ISSN 1503-299X

WORKING PAPER SERIES

No. 2/2011

MIGRATION AND DYNAMIC AGGLOMERATION ECONOMIES: REGIONAL INCOME GROWTH IN NORWAY

Jørn Rattsø and Hildegunn E. Stokke

Department of Economics

Norwegian University of Science and Technology
N-7491 Trondheim, Norway
www.svt.ntnu.no/iso/wp/wp.htm

Migration and dynamic agglomeration economies: Regional income growth in Norway[*]

Jørn Rattsø and Hildegunn E. Stokke

Department of Economics, Norwegian University of Science and Technology

jorn.rattso@svt.ntnu.no, hildegunnes@svt.ntnu.no

Abstract

The existence of agglomeration economies is well established. We study dynamic agglomeration effects and expect regional income divergence when people move from the periphery to cities. We use distribution analysis, Kernel density functions and first order Markov chains, to investigate whether data from Norwegian regions 1972-2008 with strong urbanization are consistent with agglomeration effects. The analysis shows overall income convergence, contrary to the hypothesis of agglomeration economies. Regions with large increases in population do not show systematic higher income growth. The probabilities of moving up and down the income distribution are independent of the migration pattern. We conclude that strong dynamic agglomeration effects linked to inmigration are inconsistent with the data.

JEL codes: C14, J61, O15, R11, R12

Key words: Agglomeration economies, urban evolution, migration, income convergence, Kernel density function, Markov chain

Date: February 1, 2011

[*] We appreciate data handling by Jo Jakobsen, background documentation and computer programs received from Thomas Herzfeld, discussions at the 2010 ERSA conference and at staff seminars at Statistics Norway, the Kiel Institute for the World Economy and the University of Bergen, and comments from Rolf Aaberge, Frank Bickenbach, Holger Gørg, and Magne Mogstad. The project is funded by the Norwegian Research Council.

1. Introduction

Productivity and income are higher in urban areas compared to rural. This stylized fact can be observed in all industrialized countries. Agglomeration of economic activity into cities seems to raise productivity and thereby income.

The level agglomeration effect means that large population flows to urban centers may generate economic gains. The large population shift from rural to urban areas in Norway during the last decades motivates this investigation of the agglomeration effect of migration. Only few studies of dynamic agglomeration effects are available. The importance of migration for the income process has been analyzed with distribution analysis of US states by DiCecio and Gascon (2010). They find that states that are losing position in the income distribution are losing population, while states climbing the ladder of the income distribution are gaining population. Au and Henderson (2006) argue that migration restrictions limit agglomeration in China and that insufficient agglomeration has led to first-order losses of GDP. These studies suggest that increased urban income growth and divergence are expected as a result of the large migration flows in Norway.

The analysis of agglomeration effects has typically been static and emphasizing scale and externalities (see recent overviews by Glaeser and Gottlieb, 2009 and Moretti, 2010). The mechanisms of agglomeration effects are reviewed by Duranton and Puga (2004) separating between scale, matching and learning. New economic geography has added understanding of the dynamic mechanics of urbanization and growth (see overview by Baldwin and Martin, 2004). Recent empirical evidence about agglomeration effects are summarized by Combes (2010), Puga (2011) and Redding (2010). The main empirical challenge is to identify the agglomeration effect in a setting where migration endogenously responds to urbanization and the population is heterogeneous with respect to skill and education. The studies show overwhelming support for agglomeration economies with different datasets and methodologies. Combes (2010) presents estimates with particular relevance for France: The elasticity of production per worker with respect to density corrected for endogeneity and skill sorting is about 0.02 and implies that a doubling of density gives 2% higher production per worker. The associated shift in skill composition of the urban areas offer an extra gain of 2% and the endogenous increase in density with higher income offers another 1%. A doubling of density is assumed to give 5% higher production per worker and income overall. The estimates using various instruments for density are typically in this order of magnitude.

The recent econometric studies estimating agglomeration effects have addressed the methodological challenges related to endogeneity of migration and sorting. They have used individual and firm level data to handle sorting. This is obviously a strength, but long time series of individual and firm level data are not typically available. The analyses of the level effect of population size and density for productivity and income have used instrument variables to handle the endogenous interaction of population, productivity and income. Good and valid instruments are not available as exogenous driving forces during long time periods. We have consequently looked for alternative methodological approaches. To analyze the linkages between migration and dynamic agglomeration effects we study regional level data on income per capita in Norway during 1972-2008. To capture the dynamics and the heterogeneity among regions we work with distribution analysis and Markov chains. The analysis concentrates on tests of the dynamic characteristics of regional income growth and their links to changes in the population sizes.

The analysis shows a consistent pattern of urbanization during the past four decades. People have moved from the periphery to urban centers and the urban evolution is common to most industrialized countries (Black and Henderson, 2003, for the US and overview). During this period of large shift in the population to urban areas, incomes have converged towards unimodal distribution of per capita income. The analysis is made both at the level of 396 municipalities and 89 economic regions defined by the European Union classification (NUTS-4). The relationship between migration and income growth is studied in detail for municipalities with more than 5.000 inhabitants in 2008 (about half the sample). We separate between municipalities with relatively stable population and municipalities with strong inmigration, and find that the probabilities of moving up and down the income distribution are independent of the migration pattern. The hypothesis of equal transition probabilities across municipalities with different migration pattern cannot be rejected at 5% significance level.

Strong agglomeration effects should contribute to income divergence among regions during a period of large shifts of the population to urban areas. In addition to this, the shift in the population has included a composition effect with increasing relative skill level in the urban areas. The combined agglomeration and skill-composition effect should give strong urban income growth. Our results are not consistent with strong dynamic agglomeration effects and high income urban regions taking off. Rattsø and Stokke (2011) pursue the role of education in the same dataset of regional income and conclude that there are limited human capital effects.

Section 2 discusses our approach to the analysis of agglomeration. The use of distribution analysis and Markov chains is explained in section 3. The urban evolution is documented in section 4 and shows large population shifts across regions. The income convergence in the data is shown in the Kernel distribution densities and Markov chains analyzed section 5. Section 6 offers tests of time stationarity and time independence. The investigation of dynamic agglomeration effects are pursued in section 7. Some robustness checks of the results are reported in section 8. Concluding remarks are given in section 9.

2. Investigating the effects of agglomeration

The key sources of agglomeration economies were outlined already by Alfred Marshall (see overviews by Glaeser and Gottlieb, 2009 and Moretti, 2010). He argued for the role of matching of worker skills and firms in larger labor markets, the specialization and differentiation of intermediate inputs, and the role of knowledge spillovers. The theoretical micro-foundations discussed by Duranton and Puga (2004) clarify the mechanisms of agglomeration effects. Agglomeration economies are related to different economic advantages of population density implying that the concentration of population contributes to higher productivity and income. This is consistent with the stylized fact that income and productivity levels are higher in urban compared to rural regions. But also other mechanisms are involved in explaining the observed productivity and income differences. The high income level in urban regions may result from local resources as the source of high productivity. People move to cities because they are more productive, the cities are not necessarily more productive because they have more people. And cities may offer amenities and services that motivate skilled and productive people to move to cities. Cities have productive people, they don't necessarily make them productive.

The econometric studies of agglomeration effects have started to address the endogeneity and sorting issues as discussed in the introduction. The causal effect of migration and urbanization is hard to identify because of the simultaneous determination of density and productivity and the heterogeneity of the population with respect to productivity. Standard econometric analysis of the relationship between productivity and population density is expected to overstate the importance of migration. Ciccone and Hall (1996) innovated the handling of endogeneity by using lagged population variables as instruments. Glaeser and Gottlieb (2009) have a migration equilibrium model as point of departure and apply measures of amenities as instrument to predict population. Combes et al. (2009) extend the menu of instruments and introduce geographical variables and make use of firm level data. The analyses attempt at estimating the elasticity of productivity and income with respect to population size and density. Ciccone and Hall (1996) find that a doubling of employment density increases labor productivity by 6% using US county and states data for 1988. Later studies find somewhat lower elasticities. The meta analysis of Melo et al. (2009) reveals heterogeneity in agglomeration economies, and with elasticities typically in the range of 0.02-0.10. Recent studies have looked into the composition effect by using individual data. Combes et al. (2008) conclude that skill-composition is the major explanatory factor of geographic wage disparities in France, but that agglomeration effects also are present. Overall the econometric studies show quite convincing agglomeration effects.

The econometric estimates of agglomeration economies should be understood in the context of general equilibrium effects. The standard framework of migration equilibrium is the Rosen-Roback model as laid out by Glaeser and Gottlieb (2008) and Moretti (2010). The model can capture the role of scale and externalities in urban areas leading to higher productivity and wages. But the migration equilibrium serves as a constraint on the observed income differences between urban and rural areas. If workers are perfectly mobile and with equalized indirect utilities across regions, the urban inhabitants will not experience a welfare gain with higher productivity. The higher wages will be balanced by higher congestion and housing costs. The assumption of perfect mobility is too strong. It seems realistic to assume that higher wages will reflect higher productivity, but the real income will be reduced by housing price effects. Our interest is in the dynamic process, and Rappaport (2005) has integrated the Rosen-Roback model into a neoclassical growth model of multiple regions. In this setting migration modifies the income convergence effect of the allocation of capital. The labor supply effect of inmigration contributes to lower wages. However, he concludes that migration has an ambiguous effect on income convergence. The decreased marginal product of labor following inmigration is counteracted by higher marginal product of capital when there are complementarities between the factors of production. Given the convergence factors, strong agglomeration effects are needed to reach overall income divergence. The neoclassical model can be extended to generate divergence as a result of migration with heterogeneous labor and associated skill effects and brain drain/gain, as shown by Rattsø and Stokke (2009).

We are interested in the long run growth process and the dynamics of the agglomeration effects. The desired dataset would reflect a natural experiment with large shifts in population over time independent of local income generation. But finding instruments to represent some exogenous part of the long run shift in the population pattern is difficult. We want to take benefit of a long time series of urbanization and it follows that we cannot convincingly separate between the consequences of migration for agglomeration effects, the importance of agglomeration effects for migration, and other sources of increasing productivity and income. We investigate whether our time series observations are consistent with strong agglomeration effects. The analysis addresses the relationship between the changes in the population pattern and the changes of income levels.

We conclude that observation of a positive association between inmigration and income growth is a necessary, but not sufficient, condition to have positive agglomeration economies. It is necessary assuming that the agglomeration effects are strong enough to dominate the labor supply effect of migration. But it is not sufficient, since other factors beside agglomeration economies may explain the income growth following inmigration. Nevertheless, the empirical and theoretical literature points to strong arguments for high income and productivity growth in cities with high population growth and consequent income divergence. Income convergence in a country with strong population flow from rural to urban areas indicates that the agglomeration effects are limited.

3. Distribution analysis

Our approach is to study whether the data are consistent with dominant agglomeration effects in regions with increased population. We observe large shifts in the population across 396 municipalities (and 89 labor market regions) in Norway during the past four decades. The analysis attempts at linking the expansion of the population in the urban areas with increases in the income level. A positive relationship between increased population and increasing income indicates the possibility of agglomeration effects. As will come clear, we do not find such a relationship. There is no systematic pattern of higher increases in income in regions with higher population growth. This result follows even though the endogenous migration response to higher income and skill composition effects of the labor force both should work in the direction of a positive relationship between inmigration and income growth. Also we analyze nominal incomes and counteracting urban price effects are not taken into account. It is hard to argue that agglomeration effects are important with this pattern of data.

The method is based on distribution analysis and Markov chains to capture heterogeneous processes with different growth paths from different starting points. Using distribution analysis we can in particular study the two ends of the distribution of per capita incomes – relative low income and relative high income regions. Quah (1993a, 1993b, 2001) developed this methodology, more recently applied and extended by Kremer et al. (2001). The basics of the method are presented by Shorrocks (1978). We estimate the transitional probabilities of the Markov chains by the maximum likelihood method to facilitate tests of homogeneity and dependence, as well as tests of how population changes are related to income transitions. Our discussion of the method is related to Bichenbach and Bode (2003).

The analysis is based on data for taxable income of each of the 396 municipalities and calculated per capita based on the number of residents in the beginning of the year. The data cover all years during the period 1972-2008 and it follows that we have 14.652 observations of per capita incomes in different localities at different years. Personal income measured in the tax statistic basically reflects wage income, and capital income is hard to locate at this level of disaggregation. No municipal GDP measure is available. We exploit the data at the municipal level, but municipalities are small (about 10.000 inhabitants on average) and administrative borders may cut through functional regions. To investigate functional regions we also analyze an aggregation of municipalities into common labor market areas. Following the European Union standard of NUTS-4 regions, the municipalities are aggregated to 89 labor market regions. This level of aggregation captures economic regions understood as a common labor market.

In the analysis the income level is measured relative to the average income per capita across regions in each year. The whole range of relative per capita income is divided into a finite number of N mutually exclusive income groups and in this analysis we follow the convention of working with five groups (N=5). For each region we get a sequence of variables describing the income group of that region at time *t*. The sequences are considered as independent realizations of a single homogeneous Markov chain with finite group space N. The assumption of a finite first order Markov chain implies that the probability of being in a specific income group at time *t* only depends on the group of the previous period (and not earlier periods). The transition probability, the probability of moving from group *i* to group *j* from period *t-1* to period *t*, is described by $p_{ij}(t)$. The probability is estimated based on observations of how regions move between income groups over time. The number of regions moving from group *i* to group *j* from period *t-1* to *t* is measured by $n_{ij}(t)$. The total number of regions moving from group *i* from period *t-1* to *t* is measured by $n_i(t-1)=\sum_j n_{ij}(t)$. The Markov chain can be reduced to a product of five mutually independent multinomial distributions (one for each row *i* of the transition matrix). For each time period *t*, the distribution function is:

$$f\left(n_{ij}(t)\right)=\prod_{i=1}^{5} f_i\left(n_{ij}(t)\right)=\prod_{i=1}^{5}\left[\frac{n_i(t-1)!}{\prod_{j=1}^{5} n_{ij}(t)!}\prod_{j=1}^{5} p_{ij}^{n_{ij}(t)}\right] \tag{1}$$

The transition probabilities can be estimated by maximizing the log likelihood of the T multinomials above with respect to p_{ij}:

$$f\left(n_{ij}\right)=\prod_{t=1}^{T} f\left(n_{ij}(t)\right) \tag{2}$$

Given the constraint that the sum of p_{ij} over all *j* is 1, the maximum likelihood estimator is simply the relative frequency of transitions:

$$\hat{p}_{ij}=\frac{n_{ij}}{n_i}=\frac{\sum_{t=1}^{T} n_{ij}(t)}{\sum_{t=1}^{T} n_i(t-1)} \tag{3}$$

where n_{ij} and n_i are the sums of the observed frequencies over all transition periods.

Given the initial distribution of regional income per capita across income groups, $h(0)=\left[h_1(0),h_2(0),h_3(0),h_4(0),h_5(0)\right]$ where $\sum_i h_i(0)=1$, the distribution after the first transition period can be calculated as $h(1)=h(0)\Omega$, where Ω is the estimated 5x5 Markov transition matrix. And similar, the distribution after k transition periods follows as $h(k) = h(0)\Omega^k$. Given that the matrix is regular[1], the distribution converges to the limiting distribution $h^* = \lim_{k\to\infty} h(0)\Omega^k$, which is independent of the initial distribution. This is the ergodic long-run distribution of regional incomes and is estimated based on the Markov chain matrix under the assumption that the transition dynamics remain unchanged.

To check whether the calculated estimates work as estimates of Markov transition probabilities, we test two key properties: time stationarity and time independence. The test of time stationarity investigates whether the transition probabilities are constant over time. The test divides the entire sample period with T transitions into M mutually exclusive and exhaustive subperiods and compares the transition matrices under each of the M subsamples to the entire sample. The estimators are obtained based on the distribution function above, and Bickenbach and Bode (2003, p. 369) show how the following Pearson (Q) and Likelihood Ratio (LR) test statistics have an asymptotic χ^2 distribution with degrees of freedom equal to the number of independent pairwise comparisons:

$$Q = \sum_{m=1}^{M}\sum_{i=1}^{5}\sum_{j\in A_i} n_{i|m} \frac{\left(\hat{p}_{ij|m} - \hat{p}_{ij}\right)^2}{\hat{p}_{ij}} \sim asy\chi^2\left(\sum_{i=1}^{5}(a_i - 1)(b_i - 1)\right) \quad (4)$$

$$LR = 2\sum_{m=1}^{M}\sum_{i=1}^{5}\sum_{j\in A_{i|m}} n_{ij|m} \ln\frac{\hat{p}_{ij|m}}{\hat{p}_{ij}} \sim asy\chi^2\left(\sum_{i=1}^{5}(a_i - 1)(b_i - 1)\right) \quad (5)$$

A_i is the set of nonzero transition probabilities in the ith row of the transition matrix estimated from the entire sample, while $A_{i|m}$ is the set of nonzero transition probabilities in the ith row of the matrix estimated from the mth subperiod. The total number of transitions from group i in subperiod m and the total number of transitions from group i to group j in subperiod m are given by $n_{i|m}$ and $n_{ij|m}$, respectively. The degrees of freedom is given in the last parenthesis, where a_i is the number of elements in A_i and b_i is the number of subperiods with a positive number of observations in the ith row.

A more direct investigation of the stationarity of the distribution addresses the second eigenvalue of the Markov matrix. When the second eigenvalue λ_2 is less than 1, the cross-

[1] The Markov chain is regular if for some integer k, all entries of Ω^k are positive.

sectional distribution converges to a steady state. The speed of the process towards steady state can be characterized by the asymptotic half life (*hl*) of the chain as shown by Shorrocks (1978):

$$hl = \frac{-\log 2}{\log|\lambda_2|} \tag{6}$$

The test of time independence checks the first order Markov assumption against higher order alternatives. The second order chain assumes that the probability of a region to move to group *j* at time *t* is influenced both by group *i* at time *t-1* and group *h* at time *t-2* and can be defined as p_{hij}. Again Bickenbach and Bode (2003, p 371) show that the Pearson (*Q*) and Likelihood Ratio (*LR*) test statistics have an asymptotic χ^2 distribution with degrees of freedom determined by the number of independent pairwise comparisons:

$$Q = \sum_{h=1}^{5}\sum_{i=1}^{5}\sum_{j\in C_i} n_{hi}\frac{\left(\hat{p}_{hij}-\hat{p}_{ij}\right)^2}{\hat{p}_{ij}} \sim asy\chi^2\left(\sum_{i=1}^{5}(c_i-1)(d_i-1)\right) \tag{7}$$

$$LR = 2\sum_{h=1}^{5}\sum_{i=1}^{5}\sum_{j\in C_{hi}} n_{hij}\ln\frac{\hat{p}_{hij}}{\hat{p}_{ij}} \sim asy\chi^2\left(\sum_{i=1}^{5}(c_i-1)(d_i-1)\right) \tag{8}$$

C_i is the set of nonzero transition probabilities in the *i*th row of the transition matrix estimated from the entire sample minus the first period (since the income group in the previous period is not available for the first period), while C_{hi} is the set of nonzero transition probabilities in the *i*th row of the matrix estimated from the *h*th subsample. The total number of transitions from group *i* given group *h* in the previous period and the total number of transitions from group *i* to group *j* given group *h* in the previous period are given by n_{hi} and n_{hij}, respectively. The degrees of freedom is given in the last parenthesis, where c_i is the number of elements in C_i and d_i is the number of sub-matrices with a positive number of observations in the *i*th row.

4. Population dynamics: Urban evolution

Agglomeration economies result from geographic concentration of households and firms. We concentrate on the development of the population here. The shift in the population across regions is similar to the pattern observed in most industrialized countries – migration from the rural periphery to urban areas. The urban evolution is in accordance with the influential analysis of US data by Black and Henderson (2003). Old and new urban areas grow steadily, and at the same time the relative size distribution of urban areas is fairly stable. The hierarchy of cities is reproduced even over centuries. In Norway, Oslo stays on the top, followed by Bergen, Trondheim and Stavanger, and so forth. The expansion of the cities during the period studied is assumed to lay the ground for positive agglomeration effects on income.

The urban evolution can be measured by the share of the population living in densely built-up areas and this increased from 65% in 1970 to 78% in 2008.[2] The shift in the population density basically reflects migration flows. Since the demographics of fertility and mortality are similar across Norway, the changes in population size primarily reflect migration. The aggregate population increases by 21% during 1972-2008, representing an annual population growth rate of about 0.6%. The municipalities have dramatic different development of population size. When ranking municipalities according to the population change, the bottom 25% with large outmigration had an average population decline of 28% and the top 25% with large inmigration had an average increase in the population size of 60%.[3] Outmigration is dominated by municipalities in the lower half of the income distribution, while inmigration is most common in the upper half of the distribution.

To describe the development in the distribution of population sizes, we estimate Kernel density functions for relative population sizes in 1972 and 2008, illustrated in Figure 1.[4] The horizontal axis represents population relative to the average across municipalities (measured in log), while the vertical axis gives the density of municipalities at different relative population sizes. Both the old and the new function have a single-peak distribution with the majority of municipalities located right below the average population size. The main difference between the estimated distributions is the 2008 loss of density relative to 1972 in the area around the peak. The density is added at the two ends of the distribution, in particular in the lower end. This is exactly the shift in the distribution expected when outmigration reduces the size of many regions while at the same time the largest regions expand in

[2] Densely built-up areas are defined as areas with at least 200 persons and with distances between houses of less than 50 meters. The data is from the Norwegian Social Science Data Services (NSD), www.nsd.uib.no.

[3] The unweighted average increase in population among the 396 municipalities is 9.6%, and the standard deviation is as high as 36%. Almost half the municipalities had a decrease in the absolute population size during the past four decades.

[4] The density estimates are calculated using a Gaussian kernel with bandwidth set according to Silverman's rule of thumb; $1.06\sigma B^{-0.2}$, where σ is the standard deviation of the data and B is the number of observations. This gives bandwidth equal to 0.309 and 0.36 for 1972 and 2008, respectively.

population size. The understanding is that there are many medium-sized municipalities with outmigration that become small municipalities, while large municipalities with inmigration explain the added density at population levels above the average size.

Figure 1 about here

Furthermore we rank municipalities according to population density in 1970 and consider the shares of total population by density quintiles.[5] The broad picture is urbanization. The top quintile increased its population share by 2.4%-points during 1972-2008 (from 58.8% to 61.2%). The population share of the fourth quintile remains roughly constant around 16%, while the other quintiles account for a lower share of total population in 2008 compared to 1972. The degree of urbanization among Norwegian municipalities can be compared to the findings of Beeson et al. (2001) for US counties. They find that the top decile increased its population share by 5%-points during 150 years (1840-1990), while deciles 4-9 are losing population share and the bottom three deciles actually gain relative population. When considering density deciles, we find that the top 3 deciles have higher population share in 2008 compared to 1972, and in total their share increases by 3.2% points. All other deciles are losing population share. This implies that the urbanization in Norway is broader than in the US case documented by Beeson et al., first and for all as a result of the large number of small municipalities in the dataset.

Ranking municipalities based on the population level in 1972 (rather than population density) does not affect our main finding of urbanization. Initially populous municipalities are gaining relative more population and increasing their share in total population. In addition, using data from 89 economic regions (NUTS-4) rather than municipalities, gives the same outcome. This strengthens the case for urbanization.

5. Income convergence

As argued in the introduction, the shift in the population is expected to lead to income divergence among regions due to agglomeration economies. The income growth in the expanding urban areas is stimulated by the inmigration. The income gap between rural areas in stagnation and urban areas in growth is expected to increase over time.

[5] A region's population density is measured as the share of the population living in densely built-up areas.

Barro and Sala-i-Martin (1992) started up a large literature analyzing income convergence. The estimation of structural convergence equations derived from growth models using various national and regional samples broadly support income convergence. But the convergence is heterogeneous, often conditional on other determinants of income, and sometimes restricted to convergence clubs with similar initial conditions. And the convergence is slow, often with a rate of convergence about 2% per year, which implies that the half life of the convergence process is about 35 years. De la Fuente (2002) summarizes the recent literature using econometric analysis of panel data and offers a study of Spanish regions. He concludes that panel analyses of regional data indicate income convergence and with a higher convergence rate than the early static analyses. The econometric approach faces serious challenges of endogeneity and alternative methods have been investigated. The distribution analysis developed by Quah (1993a, b) has given more support to income divergence. An important study by Magrini (1999) finds that EU regions are characterized by a tendency towards divergence, in particular because of the high growth of high income regions. As will become clear below, we find income convergence in Norway even using the distribution method.

As a simple start, we examine how the distribution of relative municipal income per capita develops over time. Figure 2 shows the estimated Kernel density functions for the first year 1972 and the last year 2008[6]. The horizontal axis represents income per capita relative to the average across municipalities, while the vertical axis gives the density of municipalities at different relative income levels. Both functions have a single-peak distribution with the majority of municipalities located close to the average level of income per capita. The estimated distributions show a clear pattern of convergence over time that constitutes the main message of this analysis. Compared to 1972, the distribution is narrower and much more concentrated around the peak in 2008.

Figure 2 about here

The most intensive use of the data estimates transition probability matrices using Markov chains for annual transitions. We have investigated annual transitions, but focus on 6-year transitions in the analysis below. The pattern is the same, and the argument for longer

[6] The density estimates are calculated using a Gaussian kernel with bandwidth equal to 0.0606 and 0.0375 for 1972 and 2008, respectively, consistent with Silverman's rule of thumb.

intervals is to avoid short term fluctuations and thereby have more stable transition paths. Since the database covers the period 1972-2008, we have six 6-year transitions and a total of 2376 observations to estimate transition probabilities. We follow the convention of discretization based on a uniform initial distribution of relative incomes across income groups, which gives the following five groups: 1) less than 89% of the average, 2) between 89% and 95%, 3) between 95% and 101%, 4) between 101% and 109%, and 5) more than 109% of the average income across municipalities. The transition probability matrix is shown in Table 1. As seen from the binomial standard errors given in parentheses, most of the estimated transition probabilities are significant. The exception is the probabilities of moving three or more income groups during a 6-year period.

Table 1 about here

The Markov matrix shows income convergence across municipalities. The distribution of per capita incomes is tending towards a point mass, rather than towards a two-point distribution. Municipalities in the lowest income group (income level relative to the average below 0.89) have 33% probability of catching-up during a 6-year period, and the high income municipalities have about 25% chance of moving down the distribution. Municipalities in income groups 2 and 4 have much higher probability of moving towards the middle of the distribution than towards the end. The probability of moving from group 4 to the high income group is 12%, compared to about 30% chance of moving down the distribution. In other words, the distribution dynamics show no tendencies of a bimodal twin peaked distribution. This pattern is confirmed by the implied ergodic (long-run) distribution given in the last row of the matrix. Municipal incomes go from a uniform distribution initially to a normal distribution in the long-run. The lowest and the highest income groups are reduced from 20% initially to less than 15%, while the middle-income group accounts for about 25% of the municipalities in the long-run. The distribution tends to accumulate in the middle, combined with thinning of both the lower and the higher tail, consistent with income convergence. Low income municipalities become richer and high income municipalities become poorer (relatively speaking), i.e. living standards converge across municipalities.

Since convergence at the municipal level represents a regional classification that is arbitrary with respect to economic conditions, we have done the same analysis at the level of 89 NUTS-4 common labor market regions. Based on regional income per capita relative to the average across regions we separate between five income groups defined to have the same number of observations in each group: 1) less than 90% of the average income per capita, 2) between 90% and 96%, 3) between 96% and 101%, 4) between 101% and 107%, 5) more than 107% of the average income per capita across regions. Table 2 shows the regional Markov matrix based on 6-year transitions. The results are consistent with the findings at the municipal level with income convergence and a single peaked distribution in the long-run. We conclude that both at the municipal and the economic region level the time series data show income convergence.

Table 2 about here

6. Tests of Markov properties: Time stationarity and independence

We investigate the Markov properties of the transition matrix at the municipal level, given in Table 1. The transition probabilities are estimated based on six 6-year transitions for 396 municipalities during 1972-2008. We test for time stationarity (constant transition probabilities over time) by applying the Pearson (Q) and the Likelihood Ratio (LR) test statistics, as described in section 3. The sample period is divided into two subperiods (1972-1990 and 1990-2008), each containing three 6-year transitions. The transition matrices for each subperiod are then compared to the full period matrix. With 18 degrees of freedom the critical value at 5% significance level equals 28.9. The test statistics are calculated to Q = 24.6 (prob = 0.14) and LR = 26.7 (prob = 0.08), which means that the null hypothesis of constant transition probabilities over time cannot be rejected, and the Pearson test even holds at 14% significance level. The detailed contributions to the Pearson test statistic from each transition in the two subperiods are shown in Table 3. The differences in transition probabilities are minor. Of the 46 comparisons (23 probabilities, 2 subperiods) more than half the probabilities contribute with less than 0.25 to the test statistic. Transition probabilities seem to be constant over time.

Table 3 about here

The stationarity of the distribution is also addressed by looking at the second eigenvalue of the transition matrix. For the Markov matrix at the municipal level the second eigenvalue equals 0.829, which implies that the distribution converges to a steady state. As explained in section 3, the asymptotic half life can be calculated based on the second eigenvalue. With an eigenvalue of 0.829 the time it takes to reach half way to the long run distribution equals 3.7 transition periods, which corresponds to about 22 years.

For the time stationary process in Table 1, we test the first order Markov assumption against both lower and higher order alternatives. Markovity of order 0 (previous income group does not matter) is clearly rejected (test statistics exceeding 5 000). To test Markovity of order 1 versus order 2, the sample is divided into five sub-samples depending on the municipality's income group in period *t-2*. The estimated transition matrices for each subsample are compared to the full sample matrix, and corresponding test statistics are calculated. The null hypothesis of a first order Markov process is rejected for both the Pearson and the Likelihood Ratio test, and the result indicates that the dynamics are more complicated than an ordinary first order Markov chain. However, many of the transition probabilities within the five subsamples are estimated based on very few observations and are therefore not expected to be reliable. We therefore keep the assumption of order 1 Markovity. But it is of interest to analyze further the dynamic properties of time dependent Markov chains, and the results above regarding the ergodic distributions may change with a richer dynamic formulation.

7. Migration and income growth: Dynamic agglomeration effects?

The analysis of the previous sections has shown the combination of massive migration flows to urban areas and income convergence among Norwegian municipalities and regions. To investigate possible agglomeration effects further we now concentrate on the municipalities and regions with high income levels and increasing population size. The disaggregation is in line with the approach of DiCecio and Gascon (2010) who find that US states climbing the ladder of the income distribution are gaining population. Their result is consistent with agglomeration economies where high income regions take off because of externalities and scale effects.

We start out taking a closer look at the dynamics within the top income group. Based on the transition matrix at the municipal level given in Table 1 in section 5, we divide group 5 into two equal subgroups: a lower group between 109% and 118% of the average income, and an upper group of more than 118% of the average income. The disaggregation does not reveal a small group of high income municipalities taking off. The probability of moving down the distribution from the lower group is about 45%, compared to less than 10% probability of moving up. Even municipalities in the upper group have 25% probability of moving down the distribution. Both subgroups account for a lower share of municipalities in the long run, in particular the upper group where the share decreases from 10% initially to less than 5% in the long run. A closer look at the data shows that 30 out of 396 municipalities are in income group 5 in all periods. Among these, the broad picture is not income take off, but movement towards the average. Only 3 of the 30 municipalities have higher relative income in 2008 compared to 1972, and they all belong to the oil region Stavanger/Sandnes. The rest have lower or about the same relative income level. On average, relative income among the top 30 group decreases from 1.39 to 1.25 during 1972-2008.

Furthermore we calculate population weighted density functions to check if high income populous municipalities are taking off, illustrated in Figure 3 below. At the beginning of the time period studied, the population weighted income distribution has two peaks and a long upper tail. The first peak is around the national average and the second is at relative income equal to about 1.4 (income per capita 40% higher than the average). When we compare the initial distribution to the 2008 distribution, we see that most of the municipalities have merged into a single peak with income per capita close to the average income level. The distribution becomes narrower over time with smaller differences between the low income and the high income municipalities. As can be seen, the long upper tail of the distribution has merged into a small peak of municipalities at relative income equal to 1.5.[7] This motivates our detailed study of the richest regions below.

Figure 3 about here

To analyze the role of migration for income growth and identify possible agglomeration effects, we calculate income transition probability matrices conditioned on the migration pattern. We focus on municipalities with more than 5.000 inhabitants in 2008, which gives 176 municipalities. The corresponding 6-year transition matrix is given in Table 4. The classification of income groups is updated to have uniform distribution initially: 1) less than 96% of the average income across all 396 municipalities, 2) between 96% and 102%, 3)

[7] The capital Oslo is excluded from this analysis. Oslo constitutes a separate peak at relative income equal to 1.8 in the initial 1972 distribution, but gradually moves to the left and joins the small peak around 1.5 in 2008.

between 102% and 108%, 4) between 108% and 119%, and 5) more than 119% of the average income. Consistent with the findings for the full sample of municipalities, we still find evidence of income convergence. High income municipalities are not taking off, and the transition probabilities indicate movement towards the average income level.

Table 4 about here

We then rank the 176 municipalities according to the percentage change in the population during 1972-2008, and split the sample into two subgroups. The bottom 50% represents municipalities with relatively stable development of the population size, while the top 50% is municipalities with high inmigration. Among the municipalities in the bottom 50%, the population on average increased by 3% during 1972-2008, while the municipalities with high inmigration had an average population increase of 59%. Municipalities with high inmigration are typically suburbs of large cities like Oslo, Bergen, Trondheim and Stavanger. Interestingly, most of these municipalities have weak or even negative relationship between the development in relative population and relative income during 1972-2008.

The Markov matrices for the two subgroups of municipalities with stable population and high inmigration are shown in panels a and b, respectively, of Table 5. The broad picture is that the migration pattern does not matter much for income transition probabilities. Municipalities with stable population that are in income group 2 at the beginning of a 6-year period have 18% chance of moving up the income distribution and 21% chance of falling behind. For municipalities with high inmigration the same probabilities are 24% and 20%, respectively. The probability of moving from group 4 to the top income group equals 6.3% for municipalities with high inmigration and 3.6% for municipalities with relatively stable population.

Table 5 about here

To statistically test for the importance of the migration pattern to income growth, we apply Pearson (Q) and Likelihood Ratio (LR) tests in similar ways as for tests of time stationarity. Comparing the matrix for each subsample to the matrix for the entire sample of 176 municipalities (Table 4) simultaneously results in test statistics of Q = 20.3 and LR = 21.1.[8] With 13 degrees of freedom, the 5% critical value equals 22.4. This implies that the null hypothesis of equal transition probabilities across migration patterns is not rejected at 5% significance level. The contributions to the Pearson test statistic from each transition in the two subsamples are given in Table 6.

Table 6 about here

The analysis above is inconsistent with dynamic agglomeration effects where high income municipalities take off because of inmigration. Possible scale effects cannot be large and widespread given these movements in the distribution of incomes. This is confirmed by the ergodic distribution of the matrix in panel b of Table 5. If the migration pattern generates agglomeration effects, high income municipalities with inmigration should remain in the top income group. Initially, 30% of the municipalities with high inmigration are in group 5, but instead of taking off, they tend to move towards the average income level. The estimated transition probabilities imply that in the long-run ergodic distribution, the top income group is significantly reduced, and contains less than 10% of the municipalities.

These findings are confirmed when looking at the level of the 89 NUTS-4 economic regions. Regions gaining population do not take off, but typically move towards the middle of the income distribution. When ranking regions based on the population change during 1972-2008, 25 regions have population growth above 25%. These are typically high income regions with declining relative income over time. On average, the relative income decreased from 1.13 in 1972 to 1.06 in 2008. Table 7 gives an overview of population change and relative income in the seven most populous regions in Norway. They all have initial income well above the average and high inmigration, but except for the 'oil capital' Stavanger/Sandnes, relative income is decreasing during 1972-2008.[9]

Table 7 about here

[8] When comparing the matrices, we exclude transitions with only one observation in the full sample matrix in Table 4. The degrees of freedom are adjusted accordingly.

[9] The population in the capital Oslo grew by only 18% during this period, indicating that most of the migration flows to this area is captured by the surrounding regions, like Lillestrøm, Drammen and Asker/Bærum.

The lack of agglomeration economies in our dataset goes against a large literature of agglomeration effects. The meta-analysis of Melo et al. (2009), covering 34 studies, finds strong support for agglomeration economies although the size of the effect depends on local economic conditions. They cover both urbanization economies linked to population size as studied here and localization economies related to industrial concentration. We can only conclude that agglomeration effects seem to be unimportant given the city sizes relevant in the data. The industrial composition may be important also. The urbanization in Norway is best described as service sector expansion linked to consumption growth based on oil revenues. It follows that consumption led expansion of urban services carry limited agglomeration effects. Melo et al. find that service industries derive considerably larger benefits from urban agglomeration compared to manufacturing industries.

Our result is fairly consistent with existing Scandinavian evidence of relevance. Aronsson et al. (2001) cannot reject that estimated convergence among Swedish counties is related to labor mobility. They include labor market conditions and find that initial unemployment has a positive effect on outmigration and subsequent income growth. The size of the effects is small and it is hard to interpret migration as a major determinant of convergence in their estimates. Persson (1997) also finds a positive, but weak, contribution of internal migration to convergence using an older and longer dataset. The contribution of Østbye and Westerlund (2007) is to look at gross migration and they conclude that gross migration adds to convergence among Swedish counties while it counteracts convergence among Norwegian counties. Their estimated convergence results are not all that consistent with this conclusion. Anyway, county level data have serious drawbacks in investigating migration effects. All 19 counties in Norway include both periphery with large outmigration and urban areas with inmigration.

8. Robustness checks: Age structure and municipality consolidations

We check whether changes in the age structure across municipalities can explain the lack of relationship between migration and income growth. If the population increase is driven by young people with relatively low income, municipalities with high inmigration are not likely to have income take off. Among municipalities with more than 5.000 inhabitants, the broad picture is a reduction in the share of inhabitants below 24 years, while the population shares for the age groups 25-64 years and above 65 years are increasing during 1972-2008. Among the top 50% of the sample with the large inmigration, the expansion of the age group 25-64 years is larger than for the full sample, while the increase in the population share of those above 65 years is smaller than the average. The growth of inhabitants between 25 and 64 years old is driven by the age group 35-44 years, which corresponds to the period of life when the income level is at its highest. The opposite is the case for the bottom 50% of municipalities with relatively stable population size; the increase in the age group 25-64 years is smaller than the average, while the increase of inhabitants above 65 years is larger than the average. These changes in the age structure should imply higher income growth in municipalities with high inmigration, and similar lower income growth in municipalities with large outmigration. When considering the full sample of 396 municipalities, the same pattern applies. The changes in the age structure of periphery municipalities with large outmigration are not favorable to high income growth. Consequently, the age composition across municipalities cannot explain the observed income convergence or the lack of income take off in municipalities with high inmigration.

To further check the robustness of our results, we experiment with an extended dataset. Our original data sample excludes municipalities that have been part of a municipality consolidation during 1972-2008. Since consolidations are more common around large city municipalities, we test whether this exclusion drives our results. By treating consolidated municipalities as one municipality during the whole time period, the data sample is extended from 396 to 416 municipalities. The main conclusions still hold. We find strong evidence of income convergence, while dynamic agglomeration effects are limited. Among municipalities with more than 5.000 inhabitants in 2008, high inmigration does not generate more income growth. The null hypothesis of equal transition probabilities across different migration patterns cannot be rejected at 5% significance level.

9. Concluding remarks

The large population flow from the periphery to urban centers during the past four decades is expected to stimulate urban income growth by way of agglomeration economies. The income level is higher in urban compared to rural regions and further urban expansion is assumed to increase productivity and income more. We do not find evidence in support of important agglomeration effects in this dataset of Norwegian municipalities and regions. Distribution analysis using Kernel densities and Markov chains convincingly shows income convergence.

Strong agglomeration effects should have showed up as income divergence given the population shift. The analysis of the relationship between population changes and income changes shows no systematic pattern. Inmigration in high income regions does not lead to a take-off in income growth. Urbanization in the Norwegian context, with cities in the range of 20.000 to 250.000 except for Oslo's 500.000, seems not to offer dynamic agglomeration economies.

The overall income convergence even at the municipal level reflects heterogeneity in the underlying income process. Small regions with resource based activities have had strong growth. In our dataset many small regions take benefit of incomes related to oil extraction, electric power production and salmon production. During the period under study the public sector has expanded in terms of revenue, spending and employment and with employment growth spread out with the use of central government grants to local governments. The grants have been part of regional policy and have led to high ratios of public to total employment in the periphery. The employment effect together with equalization of public sector wages across the country have been important factors in the income growth observed in rural regions.

The lack of income growth response to the large inmigration in the cities is somewhat puzzling. It seems like homogenous institutions across an open integrated market are conducive to convergence and that dynamic agglomeration effects related to migration are limited. Possibly agglomeration effects are of less importance in a setting where knowledge and technology are available for all regions. The lack of a relationship between population shift and income growth also may reflect the industrial structure. The literature indicates that manufacturing industries enjoy less agglomeration effects, but manufacturing industries have not been a major factor in the urbanization in the dataset. The urbanization in the data has rather been dominated by service sector growth. The service sector has been identified as a more important source of agglomeration economies in studies of industrial composition effects, possibly as a supplier of intermediate services to industry. The service employment expanding in Norway has been driven by the increased consumption level with economic growth based on increasing oil revenues. The lack of agglomeration effect even with service based urbanization may imply that consumption led expansion of urban services carry limited agglomeration effects. It is of interest to have a closer look at the production side of the agglomeration effects. Localization economies may be important although we have not found urbanization economies.

References

Aronsson T., J. Lundberg and M. Wikstrøm (2001), Regional income growth and net migration in Sweden, 1970-1995. *Regional Studies* 35, 9, 823-830.

Au C. and V. Henderson (2006), How migration restrictions limit agglomeration and productivity in China, *Journal of Development Economics* 80, 2, 350-388.

Baldwin, R. and P. Martin (2004), Agglomeration and regional growth, in V.J. Henderson and J.F. Thisse (eds), *Handbook of Regional and Urban Economics*, vol. 4, Amsterdam: Elsevier North-Holland.

Barro R. and X. Sala-i-Martin (1992), Convergence, *Journal of Political Economy* 100, 2, 223-251.

Beeson P., D. DeJong and W. Troesken (2001), Population growth in U.S. counties, 1840-1990, *Regional Science and Urban Economics* 31, 6, 669-699.

Bickenbach F. and E. Bode (2003), Evaluating the Markov property in studies of economic convergence. *International Regional Science Review* 26, 363-392.

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

Ciccone A. and R. Hall (1996), Productivity and the density of economic activity, *American Economic Review* 86, 1, 54-70.

Combes, P-P. (2010), The empirics of economic geography: How to draw policy implications, Document de Travail no 2010-41, GREQAM.

Combes P-P., G. Duranton and L. Gobillon (2008), Spatial wage disparities: Sorting matters! *Journal of Urban Economics* 63, 723-742.

Combes P-P., G. Duranton, L. Gobillon, D. Puga and S. Roux (2009), The productivity advantage of large cities: Distinguishing agglomeration from firm selection, CEPR Discussion Paper No. 7191.

De la Fuente A. (2002), On the sources of convergence: A close look at the Spanish regions, *European Economic Review* 46, 569-599.

DiCecio R. and C. Gascon (2010), Income convergence in the United States: A tale of migration and urbanization, *The Annals of Regional Science* 45, 2, 365-377.

Duranton G. and D. Puga (2004), Micro-foundations of urban agglomeration economies, ch. 48 in V. Henderson and F. Thisse (eds.), *Handbook of Regional and Urban Economics* 4, 2063-2117.

Glaeser E. and J. Gottlieb (2008), The economics of place-making policies, *Brookings Papers on Economic Activity*, 156-253.

Glaeser E. and J. Gottlieb (2009), The wealth of cities: Agglomeration economies and spatial equilibrium in the United States, NBER Working Paper 14806.

Kremer M., A. Onatski and J. Stock (2001), Searching for prosperity. *Carnegie-Rochester Conference Series on Public Policy* 55, 275-303.

Magrini S. (1999), The evolution of income disparities among the regions of the European Union, *Regional Science and Urban Economics* 29, 257-281.

Melo P., D. Graham and R. Noland (2009), A meta-analysis of estimates of urban agglomeration economies, *Regional Science and Urban Economics* 39, 332-342.

Moretti E. (2010), Local labor markets, NBER Working Paper 15947.

Persson J. (1997), Convergence across the Swedish counties, 1911-1993. *European Economic Review*, 41: 1835-1852.

Puga, D. (2011), The magnitude and causes of agglomeration economies, *Journal of Regional Science*, forthcoming.

Quah D. (1993a), Empirical cross-section dynamics in economic growth. *European Economic Review* 37, 426-434.

Quah D. (1993b), Galton's fallacy and tests of the convergence hypothesis. *Scandinavian Journal of Economics* 95, 427-43.

Quah D. (2001), Searching for prosperity: A comment, *Carnegie-Rochester Conference Series on Public Policy* 55, 305-319.

Rappaport J. (2005), How does labor mobility affect income convergence? *Journal of Economic Dynamics & Control* 29: 567-581.

Rattsø J. and H. Stokke (2009), Regional income convergence, skilled migration and productivity response: Explaining relative stagnation in the periphery. Working Paper No. 01/2009, Department of Economics, Norwegian University of Science and Technology.

Rattsø J. and H. Stokke (2011), The accumulation of education and regional income growth: Limited human capital effects in Norway, mimeo, Norwegian University of Science and Technology.

Redding, S. (2010), The empirics of new economic geography, *Journal of Regional Science* 50, 1, 297-311.

Shorrocks A. (1978), The measurement of mobility, *Econometrica* 46, 5, 1013-1024.

Østbye S. and O. Westerlund (2007), Is migration important for regional convergence? Comparative evidence for Norwegian and Swedish counties, 1980-2000, *Regional Studies* 41, 7, 901-915.

Figure 1: Kernel density estimates, 396 municipalities, log relative population size, 1972 vs. 2008.

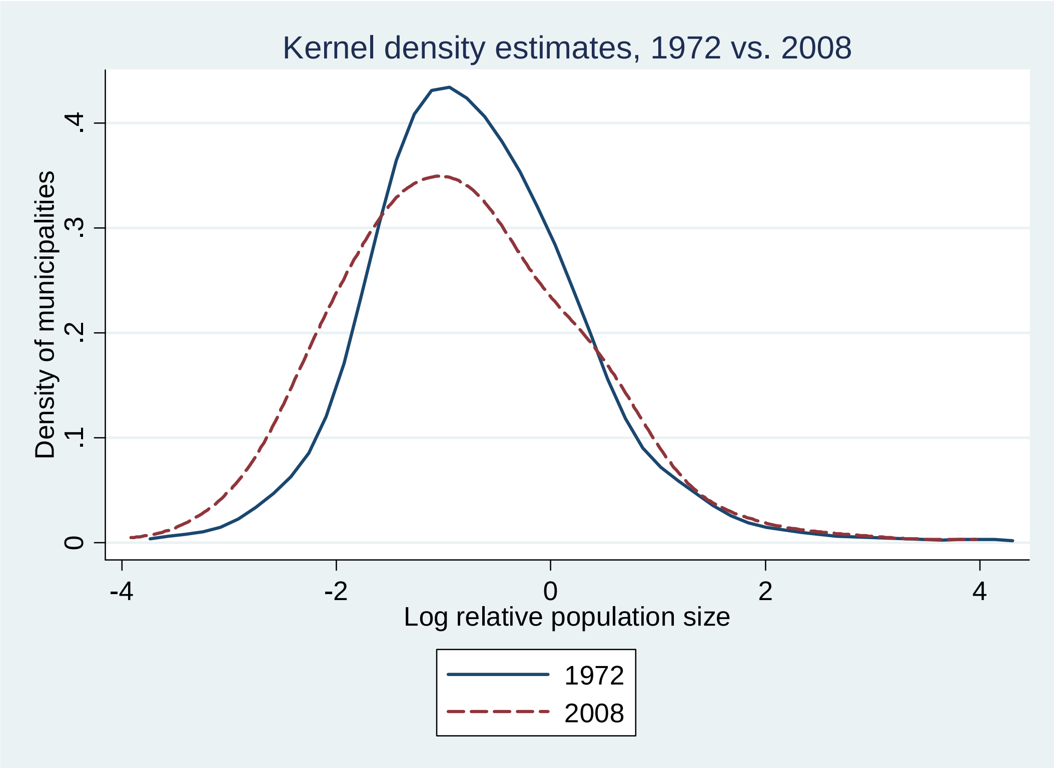

Figure 2: Kernel density estimates, 396 municipalities, relative income per capita, 1972 vs. 2008.

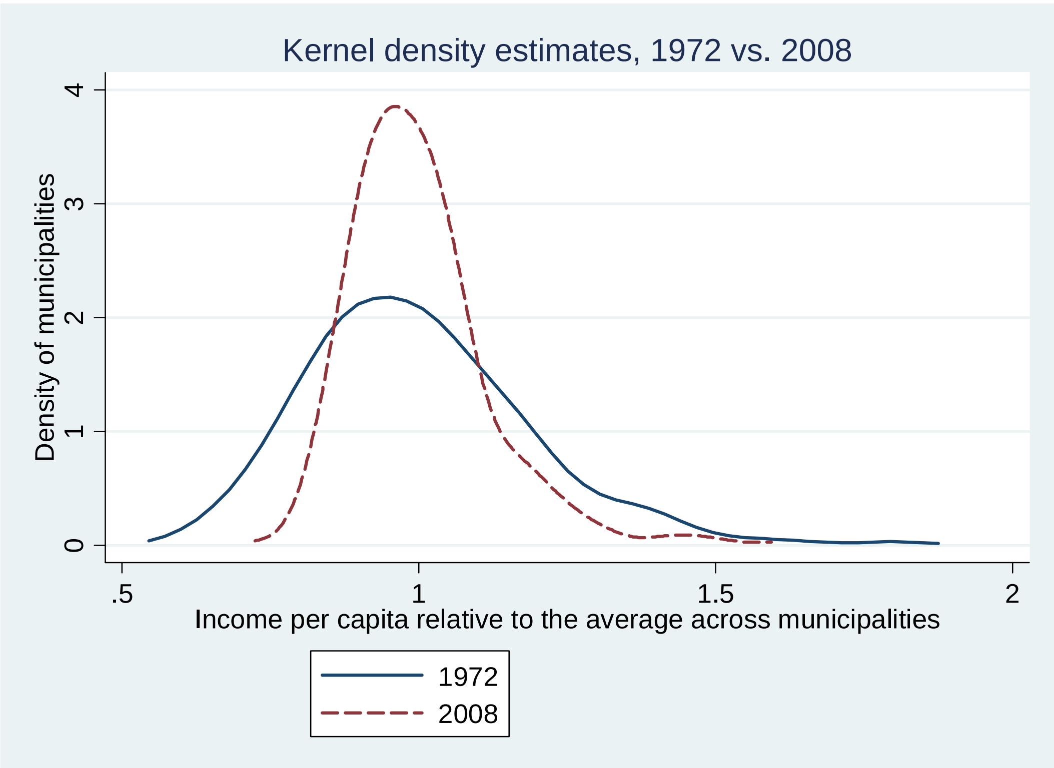

Figure 3: Population weighted Kernel density estimates, 395 municipalities (excluding the capital Oslo), relative income per capita, 1972 vs. 2008.

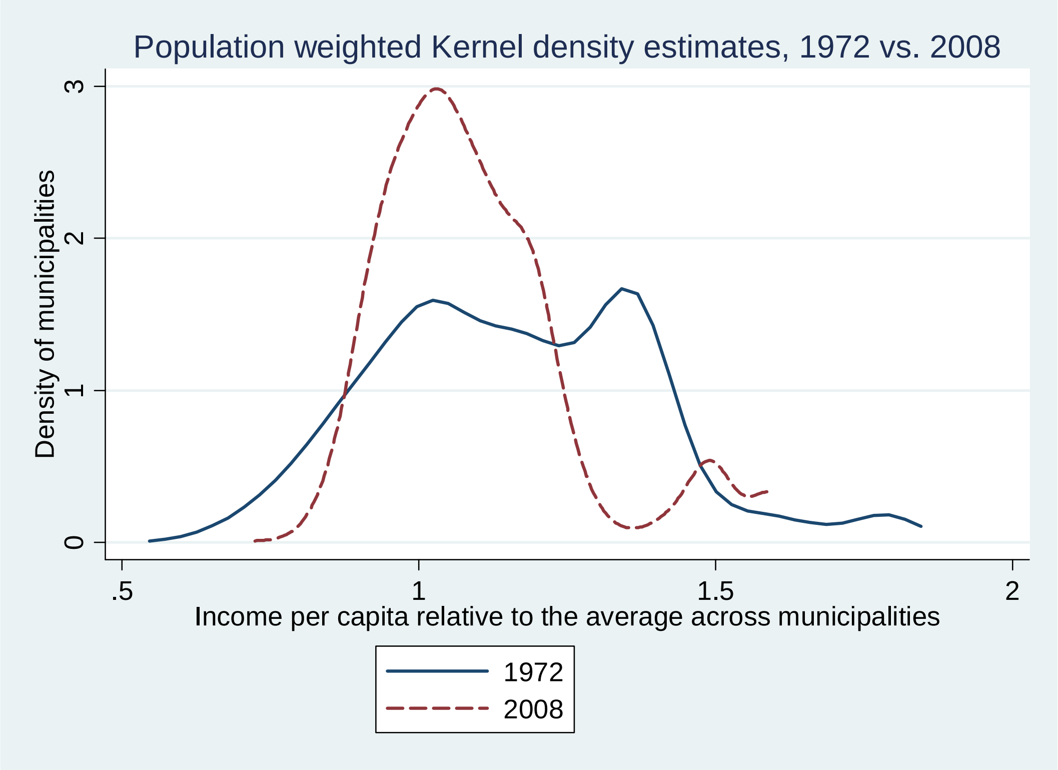

Table 1: Markov chain transition probability matrix, municipal level, 6-year transitions, 1972-2008, 2376 observations (binomial standard errors in parentheses).

Income groups	1 ≤ 0.89	2 ≤ 0.95	3 ≤ 1.01	4 ≤ 1.09	5 > 1.09	Obs.
1	**67.4** **(2.2)**	27.0 (2.0)	4.8 (1.0)	0.6 (0.4)	0.2 (0.2)	475
2	17.3 (1.7)	**54.8** **(2.3)**	22.4 (1.9)	5.1 (1.0)	0.4 (0.3)	474
3	1.7 (0.6)	22.5 (1.9)	**52.7** **(2.3)**	20.2 (1.8)	2.9 (0.8)	476
4		4.2 (0.9)	25.5 (2.0)	**58.7** **(2.3)**	11.6 (1.5)	475
5		0.4 (0.3)	1.7 (0.6)	22.3 (1.9)	**75.6** **(2.0)**	476
Initial distribution	20.0	20.0	20.0	20.0	20.0	
Ergodic distribution	13.5	23.0	25.4	23.4	14.7	

Table 2: Markov chain transition probability matrix, regional level, 6-year transitions, 1972-2008, 534 observations (binomial standard errors in parentheses).

Income groups	1 ≤ 0.9	2 ≤ 0.96	3 ≤ 1.01	4 ≤ 1.07	5 > 1.07	Obs.
1	**70.1** **(4.4)**	29.0 (4.4)		0.9 (0.9)		107
2	17.8 (3.7)	**58.9** **(4.8)**	21.5 (4.0)	1.9 (1.3)		107
3		19.8 (3.9)	**55.7** **(4.8)**	22.6 (4.1)	1.9 (1.3)	106
4		0.9 (0.9)	25.2 (4.2)	**62.6** **(4.7)**	11.2 (3.0)	107
5			1.9 (1.3)	20.6 (3.9)	**77.6** **(4.0)**	107
Initial distribution	20.0	20.0	20.0	20.0	20.0	
Ergodic distribution	13.1	22.0	25.4	24.9	14.6	

Table 3: Test of time stationarity, municipal level, 6-year transitions. Contributions of the subperiods 1972-1990 and 1990-2008 to the Pearson test statistics.

Subperiods	Income groups	Number of obs.	1 ≤ 0.89	2 ≤ 0.95	3 ≤ 1.01	4 ≤ 1.09	5 > 1.09	Sum
1972-1990	1	302	0.03	0.16	0.02	0.00	0.63	0.84
	2	200	0.07	0.30	1.19	0.00	0.03	1.59
	3	194	0.17	2.11	1.58	0.12	0.51	4.49
	4	203		3.59	0.77	0.19	1.29	5.84
	5	289		0.50	0.15	0.00	0.00	0.65
	Sum							**13.41**
1990-2008	1	173	0.05	0.28	0.05	0.01	1.13	1.52
	2	274	0.05	0.22	0.86	0.00	0.02	1.15
	3	282	0.11	1.46	1.08	0.08	0.36	3.09
	4	272		2.69	0.57	0.14	0.96	4.36
	5	187		0.79	0.24	0.00	0.00	1.03
	Sum							**11.15**
Pearson test statistic								**24.6**
Critical value at 5% significance level								28.9

Table 4: Markov chain transition probability matrix, 6-year transitions, 1972-2008, municipalities with more than 5000 inhabitants in 2008, 1056 observations (binomial standard errors in parentheses).

Income groups	1 ≤ 0.96	2 ≤ 1.02	3 ≤ 1.08	4 ≤ 1.19	5 > 1.19	Obs.
1	**81.0**	16.6	1.4	0.5	0.5	211
	(2.7)	(2.6)	(0.8)	(0.5)	(0.5)	
2	20.4	**59.7**	19.4	0.5		211
	(2.8)	**(3.4)**	(2.7)	(0.5)		
3	1.9	29.7	**51.4**	16.0	0.9	212
	(0.9)	(3.1)	**(3.4)**	(2.5)	(0.6)	
4	1.0	3.8	34.6	**55.4**	5.2	211
	(0.7)	(1.3)	(3.3)	**(3.4)**	(1.5)	
5			0.5	22.3	**77.2**	211
			(0.5)	(2.9)	**(2.9)**	
Initial distribution	20.0	20.0	20.0	20.0	20.0	
Ergodic distribution	35.3	30.5	20.4	9.9	3.9	

Table 5: Markov chain transition probability matrix according to the migration pattern during 1972-2008, 6-year transitions (binomial standard errors in parentheses).

Panel a: Municipalities with relatively stable population (528 observations).

Income groups	1 ≤ 0.96	2 ≤ 1.02	3 ≤ 1.08	4 ≤ 1.19	5 > 1.19	Obs.
1	**83.4**	15.2	0.7		0.7	145
	(3.1)	(3.0)	(0.7)		(0.7)	
2	20.9	**61.1**	18.0			139
	(3.4)	**(4.1)**	(3.3)			
3	2.8	37.6	**49.5**	10.1		109
	(1.6)	(4.6)	**(4.8)**	(2.9)		
4	1.2	3.6	35.7	**55.9**	3.6	84
	(1.2)	(2.0)	(5.2)	**(5.4)**	(2.0)	
5			1.9	31.4	**66.7**	51
			(1.9)	(6.5)	**(6.6)**	
Initial distribution	27.5	26.3	20.6	15.9	9.7	
Ergodic distribution	44.9	33.3	15.8	4.6	1.4	

Panel b: Municipalities with high inmigration (528 observations)

Income groups	1 ≤ 0.96	2 ≤ 1.02	3 ≤ 1.08	4 ≤ 1.19	5 > 1.19	Obs.
1	**75.8**	19.7	3.0	1.5		66
	(5.3)	(4.9)	(2.1)	(1.5)		
2	19.5	**56.9**	22.2	1.4		72
	(4.7)	**(5.8)**	(4.9)	(1.4)		
3	1.0	21.4	**53.4**	22.3	1.9	103
	(1.0)	(4.0)	**(4.9)**	(4.1)	(1.3)	
4	0.8	3.9	33.9	**55.1**	6.3	127
	(0.8)	(1.7)	(4.2)	**(4.4)**	(2.2)	
5				19.4	**80.6**	160
				(3.1)	**(3.1)**	
Initial distribution	12.5	13.6	19.5	24.1	30.3	
Ergodic distribution	21.5	24.7	26.6	18.5	8.7	

Table 6: Test of whether the migration pattern affects transition probabilities, 176 municipalities 1972-2008, 6-year transitions. Contributions of single subsamples to the Pearson test statistics.

Migration pattern	Income groups	Obs.	1 ≤ 0.96	2 ≤ 1.02	3 ≤ 1.08	4 ≤ 1.19	5 > 1.19	Sum
Relatively stable population	1	145	0.10	0.18	0.54			0.82
	2	139	0.02	0.05	0.15			0.22
	3	109	0.43	2.28	0.07	2.41	1.02	6.21
	4	84	0.05	0.01	0.03	0.00	0.43	0.52
	5	51				1.86	0.72	2.58
	Sum							**10.35**
High inmigration	1	66	0.22	0.38	1.19			1.79
	2	72	0.03	0.09	0.28			0.40
	3	103	0.46	2.42	0.08	2.54	1.10	6.60
	4	127	0.03	0.01	0.02	0.00	0.29	0.35
	5	160				0.60	0.24	0.84
	Sum							**9.98**
Pearson test statistic								**20.3**
Critical value at 5% significance level								22.4

Table 7: Population change and development in relative income level in the seven most populous regions in Norway.

Region	Population 1972	Population 2008	Population change	Relative income 1972	Relative income 2008
Oslo	475 563	560 484	17.9 %	1.69	1.40
Bergen	297 275	378 818	27.4 %	1.17	1.12
Stavanger/Sandnes	149 358	244 118	63.4 %	1.18	1.33
Trondheim	170 783	221 058	29.4 %	1.17	1.05
Lillestrøm	118 859	185 471	56.0 %	1.28	1.13
Drammen	126 710	164 542	29.9 %	1.23	1.10
Asker/Bærum	111 393	161 066	44.6 %	1.62	1.53