Author(s): Tervo, Hannu

Title: Cities, hinterlands and agglomeration shadows: spatial developments in Finland over 1880-2004

Year: 2010

Version:

Please cite the original version:

Tervo, H. (2010). Cities, hinterlands and agglomeration shadows: spatial developments in Finland over 1880-2004. In ERSA European Regional Science Association 19-23 August 2010 Jönköping Sweden: papers.

Cities, hinterlands and agglomeration shadows: spatial developments in Finland during 1880-2004*

Hannu Tervo
Professor in Economics
School of Business and Economics
P.O. Box 35
FIN 40014-University of Jyväskylä, Finland
hannu.t.tervo@jyu.fi

Abstract. This paper analyzes long-term spatial developments in Finland by focusing on two predictions of the new economic geography (NEG) models: the increasing persistence of locational patterns and the rising dominance of growth centers. The empirical analysis is based on regional population data from 1880 to 2004. The results support the hypotheses. Evolutions in rank and rank-size distributions during the processes of industrialization and urbanization suggest increasing persistence of regional structures. The analysis of causal processes between population centers and their hinterlands shows that these regions grew hand-in-hand in the pre-war period, whereas agglomeration shadows started to come about during the post-war period.

Keywords: Regional development; Urbanization; Growth centers; Peripheral regions; New economic geography; Rank-size rule; Granger causality; Finland

* **Acknowledgements** – I'd like to thank Pasi Mäkelä and Mika Haapanen for their help in the processing and analysis of the data and Ilkka Nummela for discussions on the topic. This research has been undertaken with the support of the Academy of Finland (research project 127049).

1. INTRODUCTION

Like many other developed countries, Finland has undergone a significant transition from an agriculture-dominated country to a technology-based information society. Agriculture and forestry were long dominant in Finland before the rapid growth of manufacturing and service industries after WWII. Regional development in the form of agglomeration and urbanization closely followed this industrialization. Each Finnish province (NUTS 3 region) developed a center of its own, which is now in a leading position in its region. Casual observation shows that these growing cities tend to dominate in their regions. New economic geography (NEG) models formalize the role of agglomeration in the dynamic formation of an urban system. NEG predicts that spatial development will become centralized and locational patterns persistent with time (Krugman, 1991). This development is due to self-reinforcing forces, which, once initiated, will also develop rapidly. The models also predict that a city always casts a so-called "agglomeration shadow" on its local hinterland (Fujita, Krugman and Venables, 1999; Dobkins and Ioannides, 2001; Partridge et al., 2009). Persistence and the role of growth centers in their regions depend on the stage of development of the economy. In the early phase, when the industrialization process is about to begin, regional structures are not yet persistent. Growth centers are only evolving, and agglomeration shadows are weak or nonexistent. Developing centers and hinterlands grow hand-in-hand as the population in the hinterlands expands to provide the urban center with agricultural products (Fujita et al., 1999, Partridge et al., 2009). In the course of industrialization and urbanization, however, persistence in locational structures increases and agglomeration shadows strengthen.

This paper analyzes long-term spatial developments in Finland by focusing on two predictions of the NEG models: the increasing persistence of locational patterns and the emerging agglomeration shadow, i.e., the rising dominance of growth centers.[1] The

---

[1] NEG models have given rise to several empirical hypotheses, including the home-market effect, i.e., how regions with a large demand for industries with increasing returns have a more than proportional share of

empirical analysis is based on regional population data from 1880 to 2004 at decade intervals. Population data offer the only reasonably good possibility for analyzing regional development over the long term. In the empirical analysis, persistence and dominance in Finnish regions are examined in the different phases of the industrialization process. We distinguish between the pre- and post-war periods to roughly capture the shift from an agriculture-based economy to a post-industrial country. WWII can be regarded as a landmark of this change in Finland, although the exact timing of the transition is difficult to specify. Roughly, in the language of Krugman's (1991) model, pre-war Finland was characterized by production dominated by immobile farmers and high transport costs, whereas the post-war economy is characterized by declining transport costs and the growing importance of footloose industrial production with increasing returns to scale. For simplicity, the analyses are based on the assumption that each of the 19 Finnish regions has a center of its own and the rest of the region forms its local hinterland. First, to analyze the persistence of locational patterns, we examine the variation in the ranks of regions over time and the evolution in rank-size distributions at different stages of development. Second, to analyze the dominance of centers and causal processes between cities and their local hinterlands before and after WWII, an extension of the Granger causality method using a panel framework is applied.

A substantial line of work has assessed the persistence of the distribution of economic activity, but empirical work related to the role of growth centers is scarce (see, however, Partridge et al., 2008 and 2009). Dobkins and Ioannides (2001), Overman and Ioannides (2001) and Black and Henderson (2003) all present analyses of urban development in the U.S over the 1900-1990 period. Davis and Weinstein (2002) analyzed the distribution of economic activity in Japan under assumptions of increasing returns, random growth and location fundamental theory and reached the conclusion that location fundamentals establish the spatial patterns of relative regional densities, but increasing returns may help

---

their production; raising local factor prices with regions with a large market potential; increasing factor flows with regions with a large market potential; raising agglomeration with a further reduction in trade costs at some critical level of transport or trade costs; and shock sensitivity, i.e., how changes in the economic environment can trigger permanent changes in the equilibrium distribution of economic activity (Head and Mayer, 2004; Brakman and Garretsen, 2006). Our hypotheses are related to the last two NEG features.

to determine the degree of spatial differentiation. Brakman, Garretsen and Schramm (2004) showed that this conclusion also holds to some degree for western German cities. They also found that city growth in western Germany did not follow a random walk, whereas in eastern Germany it did. Based on the same methodology, Bosker et al. (2007) explicitly tested for multiple equilibria in western German cities and found that multiple equilibriums are present in the evolution of the distribution of city size. Earlier analyses focusing on long-term regional development in Finland are in short supply, especially from the perspective of spatial economics. Ottaviano and Pinelli (2004; 2006) shed some light on the forces that have shaped the economic landscape of Finland by assessing the theoretical predictions of NEG models. Tervo (2009) analyzed the mutual relationships between growth processes in centers and their surrounding hinterlands from 1970 to 2004. Both of these analyses involved only the most recent few decades.

The paper is organized as follows. Section 2 presents the theoretical starting point for the empirical testing. Section 3 provides background information on long-term economic and policy development in Finland's regions. Section 4 describes the data and implementation of the study. The stability of Finland's regional structure is analyzed in Section 5, first by analyzing the development of the relative positions of sub-regions and then by estimating the rank-size equation for different years. The role of growth centers in regional development is analyzed in Section 6. This section introduces the method based on panel Granger causality tests and presents the results. Section 7 concludes.

2. THEORETICAL CONSIDERATIONS

New economic geography (NEG) models predict that spatial development is centralized due to self-reinforcing forces, which, once started, will also develop rapidly (Fujita, Krugman and Venables, 1999). A process of cumulative causation predominates, and the result is an economy with a few strong centers. NEG models formalize this cumulative causation mechanism to explain the spatial concentration of economic activities and endogenous regional differentiation of centers and peripheral regions (Ottaviano and Puga, 1998). In Krugman's (1991) original core-periphery model, there are two kinds of production, namely agriculture and manufacturing. The model has three potential stable equilibria: perfect spreading or full agglomeration in either region. According to the theory, agglomeration is linked with structural change from primary production to manufacturing and service industries. The circularity that can generate manufacturing concentration will not matter significantly if manufacturing employs only a small fraction of the population and hence generates only a small fraction of demand or if a combination of weak economies of scale and high transportation costs induces suppliers of goods and services to the agricultural sector to locate very close to their markets (Krugman, 1991, p.486). However, once these criteria are no longer met, the population will start to concentrate and regions to diverge. The circular process feeds itself, with the result being an economy with a distinct core-periphery structure. However, because of competing forces, there is a chance for changes in regional development.

The original two-region NEG model can be extended to explain how an economy can evolve from monocentrism to a multiple-city geography and even to a central-place hierarchy (Fujita et al, 1999). The possibility of partial agglomeration certainly must be considered as this situation is what predominates in the real world. The same factors that work toward the concentration of economic activity in the two-region model tend to produce fewer, larger concentrations in a continuous-space model. The evolution of concentration developed by Fujita et al. (1999) reflects the same logic as the two-region model, including the tension between the centripetal forces created by backward and forward linkages and the centrifugal force created by immobile land. In an economy with several manufacturing sectors, differences among industries in scale economies and/or transport costs can produce a hierarchical urban system.

In this paper, the stability of regional structure and the role of centers over the course of industrial change are analyzed in one country, Finland. NEG models imply that (1) regional structure is only emerging in the pre-industrialization phase, whereas it should become more persistent in the course of the industrialization process; and that (2) evolving centers support the growth of their local hinterlands, but once the economy has reached its break point and development has locked into the fortunate cities, a majority of firms and people move to these cities, causing an agglomeration shadow in their local hinterlands. The earlier discussion in regional science about the role of growth centers differentiates “backwash” or “polarization” effects from “spread” or “trickling down” effects (Myrdal, 1957; Richardson, 1978; Parr, 1999a and 1999b). Backwash effects occur as resources, especially labor, gravitate toward the center, although a growth center may also have spread effects on its hinterland due to the relocation of manufacturing plants, the decentralization of the population and the spread of innovation, investment and growth attitudes, especially in the longer term (Richardson, 1978).

3. A BRIEF GEOGRAPHICAL HISTORY OF FINLAND

Finland is a country with a small population but a large area. Many regional problems are related to the country’s large size and scattered population. The economy and society of Finland were long dominated by primary production. Family farms, including areas of forest, were typical across the entire country. Large-scale resettlement after WWII had the effect of prolonging this dominance of rural economic and social activity (Virrankoski, 1975).

The advantages that “first nature” provides were decisive in the birth of the first towns, whereas the self-reinforcing advantages of “second nature” provided the incentive for still more population and production (Krugman, 1993). The first towns lay along the south and west coasts, whereas inland towns began to form with the growth of forest industries. Sawmills and pulp and paper mills were built along rivers, mainly at their estuaries, and this development led to the establishment of new towns in many parts of Finland. For 50 years leading up to WWI, the portion of the paper industry that was based on local reserves of wood was the quickest-growing branch, and the regional spread of industries was fairly even across the country. Railway construction began in the mid-19th century, and new towns sprouted at railway junctions early in the 20th century. The interwar years (1920-38) were a time of early industrial growth, first in wood-based industries but later also in metal and engineering. The growth of the metal industry accelerated very rapidly after WWII due to massive reparations needed across the country, mainly heavy machinery and ships. Internal migration was fairly modest at the end of 19th century. The significance of interregional migration increased in the early 1900s, but as late as the 1920s, 90% of the population lived in their home province and about 70% in their home municipality (Pitkänen, 1994). In contrast, emigration was commonplace at the turn of the 20th century. Between the 1880s and 1930s, more than 400,000 people departed for distant countries, especially the U.S. and Canada.

As mentioned earlier, the processes of industrialization and urbanization began late in Finland, but they occurred rapidly; these processes were fastest in Europe in the 1960s and 1970s. Post-war economic development was rapid, and welfare gaps between much more developed economies and Finland narrowed, even partially reversing. Rapid economic expansion together with structural change had the effect of centralizing both economic activity and population. Migration streams from the countryside to towns started to grow in the 1950s. Migration trended toward the southern and central regions of the country, where the metropolitan area of Helsinki and most of the other larger towns and urban centers are located. The late 1960s and early 1970s were characterized by the depopulation of rural Finland. People moved from rural areas to cities and from the north to the south. A considerable part of this migration was directed toward Sweden. After this flow peaked at the beginning of the 1970s, regional agglomeration leveled out, and during the late 1970s and 1980s, migration slowed down. Finland was hit by a severe recession in 1991-93, and both production and employment fell sharply. The 1990s were a time of great economic flux and drastic structural change. Migration was directed from rural and smaller urban areas toward a few big centers for which growth was based on the export and high technology industries. Only in the 2000s has this new growth spread to a few other large towns.

Systematic regional policy became relevant in Finland as late as the mid-1960s. Before the rapid processes of urbanization and industrialization, regional problems were not salient. Only fragmented ideas, plans and projects for regional development were brought forward before the initiation of systematic regional policy. In the 1950s, many development projects focused on northern Finland. Four important policy orientations can be distinguished in Finnish regional policy (Tervo, 2005). First, the infrastructure policies of the 1950s and 1960s laid the foundations for the development of the sparsely populated parts of Finland. Second, early regional policy aimed to decentralize manufacturing industries and disperse them across the targeted development areas. The third successful form of regional policy has been welfare policy, specifically with regard to its spatial components. The fourth successful form of regional policy has been the regionalization of university education, which practically covers the entire country.

Despite the strong trends toward regional concentration, welfare differences between regional economies are not very large in Finland and have tended to converge in the long run. Indeed, many Finnish studies show a convergence in per capita differences. Kangasharju (1998) approximated regional income levels by taxed income per capita and found clear convergence within the 88 Finnish sub-regions between 1934 and 1993. The b convergence in Finland has been slightly above two percent annually and thus rather similar to that in many other countries. Government intervention has certainly influenced this, but as is typical, other factors have had a larger impact on regional development than those relating to regional policy alone (Tervo, 2005). Regional policies have not counteracted market forces that have increasingly steered development toward centralized structures.

4. DATA AND REGIONAL CLASSIFICATIONS

The data are municipality-based population data produced by Statistics Finland that have been reworked to match regional breakdowns in 2005. The data are from the period of 1880-2004 in decade intervals, except the last interval, which is four years long. The data from the 1880-1940 period are based on church registers, whereas from 1950 onward, data are based on population censuses. Analyses are carried out by separating the periods 1880-1940 and 1950-2004. The former describes the time before rapid industrialization and the latter the time after the break point in economic development, which is characterized by a very rapid structural change.

The empirical analyses are based on the current regional breakdown, i.e., NUTS level 3 regions and NUTS level 4 sub-regions in 2005[2], because the aim is to analyze regional development using present-day regions and regional breakdowns as have been shaped in the course of time. Black and Henderson (2003) also used contemporaneous definitions of urban population in their analysis of urban evolution in the USA. NUTS 3 level regions consist of provinces or counties, whereas NUTS 4 level sub-regions are similar to functional areas. Consequently, NUTS 4 level sub-regions represent local labor market areas reasonably well. It is true that municipalities might have been closer to functional areas than the present-day sub-regions at the outset of the study period, but since then, functional areas have become substantially larger.

There were 20 regions and 77 sub-regions in Finland in 2005. Sub-regions consist of two to thirteen municipalities, and regions consist of two to seven sub-regions. One small island region, Åland, which consists of three sub-regions, is omitted due to its special character. Åland has political autonomy and differs from other regions in continental Finland in many ways. As a consequence, the number of regions used in the study is 19, and the number of sub-regions is 74.

The formation of both regions and sub-regions, then, is based on the municipal division in 2005. The number of municipalities in 2005 was 432, whereas it was previously in excess of 600. Some municipalities were reorganized to form larger entities, and new municipalities were formed. All of these changes have been taken into account in the formation of the data. In cases in which an entire municipality was consolidated with another one, this process was simple. It was more problematic, however, if a municipality had been divided among two or more other municipalities. In these cases, the population of the municipality in the period before its division was added to the municipality that received the majority of the population. In practice, the problems following from these

---

[2] The Nomenclature of Territorial Units for Statistics (NUTS) was established by Eurostat more than 30 years ago to provide a single uniform breakdown of territorial units for the generation of regional statistics for the European Union.

issues are small because in most cases, such municipalities were absorbed into the same sub-region, which is the lowest level of regional unit used here.

After WWII, Finland ceded one tenth of its territory to the Soviet Union, and 11% of the population was evacuated and settled into the remaining area of Finland. Because the analysis is based on the present-day regional breakdown, the ceded municipalities do not appear in the data; thus, this is not problematic. However, it is clear that the settlement of the evacuees caused a sudden increase in the populations of many municipalities. By separating the two periods in the analyses, this problem is diminished.

In the analysis of the role of centers in their regions (Section 6), a distinction is made between centers and peripheral sub-regions. The center of a region is defined as the sub-region with the largest population in 2004. In all cases, this region is also the administrative center of the region. All other sub-regions in the region are defined as the periphery. In most cases, the regional centers are obvious; the center is a regional capital surrounded by a large local labor market, whereas the other sub-regions are much smaller and can be easily regarded as peripheral. However, some regions may have several fairly large sub-regions in addition to the center, in which cases the categorization is not that clear.[3] Figure 1 shows the nineteen regions and their centers. Table 1 shows some descriptive statistics.

5. DEVELOPMENT OF REGIONAL STRUCTURES

To analyze the stability of regional structures, we first examine the development of the relative positions, or ranks, of sub-regions over time in the two periods. The relevant questions are whether there is much variation in rankings and especially whether this variation has diminished in the post-war period compared to the pre-war period as NEG

---

[3] Accordingly, the center is defined as the largest sub-region in 2004; however, has this same sub-region always been the largest one? History shows that in 15 cases out of 19, the center has remained the largest sub-region throughout the period. The exceptions are Kymenlaakso, South Karelia, Northern Savo and Lapland. Apart from being the largest sub-regions in 2004, all four sub-regions are the administrative centers of their regions, for which reason they also deserve the status of center.

Figure 1. The nineteen Finnish regions and their centers

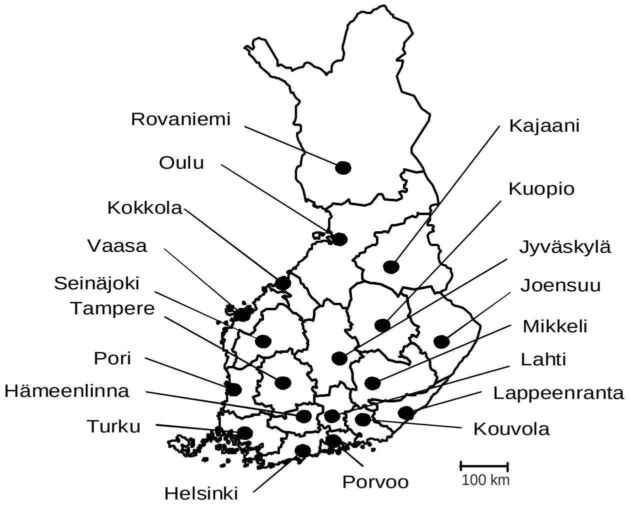


Table 1. Descriptive statistics for the nineteen NUTS 3 regions

Year	**Population in the centers**		**Population in the peripheries**	
	Mean	**S.E.**	**Mean**	**S.E.**
1880	37,975	4,421	58,169	8,003
1890	44,533	5,437	66,537	9,165
1900	52,292	7,201	73,750	9,922
1910	58,183	9,408	76,842	10,022
1920	63,157	11,311	80,020	10,502
1930	70,716	13,929	84,809	10,771
1940	83,134	19,415	89,014	11,079
1950	104,384	23,184	106,159	13,104
1960	124,180	31,343	108,740	13,079
1970	140,019	39,755	100,912	11,787
1980	152,465	45,100	98,324	11,247
1990	163,573	50,340	98,210	11,440
2000	178,009	58,345	93,324	11,034
2004	182,834	60,450	91,381	10,872

predicts. To test these questions, we calculated a simple statistic to measure changes in rankings in both periods. For each successive date, the ranks of each sub-region are compared and the absolute values of the deviations are summed. The statistic for describing changes in regional structure is the average of these sub-period sums.

For the pre-war period, this statistic has a value of 2.26 and for the post-war period a smaller value of 1.70. Thus, on an average, sub-regions encountered somewhat larger changes in rankings before than after the war. We can test the significance of the difference by using a paired samples *t* test, which is appropriate whenever two related sample means are to be compared. The *t* statistic obtained is 3.43, which, with 73 degrees of freedom, signifies a clearly significant result (p=0.001): the hypothesis of no difference between the measurements can be rejected. Thus, there is a statistically significant difference between the fluctuations of the pre- and post-war regional structures. Persistence in locational patterns increased in Finland during the processes of industrialization and urbanization.

Another interesting possibility in analyzing changes in the structure of regional size is to apply the rank-size rule, or Zipf's law, to the data at different points in time. George Zipf (1949) devised the rank-size rule to explain the sizes of cities in a country. According to Zipf's law, the total population of lower ranked cities should be a fixed proportion of the population of the largest city. For example, if the largest city in a country contained one million citizens, the second city would contain one half as many as the first, or 500,000. The third would contain one third, the fourth would be home to one quarter and so on, with the rank of the city representing the denominator of the fraction. This regularity has been found to be valid in many countries, although contrary results have been obtained for other countries (Krugman, 1996; Brakman, Garretsen and van Marrewijk, 2009).

Random growth theory (Simon, 1955; Gabaix, 1999) and locational fundamentals theory (Krugman, 1996) are able to predict the result provided by the rank-size rule, thus providing a foundation for understanding it. Random growth theory holds that a distribution of cities of quite different sizes emerges from very simple stochastic processes. Locational fundamentals theory predicts that if economic characteristics of location are randomly distributed according to the same process as in Gabaix (1999), the rank-size rule can also be accounted for. Unfortunately, these theories lack economic content. The urban hierarchy NEG model developed in Fujita et al. (1999) has no inherent tendency to simulate data in a way that resembles the rank-size rule. Instead, a NEG-based model developed by Brakman et al. (2009) that adds an additional spreading force, congestion, to the core model can replicate changes in actual rank-size distributions.

How, then, does Finland fit into this theory? As an example, consider the year 1990, at which time Helsinki, the largest sub-region, had a population of 1,030,200. The tenth largest sub-region was Kouvola, with a population of 102,200, which is roughly one tenth of the population of Helsinki. The sub-region of Porvoo is the 20th largest sub-region with a population of 64,800, whereas the rank-size rule would predict it to have a population of 51,500. The 50th largest sub-region is Pohjois-Satakunta, with a population of 28,200, whereas the rank-size prediction is 20,600. The largest deviations between the actual population data and the rank-size predictions are between the largest sub-region, Helsinki, and the next largest sub-regions. In particular, the second largest sub-region, Tampere, has a much smaller population than Zipf's law would predict. Similar findings have also been obtained in several other countries; the largest city is too large according to the rank-size rule.

Apart from these discrepancies, this example suggests that Zipf's law has some relevance in Finland. To analyze the relevance more consistently, let us take the following equation as a starting point:

$$(1) \qquad size * rank^{\alpha} = constant,$$

where α is a parameter to be estimated. The parameter α is also called the Pareto exponent. If α is statistically equal to one, then Zipf's law holds. In this case, the constant is equal to the largest sub-region. In general, the Pareto exponent is a measure of how evenly distributed the population is. The smaller the value of the exponent is, the more even in size are the sub-regions, whereas values larger than one imply more urban agglomeration. By employing logarithms and moving the size variable to the right side of the equation, we obtain an equation that can be estimated using the OLS method[4]:

$$(2) \qquad \log(size) = \log(constant) - \alpha \log(rank)$$

Unfortunately, this procedure is strongly biased in small samples (Gabaix and Ioannides, 2004; Gabaix and Ibragimov, 2007). An alternative procedure to the standard OLS estimator is the Hill estimator, but its properties in finite samples can be equally worrisome (Gabaix and Ioannides, 2004). Gabaix and Ibragimov (2007) provided a simple practical remedy for the bias in OLS estimation and proposed the use of (*rank* -1/2) in the estimation of the following equation using OLS:

$$(3) \qquad \log(size) = \log(constant) - \alpha \log(rank-1/2)$$

The shift of 1/2 reduces the bias to a leading order. The standard error is not the OLS standard error but is asymptotically $(2/n)^{1/2}\hat{a}$. Numerical results from Gabaix and Ibragimov (2007) demonstrated the advantage of the proposed approach over the standard OLS estimation procedures. Therefore, we also apply this procedure.

We estimated equation (3) using our sub-region data across fourteen different years during the study period. Two questions are examined: whether the rank-size distribution holds, i.e., what is the goodness of fit, as measured by R^2; and, if it holds, whether Zipf's law holds. It should be noted that due to the definitions used, the ceded regions are not included in the data. Naturally, the pre-war results are to some extent skewed by the absence of these regions, but this absence should not influence the results greatly.

The results show that the rank-size distribution provides a better characterization of sub-region size distributions after the war versus before the war in Finland (Table 2 and Figure 2). The goodness of fit of the rank-size distribution has risen quite evenly. For the year 1880, R^2 is 0.69, and it sinks as low as 0.65 in 1900, but it rises to 0.95 in 2004. In the pre-war period, the average R^2 is 0.78, whereas this value is 0.93 for the post-war

---

[4] Many empirical studies use *log(rank) = log(c) – a´log(size)*, but we prefer specification (2), which was also used by Zipf (1949) (Brakman et al., 2009).

period. As a result, the size distribution in contemporary Finland is reasonably well approximated by a Pareto distribution.

The results also suggest that Zipf's law did not hold in the pre-war period but gradually became valid during the process of industrialization and urbanization. The estimated values of the Pareto coefficient rise constantly from a low of 0.55 to a high of 0.88. Zipf's law is confirmed for the years from 1980 onward. The fact that the Pareto coefficient does not show constancy for Finland indicates that structural developments have taken place in the Finnish economy and implies that urban growth is not proportional (Parr, 1985; Brakman et al., 2009). Earlier studies also showed changes in the Pareto coefficient over time, especially in the U.S. (Dobkins and Ioannides, 2001; Black and Henderson, 2003). Although the population is less evenly distributed in more recent years, it should be noted that urban agglomeration is not yet high in the last year of the data because the estimated coefficient does not reach unity.

Table 2. OLS results from the estimation of the rank-size rule

Year	Constant	$\hat{a}$	S.E.	R^2
1880	11.75	0.55***	0.09	0.69
1890	11.89	0.55***	0.09	0.71
1900	12.17	0.60***	0.10	0.65
1910	12.16	0.57***	0.09	0.82
1920	12.22	0.58***	0.10	0.84
1930	12.30	0.58***	0.10	0.87
1940	12.46	0.61***	0.10	0.89
1950	12.70	0.62***	0.10	0.91
1960	12.91	0.67***	0.11	0.93
1970	13.09	0.74**	0.12	0.94
1980	13.24	0.79	0.12	0.93
1990	13.33	0.81	0.13	0.94
2000	13.46	0.86	0.14	0.94
2004	13.52	0.88	0.15	0.95

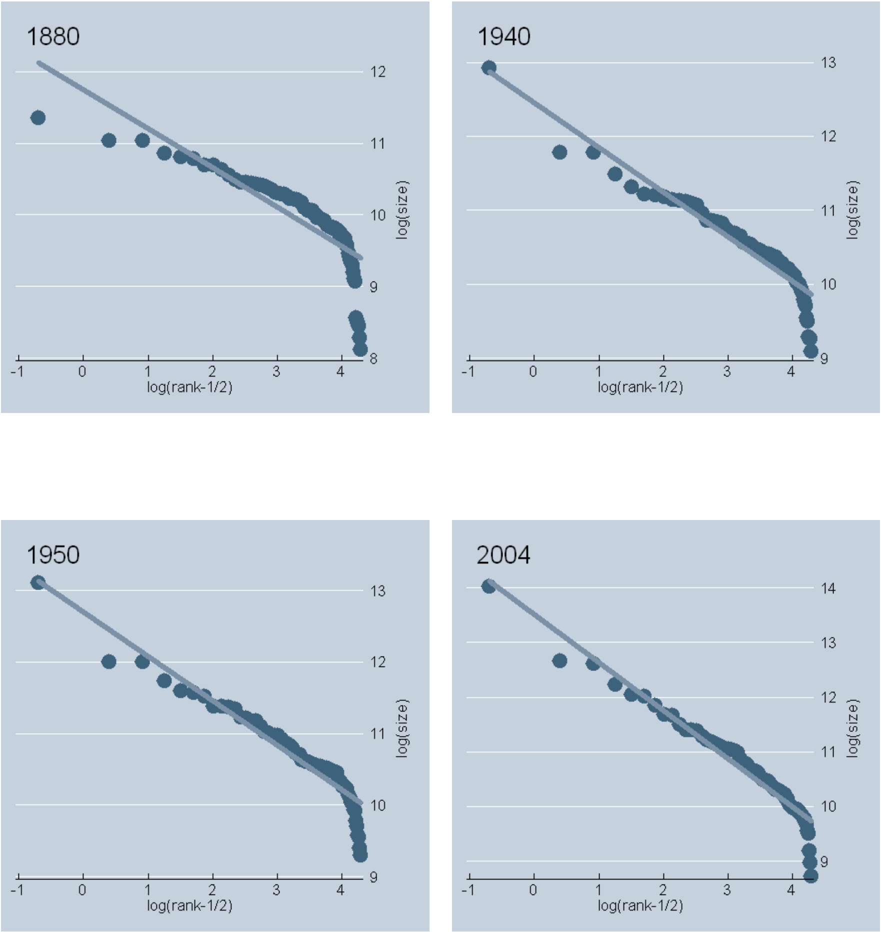
Note: The estimated equation is *log (size) = log (constant) –* α *log (rank-1/2)* (See text and Gabaix and Ibragimov, 2007).

*** Reject H_0; α=1 at 1% level of significance

** Reject H_0; α=1 at 5% level of significance

* Reject H_0; α=1 at 10% level of significance

Figure 2. Rank-size distributions and linear regression lines at four different dates



In all, except for stability, the analysis also shows substantial changes in regional structures. Especially during the first study period, new regions also found their way into a virtuous circle of growth and development. The economic landscape is not completely predetermined, but persistence in regional structures in Finland has increased with time as NEG predicts.

6. THE ROLE OF CENTERS IN THEIR REGIONS

To evaluate our second hypothesis about the rising dominance of growth centers, causal relationships between growth centers and peripheries are analyzed. The importance of cities has increased markedly in Finland. In terms of population, the population share of the 19 centers increased from 39% in 1880 to 67% in 2004 (see Table 1 and Appendix).

Comparing the periods 1880-1940 and 1950-2004, we can see that most of this increase occurred after the war. The share of the population residing in the centers increased in all regions during the post-war period, whereas it decreased in six regions during the pre-war period.

We apply a novel testing procedure based on an extension of Granger causality in the context of a panel framework, which allows for possible heterogeneity between regions. Panel Granger tests are significantly more efficient than conventional Granger tests (Baltagi, 2005; Hurlin and Venet, 2001, 2005; Hood III et al., 2008). A causal relationship may be present only in a subset of regions. The nested testing procedure, as first proposed by Hurlin and Venet (2001), has been applied, for example, in Hood III et al. (2008), Erdil and Yetkiner (2009) and Tervo (2009). It consists of three main steps, namely testing the homogeneous non-causality hypothesis; testing the homogeneous causality hypothesis; and testing the heterogeneous non-causality hypothesis.

Following Hurlin and Venet (2001; see also Hood III et al., 2008; Erdil and Yetkiner, 2009), we consider all variables to be covariance stationary and observed for T periods and N cross-section units, which consist of regions in our case. For each cross-section unit $i \in [1, N]$, the variable $x_{i,t}$ causes $y_{i,t}$ if we are better able to predict $y_{i,t}$ using all available information than if this information had not been included. In our case, the cross-section unit is a region, and $x_{i,t}$ and $y_{i,t}$ refer to changes in the populations of centers and their hinterlands, respectively. Let us consider a time-stationary vector autoregressive (VAR) representation adapted to a panel context. For each cross-section unit i (i = 1,…, N) and time period t (t = 1,…,T), we have

$$(4) \qquad y_{i,t} = \sum_{k=1}^{p} \gamma^{(k)} y_{i,t-k} + \sum_{k=0}^{p} \beta_i^{(k)} x_{i,t-k} + v_{i,t},$$

where $v_{i,t} = \alpha_i + \varepsilon_{i,t}$ are *i.i.d.* $(0, \sigma_\varepsilon^2)$ and p is the number of lags. The autoregressive coefficients $\gamma^{(k)}$ and the regression coefficient slopes $\beta_i^{(k)}$ are assumed to be constant for all lag orders. It is also assumed that $\gamma^{(k)}$ are identical for all units, whereas $\beta_i^{(k)}$ are allowed to vary across individual cross-sections. This is a panel data model with fixed coefficients. The general definitions of causality imply testing for linear restrictions on these coefficients in three main steps.

We follow the nested procedure to test different causality relationships. The Granger causality tests between the growth of centers and the growth of peripheries in 19 regions in Finland are performed for two periods, 1880-1940 and 1950-2004, with lags t and $t\text{-}1$. In our empirical application, we allow the instantaneous case; that is, $\beta_i^{(0)}x_{i,t}$ is included in the model (Hurlin and Venet, 2001). This is because the data are in intervals of ten years. If we had ignored lag t, we would have assumed that anything that happened in x during the first ten-year interval had no effect on y. Testing with longer lag length than $t\text{-}1$ cannot be carried out because of the small number of periods. For both variables, we take the natural logarithm and difference them to eliminate possible unit roots and to reach time stationarity. The tests are based on Wald statistics. To test the various hypotheses, we calculated the test statistics using the sum of squared residuals from the unrestricted model (4) and the sum of squared residuals from the requisite restricted models. The sums of squared residuals are obtained from the MLE, which in this case corresponds to the fixed effects estimator. To perform the estimations required, we used the constrained regression technique.

As a first step in exploring bi-directional Granger causality between population growth in centers and in peripheries, the homogeneous instantaneous non-causality (HINC) hypothesis is assessed. The HINC hypothesis implies the non-existence of individual causal relationships. In model (4), the corresponding test is defined by

$$(5) \qquad H_o\!: \beta_i^{(k)} = 0 \ \forall\ i \in [1, N],\ \forall\ k \in [0, p]$$

$$H_1\!: \exists (i, k) \,/\, \beta_i^{(k)} \neq 0\ .$$

For testing Np linear restrictions in (5), the following Wald statistic is computed:

$$(6) \qquad F_{HINC} = \frac{(RSS_2 - RSS_1)/Np}{RSS_1/(NT - N(1+p) - p)} ,$$

where RSS_2 denotes the restricted sum of squared residuals obtained under H_o and RSS_1 corresponds to the residual sum of squares of model (4). If the individual effects α_i are assumed to be fixed, the sum of squared residuals is obtained from the maximum likelihood estimation (MLE), which in this case corresponds to the fixed effects (FE) estimator. The results are shown in Table 3. The test statistics are significant with lag *t* but not with lag *t-1*. Overall, these results allow us to reject the HINC hypothesis. Thus, for at least one region (and possibly all of them), there is statistical evidence of Granger causality from growth in centers to growth in peripheries and *vice versa*.

Table 3. Test results for homogeneous instantaneous non-causality (HINC hypothesis) and homogeneous causality (HC hypothesis)

Lags	F_{HINC}		F_{HC}	
	1880-1940	1950-2004	1880-1940	1950-2004
Causality from center to hinterland				
Lag *t*	3.30***	5.59***	1.25	3.24***
Lag *t-1*	0.88	0.40	0.51	0.42
Causality from hinterland to center				
Lag *t*	4.73***	5.02***	1.25	2.79***
Lag *t-1*	0.74	1.21	0.54	0.48

***	Reject H_0 at 1% level of significance
**	Reject H_0 at 5% level of significance
*	Reject H_0 at 10% level of significance

Given the rejection of the HINC hypothesis, the next step is to test the hypothesis of homogeneous causality (HC). The F_{HC} test statistic is calculated using the sum of squared residuals from the unrestricted model described above (RSS_1) and the sum of squared residuals (RSS_3) from a restricted model in which the slope terms are constrained to equality for all panel members in the sample. Thus, the hypotheses are

(7) $$H_o: \forall k \in [0, p] / \beta_i^{(k)} = \beta^{(k)} \ \forall i \in [1, N]$$

$$H_1: \exists k \in [0, p], \ \exists (i, j) \in [1, N] / \beta_i^{(k)} \neq \beta_j^{(k)} ,$$

and the test statistic is

(8) $$F_{HC} = \frac{(RSS_3 - RSS_1) / p(N - 1)}{RSS_1 / (NT - N(1 + p) - p)} .$$

As in the case of HINC, if the individual effects α_i are assumed to be fixed, the ML estimator is consistent with the FE estimator. The results are shown in Table 3. For the first period, 1880-1940, the test statistics are not significant, whereas for the second period, 1950-2004, they both are significant at the 0.1% level. These results, interestingly, suggest that causal processes were homogenous and positive between the growth of centers and their hinterlands during the 1880-1940 period, but they were either heterogeneous in both cases or did not exist across all regions during 1950-2004.

The next step in the search for Granger causality is to determine the contributions of individual regions to the existence of causality in the post-war period; i.e., to test the heterogeneous non-causality hypothesis (HENC). For the pre-war period, the test results indicate causal relationships to be homogenous and present for all regions so that HENC testing is not needed. For the post-war period, it may still be possible that causal relationships exist for one or more regions. The F_{HENC} statistic is calculated using RSS_1 obtained above in addition to the sum of squared residuals ($RSS_{2,i}$) from a model in which the slope coefficient for panel member *i* in question is set to zero. The hypotheses in this case are

(9) $$H_o: \exists i \in [1, N] / \forall k \in [0, p] \ \beta_i^{(k)} = 0$$

$$H_1: \forall i \in [1, N], \ \exists k \in [0, p] / \beta_i^{(k)} \neq 0.$$

To test these hypotheses, the following statistic is calculated:

$$(10) \qquad F_{HENC} = \frac{(RSS_{2,i} - RSS_1)/p}{RSS_1/(NT - N(1+2p)+p)} \, .$$

These N individual tests identify the regions for which there are no causal relationships. The results for each region are presented in Table 4. The test statistics are only calculated with lag t as the previous results suggest this result to be the strongest.

Table 4. Test results for heterogeneous non-causality (HENC hypothesis, lag *t*); 1950-2004 period

Region	*Causality from center to hinterland*		*Causality from hinterland to center*	
	F_{HENC}	(Sign of the effect)	F_{HENC}	(Sign of the effect)
Uusimaa	0.71	(+)	3.70*	(+)
Itä-Uusimaa	0.19	(+)	0.11	(+)
Southwest Finland	2.28	(-)	4.64**	(-)
Satakunta	1.43	(+)	2.30	(+)
Häme	0.53	(+)	0.55	(+)
Tampere Region	0.03	(-)	0.11	(-)
Päijät-Häme	0.66	(+)	1.96	(+)
Kymenlaakso	5.66**	(+)	4.55**	(+)
South Karelia	0.39	(+)	0.74	(+)
Etelä-Savo	3.46*	(+)	0.78	(+)
Northern Savo	0.04	(+)	0.02	(+)
Northern Karelia	10.27***	(+)	3.08*	(+)
Central Finland	2.99*	(+)	0.49	(+)
Southern Ostrobothnia	0.74	(+)	0.18	(+)
Ostrobothnia	0.01	(-)	0.00	(-)
Central Ostrobothnia	1.81	(-)	0.67	(-)
Northern Ostrobothnia	0.35	(+)	0.38	(+)
Kainuu	34.60***	(+)	16.93***	(+)
Lapland	16.65***	(+)	32.59***	(+)

*** Reject H_0 at 1% level of significance
** Reject H_0 at 5% level of significance
* Reject H_0 at 10% level of significance

The results show there to be few significant results, and some of the relationships are even negative. For three northern and eastern centers, namely Joensuu, Kajaani and Rovaniemi, the test results still indicate strong positive causal relationships from center to periphery. Interestingly, these centers are located in some of the least developed and poorest regions of Finland. Overall, these results suggest that causal processes are changing. Earlier results by Tervo (2009) based on yearly changes during the 1970-2004 period show that large, rapidly growing centers have had negative effects on their peripheries, whereas the effects have been positive for those regions that have slowly growing (or weak) centers. Our new results from a longer post-war period with data in decade intervals clearly illustrate this change.

7. CONCLUSIONS

Utilizing regional population data for the years of 1880-2004, this paper aimed to test two predictions of the new economic geography (NEG) models by analyzing the evolution of regional structures and spatial interactions between developing cities and their local peripheries in Finland in the pre- and post-war periods. The paper, then, is one of the first studies in which long-term regional growth in a country has been analyzed from this perspective.

The results indicate that the rank-size distribution provides a better characterization of sub-region size distributions after the war than before the war in Finland and that Zipf's law has only gradually become valid during the process of industrialization and urbanization. Furthermore, the analyses of the changes in rankings of the sub-regions also suggest an increase in the stability of regional structures during the post-war period. Thus, persistence in locational patterns increased in Finland during industrialization and urbanization as predicted by the NEG models.

In the analysis of causal processes between the two types of regions, the test results showed causal relationships to be homogenous and positive for the pre-war period of 1880-1940. This result reflects the situation when Finland was still dominated by primary production and internal migration was not yet extensive; in this context, centers and their hinterlands grew hand-in-hand. Rapid industrialization and urbanization only started after the war, which brought with it more clearly imbalanced regional development. Both the homogenous non-causality hypothesis and the homogenous causality hypothesis were rejected for the post-war period of 1950-2004, during which industrialization explosively accelerated and internal migration from rural areas to towns increased. Thus, we reach the result that causal processes between the centers and their hinterlands exist and are heterogeneous. The results suggest some insights regarding the agglomeration shadow that cities cast on their local hinterlands. Government intervention attempted to minimize unbalanced regional growth, but market forces were clearly stronger. These results are in line with the results of Tervo (2009), who showed that large and rapidly growing centers in Finland have had backwash effects on their hinterlands since the 1970s.

REFERENCES

Baltagi, B.H., 2005. *Econometric Analysis of Panel Data*. Wiley, New York.

Black, D., Henderson, V., 2003. Urban evolution in the USA. *Journal of Economic Geography,* 3, 343-372.

Bosker, M., Brakman, S., Garretsen, H., Schramm, M., 2007. Looking for multiple equilibria when geography matters. German city growth and the WWII shock. *Journal of Urban Economics* 61,152-169.

Brakman, S, Garretsen, H., 2006. New economic geography: closing the gap between theory and empirics. *Regional Science and Urban Economics,* 36, 569-572.

Brakman, S, Garretsen, H., van Marrewijk, C., 2009. *The New Introduction to Geographical Economics*. Cambridge University Press, Cambridge.

Brakman, S. Garretsen, H., Schramm, M., 2004. The strategic bombing of German cities during WWII and its impact on city growth. *Journal of Economic Geography* 4, 201-218.

Davis, D.R., Weinstein, D.E., 2002. Bones, bombs and break points: the geography of economic activity. *American Economic Review* 92, 1269-1289.

Dobkins, L.H., Ioannides, Y.M., 2001. Spatial interactions among U.S. cities: 1900-1990. *Regional Science and Urban Economics,* 31, 701-731.

Erdil, E., Yetkiner, I.H., 2009. The Granger-causality between health care expenditure and output: a panel data approach. *Applied Economics*, 41, 511-518.

Fujita, M., Krugman, P., Venables, A.J., 1999. *The Spatial Economy – Cities, Regions, and International Trade*. The MIT Press, Cambridge.

Gabaix, X., 1999. Zipf's law for cities: an explanation. *Quarterly Journal of Economics* 114, 739-767.

Gabaix, X, Ibragimov, R., 2007. Rank – ½: a simple way to improve the OLS estimation of tail exponents. NBER Technical Working Paper Series 342, Massachusetts.

Gabaix, X., Ioannides Y., 2004. The evolution of city size distributions. In: Henderson, V., Thisse J. (Eds.) *Handbook of Regional and Urban Economics*, Volume 4. North-Holland, Amsterdam, pp. 2341-2378.

Head, K., Mayer, T., 2004. The empirics of agglomeration and trade. In: Henderson, V., Thosse, J-F. (Eds.), *Handbook of Regional and Urban Economics*, Volume 4.. North-Holland, Amsterdam, pp. 2609-2665.

Hood III, M.V., Kidd, Q., Morris, I.L., 2008. Two sides of the same coin? Employing Granger causality tests in a time-series cross-section framework. *Political Analysis* 16, 324-344.

Hurlin, C., Venet, B., 2001. Granger causality tests in panel data models with fixed coefficients: Mimeo, University of Paris IX.

Hurlin, C., Venet, B., 2005. Testing for Granger causality in heterogeneous panel data models. *Revue Economique*, 56, 1-11.

Kangasharju, A., 1998. Regional economic differences in Finland: variations in income growth and firm formation. Pellervo Economic Research Institute Publications 17, Helsinki.

Krugman, P., 1991. Increasing returns and economic geography. *Journal of Political Economy* 99, 483-499.

Krugman, P., 1993. First nature, second nature, and metropolitan location. *Journal of Regional Science*, 33, 129-144.

Krugman, P., 1996. *The Self-Organizing Economy. Blackwell,* Oxford.

Myrdal, G., 1957. *Economic Theory and Under-Developed Regions*. Duckworth, London.

Ottaviano, G.I.P., Pinelli D., 2004. The challenge of globalization for Finland and its regions: The new economic geography perspective. Prime Minister's Office Publications 24. Edita, Helsinki.

Ottaviano, G.I.P., Pinelli D., 2006. Market potential and productivity: evidence from Finnish regions. *Regional Science and Urban Economics* 36, 636-657.

Ottaviano, G.I P., Puga, D., 1998. Agglomeration in the global economy: a survey of the ‘New Economic Geography’. *World Economy* 21, 707-731.

Overman, H.G., Ioannides, Y. (2001) Cross-sectional evolution of the US city size distribution. *Journal of Urban Economics*, 49, 543-566.

Parr, J.B., 1999a. Growth-pole strategies in regional economic planning: a retrospective view. Part 1, origins and advocacy. *Urban Studies* 36, 1195-1215.

Parr, J.B., 1999b. Growth-pole strategies in regional economic planning: a retrospective view. Part 2, implementation and outcome. *Urban Studies* 36, 1247-1268.

Partridge, M.D., Rickman, D.S., Ali, K., Olfert, M.R., 2008. Lost in space: population growth in the American hinterlands and small cities. *Journal of Economic Geography*, 8, 727-757.

Partridge, M.D., Rickman, D.S., Ali, K., Olfert, M.R., 2009. Do New Economic Geography agglomeration shadows underlie population dynamics across the urban hierarchy. *Papers in Regional Science*, 88, 445-466.

Pitkänen, K., 1994. Historical development patterns of Finnish population (In Finnish: Suomen väestön historialliset kehityslinjat). In: Koskinen S., Martelin T., Notkola I-L., Notkola V. Pitkänen K. (Eds.) Finland’s population (in Finnish: Suomen väestö). Gaudeamus, Helsinki.

Richardson, H.W., 1978. *Regional and Urban Economics.* Penguin, New York.

Simon, H., 1955. On a class of skew distribution functions. *Biometrika* 42, 425-440.

Tervo, H., 2005. Regional policy lessons from Finland. In: Felsenstein, D., Portnov, B.A. (Eds.), *Regional Disparities in Small Countries*. Springer-Verlag, Berlin.

Tervo, H., 2009. Centres and peripheries in Finland: Granger causality tests using panel data. *Spatial Economic Analysis*, 4, 377-390.

Virrankoski, P., 1975. The economic history of Finland (In Finnish: Suomen taloushistoria). Otava, Helsinki.

Zipf, G., 1949. *Human Behaviour and the Principle of Least Effort.* Addison Wesley, Cambridge, Massachusetts.

Appendix. Rankings and populations of centers and their regions in 1880, 1940, 1950 and 2004

Center / *Region*	1880		1940		1950		2004	
Helsinki	1	84,569	1	411,019	1	491,594	1	1,224,257
- Uusimaa	*7*	*123,896*	*1*	*475,531*	*1*	*574,025*	*1*	*1,346,958*
Pori	2	61,805	4	97,365	4	124,012	7	138,615
- Satakunta	*10*	*110,316*	*8*	*174,653*	*7*	*220,815*	*7*	*230,702*
Turku	3	61,789	3	129,711	3	162,711	3	296,858
- Southwest Finland	*1*	*176,787*	*2*	*282,314*	*3*	*344,286*	*3*	*453,745*
Mikkeli	4	51,594	13	63,551	12	75,317	16	71,846
- Etelä-Savo	*4*	*132,669*	*9*	*164,529*	*10*	*196,386*	*15*	*161,381*
Joensuu	5	49,375	5	81,321	6	105,168	9	115,360
- Northern Karelia	*11*	*104,194*	*10*	*163,453*	*9*	*198,775*	*13*	*168,615*
Vaasa	6	48,224	9	69,103	14	71,552	12	88,798
- Ostrobothnia	*6*	*129,260*	*11*	*150,144*	*13*	*159,535*	*12*	*173,435*
Tampere	8	43,949	2	130,729	2	163,424	2	316,023
-Tampere Region	*3*	*140,797*	*3*	*279,817*	*2*	*352,138*	*2*	*464,976*
Lahti	9	41,281	8	71,172	5	107,757	5	169,386
- Päijät-Häme	*13*	*64,435*	*16*	*94,418*	*15*	*139,050*	*8*	*198,685*
Jyväskylä	10	38,277	10	68,166	7	100,569	6	163,390
- Central Finland	*9*	*110,490*	*7*	*176,519*	*6*	*234,920*	*5*	*267,182*
Hämeenlinna	11	36,008	15	52,192	13	73,506	11	89,053
- Häme	*12*	*65,464*	*15*	*106,912*	*14*	*144,220*	*14*	*167,630*
Kuopio	16	33,291	14	57,486	15	70,767	8	118,050
- Northern Savo	*2*	*145,479*	*5*	*215,143*	*5*	*258,737*	*6*	*251,095*
Oulu	17	32,822	12	65,702	11	83,449	4	202,898
- Northern Ostrobothnia	*5*	*132,606*	*4*	*230,142*	*4*	*272,461*	*4*	*374,928*
Kouvola	18	32,028	11	66,793	8	87,517	10	97,563
- Kymenlaakso	*14*	*60,387*	*12*	*139,825*	*11*	*173,689*	*11*	*185,541*
Porvoo	28	27,115	27	38,056	27	45,805	15	73,795
- Itä-Uusimaa	*16*	*43,640*	*18*	*59,003*	*18*	*69,780*	*17*	*92,442*
Seinäjoki	30	26,155	28	38,043	29	44,537	19	64,791
- South Ostrobothnia	*8*	*124,437*	*6*	*185,495*	*8*	*210,324*	*9*	*194,076*
Kajaani	45	18,678	25	42,133	24	51,629	24	58,648
- Kainuu	*18*	*35,585*	*17*	*71,923*	*17*	*87,360*	*18*	*85,965*
Kokkola	46	18,648	48	27,793	45	36,181	25	52,355
- Central Ostrobothnia	*19*	*30,897*	*19*	*47,519*	*19*	*58,582*	*19*	*70,482*
Lappeenranta	66	10,884	24	42,274	25	50,802	17	69,790
- South Karelia	*15*	*55,708*	*14*	*119,830*	*16*	*137,675*	*16*	*135,800*
Rovaniemi	72	5,041	51	26,935	42	36,994	21	62,371
- Lapland	*17*	*39,698*	*13*	*133,633*	*12*	*167,552*	*10*	*186,443*