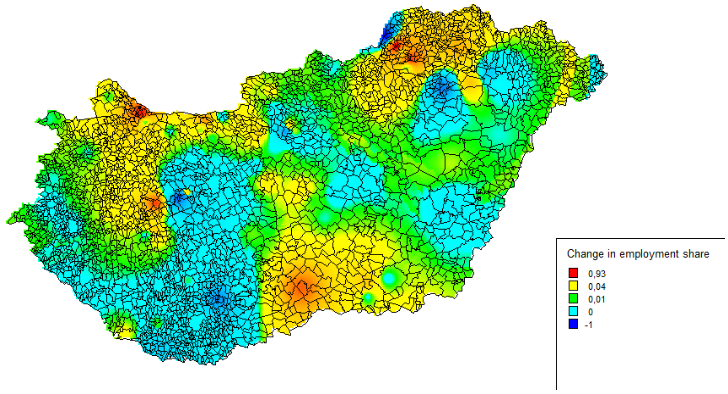


Figure 9: Changes in the share of chemicals in total employment

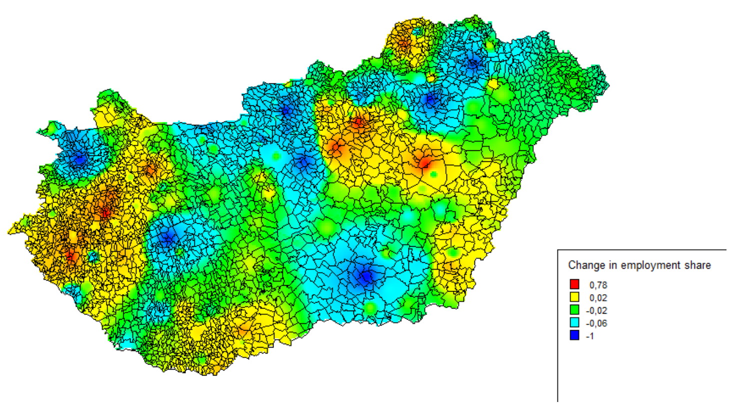


Figure 10: Changes in the share of non-metallic minerals in total employment

forints in 1986), while Figure 11 presents the evolution of employment shares for non-metallic minerals, a much less open industry (in which export per worker was only 49.000 forints). There are a couple of differences between the two distributions but the most striking one is probably that we can observe more differences within both the Western and the Eastern part of the country for non-metallic minerals than for chemicals. Whether, of course, this is due to some extent to differences in the two sectors' openness, is an open question.

The main problem with this approach is, of course, that industries are different not only in their openness but in many other factors. There is no hope to separate the effects of these factors from the effect of trade openness if one exclusively relies on the examination of smoothed maps. In addition, there might have been differences in both the degree of openness and its effect *within* sectors. We can fortunately observe the former kind of heterogeneity as we have data on export per worker in 1986 for each plant in every industry. Concerning the latter, i.e. that also the effect of openness may vary within industries, what we see is that, as documented by Nemes-Nagy (2004) and partially also confirmed in this thesis, spatial reordering during the transition took place along three dimensions: polarization between the Eastern and the Western part of the country, polarization between Budapest and the rest of the country, and increasing urban-rural differentials. Now assume that these changes in spatial structure were exclusively caused by differences in trade openness. This would imply that either (1) urban and rural, Eastern and Western regions as well as Budapest and the rest of the country were initially endowed with different degrees of trade openness, or (2) the effect of trade openness was different across these regions. (1) can be justified by comparing average trade openness in these regions; still, (2) can be true even in this case. Similarly, even if trade openness was not the only determinant of spatial reordering (surely it was not), (2) may hold. Therefore, it is reasonable to assume *ex ante* that the effect of trade openness was different between East and West, between Budapest and the other regions, as well as across areas urbanized to a different extent; we can then learn *ex post*, as one of the results of the analysis, whether these effects differed in reality. Finally, notice that the effect may be higly nonlinear, that is, a region with 5% of its workers employed in agriculture might not have experienced the same increase or decrease due to trade as one with 70% employed in agriculture, even other things being equal; therefore, it would be nice to compare regions in which employment shares of the particular industry were similar prior to the transition.

To sum up, I want to compare regions which were similarly urban or rural, were located in the same part of the country (East, West or Central), and had a similar fraction of employees working in the specific industry. Let us look at a couple of such examples first.

1) Esztergom and Dorog are neighbour subregions in Komárom-Esztergom county. As a consequence, they are both located in Western Hungary, about equally far from Budapest.

Both subregions had 19000 employees in 1986, with 12000 and 9000 workers employed in the seats of the subregions (the towns of Esztergom and Dorog). Hence, the degree to which the two subregions were urbanized seems to be very similar. In addition, the share of workers employed in the chemicals industry in 1986 was 14.2% for Esztergom and 11.4% for Dorog. Besides these similarities, export per worker in the sector was 168.700 forints in Esztergom but almost six times as large (951.300 forints) in Dorog. Until 1994, the share of the chemicals industry dropped to 8.0% in Esztergom but increased to 15.3% in Dorog. That is, the same industry improved in the region in which its plants had more trade relations while it fell back in the other region.

2) The subregions of Dunakeszi and Szentendre are both located in the Northern part of Pest county. Their distance to Budapest is similar. There were about 13000 employees in both in 1986, 4.5% working in the paper industry in Dunakeszi and 3.3% working in the same sector in Szentendre. The paper industry in Szentendre was, however, more open to trade, and especially to trade with the East: export per worker was 104.800 forints in Szentendre whereas only 42.700 in Dunakeszi, the difference being made up almost fully by rubel exports. The employment shares of 1994 were as follows: 3.3% worked in the sector in Dunakeszi and 5.2% in Szentendre. In sum, the subregion with more exports in the specific industry, and more exports to the East in particular, managed to overcome the one in which the specific industry was less open.

3) The third example sheds light on possible "crowding-out" effects across industries. The subregions of Orosháza and Hódmezővásárhely are also next to each other, located in the Southeastern part of the country. Orosháza had 23000 employees in 1986 while Hódmezővásárhely had 28000; out of these, 15.4% worked in the sector using non-metallic minerals in Orosháza and 14.5% in Hódmezővásárhely. Openness to trade in the sector was similar across the two subregions, export per worker being 75.000 in Orosháza and 57.000 in Hódmezővásárhely in 1986. The share of rubel exports was also similar. However, other sectors in Orosháza were definitely more open to trade, overall export per worker being 61.200 in the town as opposed to only 32.600 in Hódmezővásárhely. As a consequence, other sectors might have decreased less than proportionally and the share of employment in non-metallic minerals in Orosháza was below that in Hódmezővásárhely in 1994 (9.7% vs 10.2%). Of course, not only crowding-out effects but also positive spillovers are possible; if a sector starts improving due to it being more open to trade, it can affect other sectors to which it is closely related (think of forestry and the paper industry, for instance).

4) Nevertheless, it is not difficult to find counterexamples to the patterns previously described. For instance, the subregions of Kőszeg and Szentgotthárd are both located in Transdanubia, next to the Austrian border. The number of employees was 5000 in Kőszeg and

6000 in Szentgotthárd in 1986. Both Kőszeg and Szentgotthárd were "textile regions" before the transition, 34.0% of Szentgotthárd employees and an even higher fraction, 52.2%, of Kőszeg employees working in this sector in 1986. However, Szentgotthárd had the advantage of being more open to trade both at the sectoral (export per worker 190.400 forints vs 31.700 forints in Kőszeg) and at the overall economic level (export per worker 82.200 vs 18.500 forints in Kőszeg). (The share of rubel and nonrubel exports was about the same in the two regions.) Nevertheless, textile industry fell back in Szentgotthárd, its share in 1994 being less than third of its pre-transition share. Although Kőszeg also experienced some drop, textile employment took 36.8% of total employment even after the transition.

Of course, whether these examples provide any useful information on the aggregate effect of trade openness is questionable. Other factors which we cannot observe, such as the performance of local governments, may be responsible for these changes in sectoral employment shares. Therefore, I turn to reporting average effects by using a regression setup. This has the advantage that idiosyncratic regional effects can be expected to cancel as a result of taking averages.

It is not a simple task to choose the specification of the regression model and the units of observation due to the following problems. First of all, it would be nice to build on settlement-level data because this would definitely increase the number of observations. However, except if one uses advanced spatial econometric techniques, this does not take into account the fact that spillover and crowding-out effects are certainly at work not only within settlements' borders but also between neighboring settlements. Therefore, I decided to use subregion-level data in order to control, at least partially, for these cross-settlement effects. However, this radically diminishes the number of observations, especially if one would like to do the analysis separately for individual sectors and country parts so that heterogeneities can be observed.[12] Therefore, I finally decided to run pooled regressions, including each and every industry. Heterogeneity across sectors is captured merely by the differentials in sectoral exports (as well as their possible differentials in terms of location and pre-transition share); of course, we have to keep in mind that this might increase the absolute values of estimated coefficients on export per worker variables.

There is yet another issue with pooled regressions. The distribution of the number of sectors in subregions is highly skewed to the left, that is, there are only a few subregions with most of the industries operating in them while the number of regions with just a couple of industries is small. As a consequence, the regression assigns a much larger weight to those

[12] Notice that most of the sectors did not operate in each subregion in 1986 and the regression omits these regions since no export per worker data is available for them, decreasing the number of observations even further.

regions in which the number of industries operating was large, simply due to having more observations for these regions. To handle this problem, what I decided to put on the right-hand side is not regional but aggregate export per worker in each sector; as a result, 22 observations belong to every subregion. This, of course, has the disadvantage that locational differences in openness *within industries* are not taken into account. Notice that this, of course, does not mean that locational differentials in openness are totally taken out; they are included both directly, through average export per worker of industries in the specific subregion, and also indirectly through the composition of industries in the region, some of which are more and some which are less open at the aggregate. However, it is true that excluding a part of these differentials probably results in an attenuation bias in the estimated coefficients on openness indicators.

Based on these considerations, the estimated equations have the following form:

$$\log E_{i,r}^{94} = \beta_0 + \beta_1 \log E_{i,r}^{86} + \beta_2 \log X_i + \beta_3 \log X_r + \beta_4 \log N_r + \beta_5 \log D_r + \beta_6 \mathbf{I}_{i,r} + \varepsilon_{i,r}$$

where $E_{i,r}^{94}$ is the employment share of industry i in subregion r in 1994 (i.e., the number of people employed in the specific sector divided by total employment in the region), X_i is average export per worker in industry i in 1986, while X_r is average export per worker in the subregion in the same year. N_r, total employment in the subregion in 1986 and D_r, the average distance of the subregion's settlements from the closest county seat (i.e., from a "large town") are intended to control for urbanization effects. Vector $\mathbf{I}$ includes all the interaction terms between export variables and urbanization variables in order to capture the possibly different effect of openness across different levels of urbanization. Finally, $E_{i,r}^{86}$ is the 1986 share of industry i in region r, controlling for the effect of different "endowments" of sectors at the starting point of the transition process; an alternative model would consist of regressing the *change* in shares on export and urbanization variables, but notice that this specification is a more general one as we do not ex ante require the β_1 parameter to take any specific value.

To control for heterogeneity and also to see whether the estimated coefficients are different between the Western and the Eastern part of the country, I run these regressions separately for Transdanubia and Eastern Hungary.[13] As a final point, to observe discrepancies between the effect of exports to the West and those to the East, I run the regressions also by including rubel and nonrubel exports separately instead of overall exports (i.e., the X variables).

[13] It is impossible to run the regression for Central Hungary because of the small number of observations.

Table 9

Regression results (with standard errors in parentheses)

Dependent variable is sectoral employment share in the subregion in 1994

	Western Hungary		Eastern Hungary	
Regressor	(1)	(2)	(1)	(2)
$\log E_{i,r}^{86}$	0.713 (0.017)	0.718 (0.017)	0.743 (0.015)	0.742 (0.015)
$\log X_i$	−0.496 (0.911)		0.931 (0.854)	
$\log X_r$	4.457 (3.081)		3.095 (2.387)	
$\log X_i^{rub}$		0.451 (0.821)		−0.747 (0.740)
$\log X_i^{nonrub}$		−1.040 (1.526)		1.884 (1.472)
$\log X_r^{rub}$		2.797 (2.835)		1.848 (2.271)
$\log X_r^{nonrub}$		1.753 (3.140)		0.733 (2.283)
$\log N_r$	0.492 (0.387)	0.415 (0.487)	−0.317 (0.302)	−0.085 (0.354)
$\log D_r$	−1.361 (0.903)	−1.642 (1.110)	−0.796 (0.698)	−0.085 (0.354)
$\log X_i \log N_r$	0.074 (0.035)		0.000 (0.298)	
$\log X_i \log D_r$	−0.005 (0.069)		−0.076 (0.062)	
$\log X_r \log N_r$	0.016 (0.104)		−0.126 (0.079)	
$\log X_r \log D_r$	−0.428 (0.247)		−0.175 (0.193)	
$\log X_i^{rub} \log N_r$		−0.053 (0.033)		−0.008 (0.027)
$\log X_i^{rub} \log D_r$		0.003 (0.062)		0.080 (0.053)
$\log X_r^{rub} \log N_r$		−0.155 (0.111)		−0.082 (0.059)
$\log X_r^{rub} \log D_r$		−0.120 (0.243)		−0.098 (0.197)
$\log X_i^{nonrub} \log N_r$		0.138 (0.056)		0.008 (0.050)
$\log X_i^{nonrub} \log D_r$		−0.007 (0.114)		−0.174 (0.105)
$\log X_r^{nonrub} \log N_r$		0.163 (0.118)		0.044 (0.095)
$\log X_r^{nonrub} \log D_r$		−0.309 (0.282)		−0.110 (0.194)
Constant	8.785 (11.648)	12.777 (14.603)	10.542 (9.164)	10.854 (10.405)
# of observations	1452	1452	1518	1518
R^2	0.559	0.563	0.609	0.611

The results can be seen in Table 9. The estimated coefficients seem to confirm most of our expectations. In the case of the East, both sectoral and regional openness is positively connected to the evolution of sectoral employment share in the region, with a 1% increase in sectoral openness resulting in an almost proportional increase in the employment share on average, while the effect is three times as large for regional openness. The latter result might imply that spillover effects play an important role in the process of shaping industries' spatial patterns through openness to international trade. As for the interaction terms, their coefficients are all very close to 0 but (apart from one of them) seem to be in line in the expectations, i.e., the coefficients on the openness indicators were greater in more urbanized regions. Concerning the decomposition of exports to rubel and non-rubel exports, we again obtained reasonable estimates for the coefficients as regions in which non-rubel exports were greater seem to have improved to a larger extent, whereas the coefficient of rubel exports is negative in the case of sectoral openness but positive for regional openness.

The results of regressions on Western Hungary, are somewhat less clear but many of the estimated coefficients still have the expected sign. Sectoral export per worker seems to be negatively connected to industrial performance but notice that the magnitude of the estimated coefficient is small, and it becomes positive if we look at regional openness. In the regression with decomposed exports, the signs of three estimated coefficients out of four are surprising (both sectoral and regional rubel exports are positively connected to performance, and sectoral non-rubel exports are negatively connected to performance), but the magnitudes are quite small in these cases as well, so it can happen that these effects are equal to 0 in reality. The most important observation regarding the "Western" regressions might be, however, that overall regional openness has a greater coefficient on the 1994 employment share than for the corresponding "Eastern" regression; a 1% increase in export per worker in the subregion is associated with a 4.5% increase in sectoral share. That is, regional openness seems to be dominant in the case of Western regions, whereas sectoral openness might be more important (though still to a lesser extent than regional openness) in Eastern Hungary.

It is important to note that, of course, the results of these regressions should not be taken too seriously. First of all, as the trade liberalization process was to some extent already in progress in 1986, exports in this year measure "pre-transition openness" only with an error. Even more importantly, unobserved variables such as state subsidies to plants, local governments' performance, etc. are very likely to affect the changes in sectoral shares and may also be correlated with the openness of regions. All these factors imply that the OLS estimators of true effects (as well as those of standard errors) are biased and inconsistent, and significance tests are invalid. Hence, one should not look at the estimated coefficients as measures of true effects but merely as measures of average correlation between the variable of interest (e.g., the particular industry's openness to trade) and the 1994 share of the industry, once the values of all the other variables, that is, employment share in 1986, total employment and distance from the county seat, are identical. However, the observed heterogeneity in these average correlations between country parts as well as the fact that the magnitude of these effects seems to be quite large makes the results interesting and definitely calls for further research in the topic.

7 Conclusions

In this thesis, I reported stylized facts on the geographical rearrangement of economic activities in Hungary during the transition period. To this end, I used a linked employer-employee dataset which has two advantages: data were collected both before and after the transition, and it includes the location of employees' workplaces up to the settlement level. What I found, over and above the changes already known in the literature, is twofold. First, I showed that huge intersectoral heterogeneity exists in concentration tendencies across different industries, with the "biggest" industries (in terms of country-level employment share) becoming more concentrated but many sectors becoming more dispersed. At the same time, I found almost all industries moving toward the Western part of the country but with, again, inter-industry differentials regarding both the magnitude of these movements and movements *within* country parts.

Second, I focused on the correlation between (1) regional discrepancies in the changes of sectors' performance (measured as changes in employment shares) and (2) differences in the degree of pre-transition openness to international trade. I found that industries and regions which were more open, especially to trade with Western countries, managed to improve in general, but differentials in openness across sectors seem to play a larger role in the East whereas differentials across regions is the primary factor responsible for the correlation in the West.

All these results might constitute a basis for further investigations. Disentangling the effects of openness and those caused by other factors seems to be an especially challenging task. This would, of course, require more sophisticated methods so that the effects can be clearly separated from one another. Another way to proceed is building up theoretical models which could suggest answers to questions raised by this paper, such as what can be the reason for the observed heterogeneities or what made sectors and regions reacting in a different way to openness to trade.

8 References

Adam, J., 1999. *Social costs of transformation to a market economy in post-socialist countries. The cases of Poland, the Czech Republic and Hungary.* Macmillan.

Barrios, S., Bertinelli, L., Strobl, E., Teixeira, A., 2005. The dynamics of agglomeration: evidence from Ireland and Portugal. *Journal of Urban Economics* 57, 170–88.

Barta, Gy., 2003. A magyar ipar területi folyamatai. In *Közelkép. Munkaerőpiaci egyenlőtlenségek és földrajzi mobilitás Magyarországon* (ed. Cseres-Gergely, Zs.), pp. 144–58.

Behrens, K., Gaigné, C., Ottaviano, G., Thisse, J., 2007. Countries, regions and trade: On the welfare impacts of economic integration. *European Economic Review* 51, 1277–1301.

Brülhart, M., Traeger, R., 2005. An account of geographic concentration patterns in Europe. *Regional Science and Urban Economics* 35, 597–624.

Combes, P., Mayer, T., Thisse, J., 2008. *Economic geography. The integration of regions and nations.* Princeton University Press.

Crozet, M., Koenig Soubeyran, P., 2004. EU enlargement and the internal geography of countries. *Journal of Comparative Economics* 32, 265–79.

Dumais, G., Ellison, G., Glaeser, E., 2002. Geographic concentration as a dynamic process. *Review of Economics and Statistics* 84, 193–204.

Duranton, G., Overman, H., 2005. Testing for localization using micro-geographic data. *Review of Economic Studies* 72, 1077–106.

Ellison, G., Glaeser, E., 1997. Geographic concentration in U.S. manifacturing industries: a dartboard approach. *Journal of Political Economy* 105, 889–927.

Feser, E., 2000. On the Ellison–Glaeser geographic concentration index. *http://www.urban.illinois.edu/f*

Hare, P., 2001a. Hungary: In transition to a market economy. *The Journal of Economic Perspectives* 5, 195–201.

Hare, P., 2001b. Post-communist economic transformation: Hungary vs Poland. *AEA Papers and Proceedings* 81, 329–333.

Hungarian Central Statistical Office, 2009. *Gazetteer of the Republic of Hungary. http://portal.ksh.hu/pl*

Flyer on the Harmonized Hungarian Wage Survey, 2006.

Kertesi, G., Köllő, J., 1997. Reálbérek és kereseti egyenlőtlenségek, 1986–1996. A bérszerkezet alakulása Magyarországon, I. rész. *Közgazdasági Szemle* 44, 612–34.

Krugman, P., 1991. Increasing returns and economic geography. *Journal of Political Economy* 99, 483–99.

Krugman, P., Livas Elizondo, R., 1996. Trade policy and the Third World metropolis. *Journal of Development Economics* 49, 137–50.

Major, K., Nemes-Nagy, J., 1999. Területi jövedelemegyenlőtlenségek a kilencvenes években. *Statisztikai Szemle* vol. 5/1999, 397–421.

Montfort, P., Nicolini, R., 2000. Regional convergence and international integration. *Journal of Urban Economics* 48, 286–306.

Nemes-Nagy, J., 2004. Elements of regional disparities in the new regional pattern. In *The Region. Regional development, policy, administration and e-government* (ed. Enyed, Gy. and Tózsa, I.). Akadémiai Publ., pp. 62–79.

Nemes-Nagy, J., Németh, N., 2003. Az átmeneti és az új térszerkezet tagoló tényezői. *http://econ.core.hu/doc/KTI_ books/kti6_ 4.pdf*

Paluzie, E., 2001. Trade policies and regional inequalities. *Papers in Regional Science* 80, 67–85.

Quadrado, L., Heijman, W., Folmer, H., 2001. Multidimensional analysis of regional inequality: The case of Hungary. *Social Indicators Research* 56, 21–42.

Sanguinetti, P., Volpe Martincus, C., 2009. Tariffs and manufacturing location in Argentina. *Regional Science and Urban Economics* 39; 155–167.

9 Appendix

9.1 Detailed information on employers and employees included in the Harmonized Hungarian Wage Survey

Table A1

Year	Budget institutions	Companies
1986	all	all with over 20 employees
1989	all	all with over 20 employees
1992	all	all with over 20 employees
1993	no data available	all with over 20 employees
1994	all	all with over 20 employees
1995	all	all with over 20 employees, random sample of those with 11 to 20 employees
1996	all	all with over 20 employees, random sample of those with 11 to 20 employees

Year	Budget inst employees	Company employees
1986	all	complicated selection, different for different skill groups – see HWS (2006)
1989	all	complicated selection, different for different skill groups – see HWS (2006)
1992	all	those born on 5th, 15th or 25th of any month
1993	NA	those born on 5th, 15th or 25th of any month
1994	all	those born on 5th, 15th or 25th of any month
1995	those born on 5th, 15th or 25th of any month	all for companies with <20 employees, those born on 5th, 15th or 25th for others
1996	those born on 5th, 15th or 25th of any month	all for companies with <20 employees, those born on 5th, 15th or 25th for others

9.2 List of the 150 subregions of Hungary (as of 1997)

Table A2

Subregion	County	Subregion	County	Subregion	County
Budapest	Budapest	Szikszó	BAZ*	Pétervására	Heves
Komló	Baranya	Tiszaújváros	BAZ*	Dorog	KE***
Mohács	Baranya	Csongrád	Csongrád	Esztergom	KE***
Sásd	Baranya	Hódmezővásárhely	Csongrád	Kisbér	KE***
Sellye	Baranya	Kistelek	Csongrád	Komárom	KE***
Siklós	Baranya	Makó	Csongrád	Oroszlány	KE***
Szigetvár	Baranya	Mórahalom	Csongrád	Tata	KE***
Pécs	Baranya	Szeged	Csongrád	Tatabánya	KE***
Pécsvárad	Baranya	Szentes	Csongrád	Balassagyarmat	Nógrád
Baja	Bács-Kiskun	Bicske	Fejér	Bátonyterenye	Nógrád
Bácsalmás	Bács-Kiskun	Dunaújváros	Fejér	Pásztó	Nógrád
Kalocsa	Bács-Kiskun	Enying	Fejér	Rétság	Nógrád
Kecskemét	Bács-Kiskun	Gárdony	Fejér	Salgótarján	Nógrád
Kiskőrös	Bács-Kiskun	Mór	Fejér	Szécsény	Nógrád
Kiskunfélegyháza	Bács-Kiskun	Sárbogárd	Fejér	Aszód	Pest
Kiskunhalas	Bács-Kiskun	Székesfehérvár	Fejér	Cegléd	Pest
Kiskunmajsa	Bács-Kiskun	Csorna	GyMS**	Dabas	Pest
Kunszentmiklós	Bács-Kiskun	Győr	GyMS**	Gödöllő	Pest
Jánoshalma	Bács-Kiskun	Kapuvár	GyMS**	Monor	Pest
Békéscsaba	Békés	Mosonmagyaróvár	GyMS**	Nagykáta	Pest
Mezőkovácsháza	Békés	Sopron	GyMS**	Ráckeve	Pest
Orosháza	Békés	Tét	GyMS**	Szob	Pest
Sarkad	Békés	Balmazújváros	Hajdú-Bihar	Vác	Pest
Szarvas	Békés	Berettyóújfalu	Hajdú-Bihar	Budaörs	Pest
Szeghalom	Békés	Debrecen	Hajdú-Bihar	Dunakeszi	Pest
Miskolc	BAZ*	Hajdúböszörmény	Hajdú-Bihar	Gyál	Pest
Edelény	BAZ*	Hajdúszoboszló	Hajdú-Bihar	Piliscsaba	Pest
Encs	BAZ*	Polgár	Hajdú-Bihar	Szentendre	Pest
Kazincbarcika	BAZ*	Püspökladány	Hajdú-Bihar	Barcs	Somogy
Mezőkövesd	BAZ*	Eger	Heves	Csurgó	Somogy
Ózd	BAZ*	Heves	Heves	Fonyód	Somogy
Sárospatak	BAZ*	Füzesabony	Heves	Kaposvár	Somogy
Sátoraljaújhely	BAZ*	Gyöngyös	Heves	Lengyeltóti	Somogy
Szerencs	BAZ*	Hatvan	Heves	Marcali	Somogy

Subregion	County	Subregion	County	Subregion	County
Nagyatád	Somogy	Szolnok	JNSz*****	Vasvár	Vas
Siófok	Somogy	Tiszafüred	JNSz*****	Ajka	Veszprém
Tab	Somogy	Törökszentmiklós	JNSz*****	Balatonalmádi	Veszprém
Baktalórántháza	SzSzB****	Bonyhád	Tolna	Balatonfüred	Veszprém
Csenger	SzSzB****	Dombóvár	Tolna	Pápa	Veszprém
Fehérgyarmat	SzSzB****	Paks	Tolna	Sümeg	Veszprém
Kisvárda	SzSzB****	Szekszárd	Tolna	Tapolca	Veszprém
Mátészalka	SzSzB****	Tamási	Tolna	Várpalota	Veszprém
Nagykálló	SzSzB****	Celldömölk	Vas	Veszprém	Veszprém
Nyírbátor	SzSzB****	Csepreg	Vas	Zirc	Veszprém
Nyíregyháza	SzSzB****	Körmend	Vas	Keszthely	Zala
Tiszavasvár	SzSzB****	Kőszeg	Vas	Lenti	Zala
Vásárosnamény	SzSzB****	Őriszentpéter	Vas	Letenye	Zala
Jászberény	JNSz*****	Sárvár	Vas	Nagykanizsa	Zala
Karcag	JNSz*****	Szentgotthárd	Vas	Zalaegerszeg	Zala
Kunszentmárton	JNSz*****	Szombathely	Vas	Zalaszentgrót	Zala

(Abbreviations stand for: * Borsod-Abaúj-Zemplén; ** Győr-Moson-Sopron; *** Komárom-Esztergom; **** Szabolcs-Szatmár-Bereg; ***** Jász-Nagykun-Szolnok.)

9.3 Rankings of counties with respect to net revenue and profit per worker

Table A3

Rankings of counties with respect to net revenue per worker

County	Net revenue per worker in 1986, thousand HUF	County	Net revenue per worker in 1994, thousand HUF
Heves	949	Fejér	2642
Pest	936	Budapest	2519
Budapest	888	Győr-Moson-Sopron	2517
Komárom-Esztergom	795	Vas	2045
Csongrád	763	Komárom-Esztergom	1946
Fejér	754	Tolna	1915
Borsod-Abaúj-Zemplén	728	Zala	1657
Békés	706	Jász-Nagykun-Szolnok	1645
Győr-Moson-Sopron	640	Veszprém	1544
Bács-Kiskun	629	Hajdú-Bihar	1482
Szabolcs-Szatmár-Bereg	608	Borsod-Abaúj-Zemplén	1481
Tolna	599	Pest	1478
Hajdú-Bihar	597	Heves	1425
Jász-Nagykun-Szolnok	594	Bács-Kiskun	1406
Somogy	591	Békés	1390
Veszprém	575	Somogy	1381
Vas	568	Csongrád	1350
Baranya	529	Baranya	1281
Zala	501	Szabolcs-Szatmár-Bereg	1260
Nógrád	499	Nógrád	959

Table A4

Rankings of counties with respect to profit (before tax) per worker

County	Profit per worker in 1986, thousand HUF	County	Profit per worker in 1994, thousand HUF
Budapest	49	Vas	138
Komárom-Esztergom	46	Győr-Moson-Sopron	76
Pest	40	Veszprém	50
Győr-Moson-Sopron	38	Tolna	46
Heves	38	Budapest	42
Vas	35	Jász-Nagykun-Szolnok	41
Fejér	35	Komárom-Esztergom	38
Békés	34	Szabolcs-Szatmár-Bereg	35
Tolna	34	Zala	35
Csongrád	34	Heves	34
Borsod-Abaúj-Zemplén	34	Somogy	33
Bács-Kiskun	33	Baranya	31
Hajdú-Bihar	32	Csongrád	30
Jász-Nagykun-Szolnok	32	Fejér	27
Veszprém	32	Bács-Kiskun	25
Zala	31	Pest	24
Nógrád	30	Hajdú-Bihar	24
Baranya	28	Borsod-Abaúj-Zemplén	13
Szabolcs-Szatmár-Bereg	28	Békés	8
Somogy	28	Nógrád	−4