

Income convergence, migration and geography: Distribution analysis of regions 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

Income convergence among regions and countries and the role of migration depends on economic and institutional conditions. We investigate movements in the income distribution among regions in an integrated market with high mobility of labor, capital, knowledge and technology – municipalities and economic regions in Norway. Large and persistent population flows from the periphery towards urban centers characterize the economic development. The imbalance between periphery and cities indicate the possibility of income divergence. Kernel density functions and finite first order Markov chains are estimated and tested, in particular with respect to the role of migration. The results show convergence towards a narrower unimodal distribution of per capita income in this setting. Interestingly, migration does not seem to be an important part of the convergence process. Catching up of low income regions is independent of migration. Immigration in high income regions has a weak positive effect on income growth, and this holds back convergence. But overall the results are hardly consistent with strong agglomeration and scale effects and high income regions taking off. The convergence process is slow, and the evolving regional income pattern seems to reflect differences in permanent regional factors.

JEL codes: C26, D92, F2, J61, O15, O18, R11, R23

Key words: Convergence, migration, divergence, agglomeration, Kernel density functions, Markov chains

First draft, January 28, 2010

* The project is funded by the Norwegian Research Council. We appreciate data handling by Jo Jakobsen.

1. Introduction

The controversy over income convergence can be clarified by studying conditions and mechanisms conducive to convergence. We study a setting with regions in an integrated market with institutional homogeneity and high mobility of labor, capital, knowledge and technology – municipalities and economic regions in Norway. They experience the same geographic concentration of population density and economic activity as in other countries, and casual observation indicates divergence – urban areas expand while the periphery is weakened.

The standard neoclassical growth model predicts income convergence based on capital flows. It has been assumed that labor outmigration will speed up income convergence by tightening labor markets in low income regions. But closer examination of the model finds that migration has an ambiguous effect on income convergence. The increased marginal product of labor following outmigration is counteracted by reduced marginal product of capital when there are complementarities between the factors of production (Rappaport, 2004, 2005). Another literature arguing for income convergence effect of migration concentrates on the migration equilibrium of households. Households move as a response to income differences and the migration is expected to equalize income (utility) across regions. Glaeser and Gottlieb (2009) integrate migration equilibrium in an analysis of US cities.

In the neoclassical growth model migration can induce stronger productivity effects with heterogeneous labor and associated skill effects and brain drain/gain (Rattsø and Stokke, 2009). Also the receiving end of regions with immigration may experience more than declining marginal productivity of labor. Glaeser and Gottlieb (2009) combine migration and agglomeration and Baldwin and Martin (2004) summarize a large literature emphasizing agglomeration effects generating economic growth. This competing theoretical approach of income divergence predicts more dramatic consequences of migration. Reichlin and Rustichini (1998) introduce labor scale effects in a migration growth model and obtain strong divergence. Regions with outmigration experience loss of productivity, while receiving regions gain productivity. Income convergence may or may not speed up with migration.

Barro and Sala-i-Martin (1992a) started up a large literature with econometric analysis of 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 dynamic panel analysis and offers a study of Spanish regions. He concludes that panel analyses of regional data indicates income convergence and with a higher convergence rate than the early static analyses. The econometric approach faces serious challenges of endogeneity, and the literature has moved to alternative methods and investigation of channels of effects.

The econometric literature on the effects of migration for growth also is inconclusive. The early regression analyses based on the neoclassical model, notably Barro and Sala-i-Martin (1992b), conclude that migration does not account for a sizable effect of income convergence among regions within countries. Recent studies applying the regression method and using regional data give some support to the agglomeration models and endogenous growth. Ozgen, Nijkamp and Poot (2009) make a meta-analysis of econometric analyses and conclude that the migration effect is more consistent with endogenous self-enforcing growth. Fingleton and Fischer (2008) explicitly test the neoclassical model against new economic geography and find more support for the latter using data for European regions.

The distribution analysis developed by Quah (1993a, b) has given more support to income divergence. Magrini (1999) finds that EU regions are characterized by a tendency towards divergence, in particular because of the high growth of high income regions. The importance of migration for the income process has been analyzed with this method in a study 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. Their results suggest that strong divergence is expected among Norwegian regions experiencing large population flows.

The possible linkage between migration and convergence/divergence primarily affects the performance at the ends of the income distribution – the relative low income and the relative high income regions. Relative low income regions with net outmigration may gain per capita income from reduced labor supply or loose from associated capital outflow. Relative high

income regions with net immigration experience the opposite adjustments. Heterogeneous labor and scale economies may add stronger effects of population flows. Brain drain and gain and agglomeration may strengthen the forces of divergence. To capture the dynamics of the ends of the income distribution we work with distribution analysis and Markov chains. We estimate and test the dynamic characteristics of the regional income levels and compare with existing studies under similar and different conditions.

We study regional convergence in Norway based on data on income per capita during 1972-2003. The results indicate income convergence with a movement towards unimodal distribution of per capita income. The analysis is made both at the level of 401 municipalities and 89 economic regions defined by the European Union classification (NUTS-4). Income convergence may reflect the fact that available knowledge and technology can be taken into use all over Norway. Interestingly, migration seems not to be important for the convergence process. Catching up of low income regions is independent of migration. Immigration in high income regions has a weak positive effect on income growth, and this holds back convergence. High income regions moving up the income ladder have somewhat higher immigration compared to high income regions moving down. But overall the results are hardly consistent with strong agglomeration and scale effects and high income regions taking off. The convergence process is slow, and there is a distinct geographic pattern in the income development. The geographic variation is consistent with a story where fixed structures (natural resources, industrial structure, infrastructure and else) influence growth conditions.

Section 2 discusses methodological challenges with distribution analysis and the Markov method. The basic Kernel distribution densities and Markov chains documenting income convergence are shown in section 3. Section 4 offers tests of time stationarity and time independence. Migration and convergence are analyzed in section 5. Section 6 considers spatial patterns of convergence. Concluding remarks are given in section 7.

2. Methodological issues

Our main motivation to use Markov chains is the importance of the two ends of the distribution of per capita incomes – relative low income and relative high income regions. Income convergence and divergence are heterogeneous processes with different growth paths from different starting points. In addition the challenges of the econometric methods are

serious, and in particular with endogenous migration. We take benefit of the methodological innovations of Quah (1993a, 1993b, 2001) using Markov chains, more recently developed 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. Our discussion of the method is related to Bichenbach and Bode (2003).

The analysis is based on data for taxable income of each of 401 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 1972-2003 and it follows that we have 12832 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.

The database covers per capita income Y_{kt} for municipality k in year t , 401 municipalities ($k = 1, \dots, 401$) for the period 1972 to 2003 ($t=1972, \dots, 2003$). In general terms we describe these as regions. In the analysis the income level is measured relative to the average income per capita across regions in each year, and we describe this relative per capita income by y_{kt} . The whole range of relative per capita income is divided into a finite number of N mutually exclusive states and in this analysis we follow the convention of working with quintiles ($N=5$). For each region k we get a sequence of variables describing the states of that region at time t . The sequences are considered as independent realizations of a single homogeneous Markov chain with finite state space N . The assumption of a finite first order Markov chain implies that the probability of being in a specific quintile at time t only depends on the quintile of the previous period (and not earlier periods). The transition probability, the probability of moving from quintile i to quintile 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 quintiles over time. The number of regions moving from quintile i to quintile j from period $t-1$ to t is measured by

$n_{ij}(t)$. The total number of regions moving from quintile 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 for each row i and time period t of the transition matrix and the distribution function is:

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

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

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

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:

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

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 quintiles, $h(0) = [h_1(0), h_2(0), h_3(0), h_4(0), h_5(0)]$ 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 m transition periods follows as $h(m) = h(0)\Omega^m$. Given that the matrix is regular¹, the distribution converges to the limiting distribution $h^* = \lim_{m \rightarrow \infty} h(0)\Omega^m$, 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.

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

To check whether the calculated estimates work as estimates of Markov transition probabilities, we test two key properties: time stationarity and 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:

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

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

A_i is the set of nonzero transition probabilities in the i th row of the transition matrix estimated from the entire sample, while $A_{i|m}$ is the set of nonzero transition probabilities in the i th row of the matrix estimated from the m th subperiod. The total number of transitions from quintile i in subperiod m and the total number of transitions from quintile i to quintile 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 i th 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-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):

$$hl = \frac{-\log 2}{\log |\lambda_2|} \quad (6)$$

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

quintile j at time t is influenced both by quintile i at time $t-1$ and quintile 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:

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

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

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 quintile 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 quintile i given quintile h in the previous period and the total number of transitions from quintile i to quintile j given quintile 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.

3. Kernel densities and Markov chains show convergence

The analysis of income convergence applies Kernel density functions as well as Markov chain transition matrices. The most intensive use of the data estimates Markov chains using annual transitions, and this replicates the transition probability matrices suggested by Quah (1993a, b) for studies of cross-country income dynamics. We have investigated annual transitions, but report only full-period transition and 4-year transitions in Table 1 below. The pattern is the same, and the argument for 4-year intervals is to avoid short term fluctuations and thereby have more stable transition paths. We apply the average income per capita over each 4-year period rather than the income level in the first year of each period, and thereby also avoid single year outliers.

As a simple start we examine how the distribution of relative municipal income per capita develops over time. Figure 1 shows the estimated Kernel density functions for the first 4-year period 1972-75 and the last 4-year period 2000-03². 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 1972-75, the distribution is narrower and much more concentrated around the peak in 2000-03.

Figure 1 about here

To estimate Markov chain matrices we follow the convention of discretization based on quintiles in the first period and reach about 80 municipalities in each of the following income quintiles: 1) less than 82.7% of the average, 2) between 82.7% and 92.4%, 3) between 92.4% and 101%, 4) between 101% and 113.5%, and 5) more than 113.5% of the average across municipalities. The number of municipalities in the five quintiles in the last 4-year period shows a sharp movement towards the middle income quintiles, consistent with the findings in Figure 1. In 2000-03 half the high income quintile of the initial period (40 municipalities) has per capita income above 113.5% of the average. At the bottom end only 11 (out of 81) municipalities have per capita income below 82.7% in the last period. The municipalities have moved towards the middle income quintiles and now 130 municipalities have per capita income between 92.4 and 101%. Given this pattern from beginning to end it will be hard to avoid convergence in this dataset.

Table 1 about here

The Markov matrices show a pattern of income convergence across municipalities. The distribution of per capita incomes is tending towards a point mass, rather than towards a two-point distribution. A municipality that is in the relative low income quintile initially (income level relative to the average below 0.827) has a high probability of transition to a higher

² 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 0.059 and 0.04 for 1972-75 and 2000-03, respectively.

income quintile during the 28 year period from 1972-75 to 2000-03. Using the full period transition in panel a, the chance of still being in the lower end of the distribution in the end period is only 14%. There is 37.5% probability that initially high income municipalities (relative income above 1.135) remain at the top quintile. For income quintiles 2 and 4 the probability of either remaining in that quintile or moving to the middle income quintile 3 equals about 85%. In this consideration of the full period 1972-75 to 2000-03, there are no transitions from income quintile 2 to the lowest income quintile, and the probability of moving from quintile 4 to the high income quintile is below 10%. The probability of middle-income municipalities to remain in the mid-quintile is almost 60%. In other words, the distribution dynamics show no tendencies of a bimodal twin peaked distribution.

The 4-year transition probability matrix for the period 1972-75 to 2000-03 is shown in panel b. The calculation is based on 2807 observations (compared to 401 observations for the full-period transitions), which gives more reliable results. As expected, the probability of remaining in the same income quintile is higher when the transition period is shorter (4-year transitions rather than 28-year transition). The overall picture is still one of convergence. Municipalities in the low income quintile 1 have about 40% probability of catching-up during a 4-year period, and the high income municipalities have 20% chance of moving down the distribution. The probability of moving from quintile 4 to the high income quintile is only 4.8%. Similarly, the chance that municipalities in income quintile 2 move down the distribution is 3.7%.

These patterns are confirmed by the implied ergodic distributions, which are given in the last row of the matrices. The ergodic distribution represents the long-run distribution of municipal incomes. In the full period matrix, the lowest income quintile is practically disappearing in the long-run, while the middle-income quintile accounts for 45% of the municipalities. 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 tend to become richer and high income municipalities tend to become poorer (relatively speaking), i.e. living standards converge across municipalities. The ergodic distribution is independent of the length of transitions. Whether the expected long-run distribution of municipal incomes is based on the full period 28-year transition matrix or the 4-year transition matrix for the same period does not matter much. In both cases, municipal incomes seem to be normally

distributed with thinning of the tails and the majority of municipalities gathering in the middle.

Given the highly variable population sizes at the municipal level, an obvious concern with the results above is that convergence is driven by many small municipalities, while there is divergence when we take the size of the population into account. More than half the municipalities have less than 5 000 inhabitants in 2003, and about $\frac{3}{4}$ have less than the average level of 10 000 inhabitants. Five municipalities represent outliers with population sizes above 100 000, with the capital Oslo as the largest (about 500 000 inhabitants in the municipality of Oslo). To check the robustness of the convergence pattern we calculate population weighted transition matrices and density functions.

The population weighted Markov matrix is shown in Table 2. The estimated probabilities are based on 4-year transitions and can be compared to the matrix in panel b of Table 1. Controlling for population size affects the upper end of the distribution, but does not change transition probabilities very much. The main difference is that initially high income municipalities have a higher chance of remaining high income (the probability increases from 80% to 92.4%). This is driven by many large municipalities in the high income quintile. The calculated ergodic distribution still implies income convergence with a single peak, but now located at a somewhat higher income level (quintile 4 with relative income between 1.01 and 1.135, rather than quintile 3).

Table 2 about here

The estimated Kernel density functions when municipalities are weighted by the population level are shown in Figure 2. The population weighting clearly affects the income distribution. At the beginning of the time period studied (1972-75) the population weighted income distribution has three peaks; the first around the national average, the second at relative income equal to about 1.35 (income per capita 35% higher than the average), and the third at relative income of 1.75. The second peak in Figure 2 is reflected in the Markov matrix in Table 2 by the larger share of the municipalities located in the high income quintile in the long-run (increases from 8.4% to 19.3% when the population size is taken into account).

Figure 3 excludes Oslo and shows that the third peak basically is driven by the populous capital. When we compare the 1972-75 distribution to the 2000-03 distribution, we see that the first two peaks (with relative income of 1 and 1.35) have merged into a single peak where income per capita is close to the average income level (around 1-1.1). The distribution becomes narrower over time with smaller differences between the low income and the high income municipalities. There is still a trace of some large rich municipalities at the upper end of the distribution that are not part of the overall income convergence. The result is different from the bimodal distributions analyzed in the literature discussed as the result of a poverty trap.

Figures 2 and 3 about here

Convergence at the municipal level represents a regional classification that is arbitrary with respect to economic conditions. Municipal borders cut through economic regions. There exists an aggregation of municipalities to economic regions that are defined by common labor markets (NUTS-4). Based on regional income per capita relative to the average across regions we separate between five income quintiles defined according to the quintiles in 1972-75: 1) less than 90% of the average income per capita, 2) between 90% and 101%, 3) between 101% and 106.7%, 4) between 106.7% and 120%, 5) more than 120% of the average income per capita across regions. Panels a and b in Table 3 show the Markov matrices based on the full period transition from 1972-75 to 2000-03 and the 4-year transitions for the same period, respectively. Due to few observations when considering the 28-year transition (89 regions), we focus on the 4-year transition matrix.

Table 3 about here

The results are consistent with the findings at the municipal level with income convergence and a single peaked distribution in the long-run. The ergodic distribution is skewed to the right with 53.5% of the regions ending up in income quintile 2 (between 90% and 101% of the average income per capita). This corresponds well with the peak at income quintile 3 at the municipal level (between 92.4% and 101% of the average). We conclude that both at the municipal and the economic region level the time series data show income convergence.

Previous research of income convergence in the Scandinavian countries has estimated convergence in standard fashion and used county level data. Aronsson et al. (2001) find convergence among Swedish counties, but Østbye and Westerlund (2007) estimate negligible convergence for Sweden and Norway (using different model specifications). Compared to our analysis, the small number of counties represents much less information about income distribution and a classification of regions that do not capture the differences between periphery and urban centers and functional regions very well.

4. Tests of Markov properties: Time stationarity and independence

We investigate the Markov properties of the 4-year transition matrix at the municipal level, given in panel b of Table 1. The transition probabilities are estimated based on seven 4-year transitions for 401 municipalities during 1972/75-2000/03 (2807 observations). 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 2. The sample period is divided into four subperiods, each containing two 4-year transitions, except the last subperiod which has one transition. The transition matrices for each subperiod are then compared to the full period matrix. The null hypothesis of time stationarity cannot be rejected at the 5% significance level by the Pearson and the Likelihood Ratio tests. The LR test also holds at the 10% level. The test statistics are $Q = 45.31$ (prob = 0.07) and $LR = 42.75$ (prob = 0.12). The tests have 33 degrees of freedom. We have also tested for time stationarity by dividing the sample period into two subperiods. The main conclusion of constant transition probabilities still holds.

The detailed contributions to the Pearson test statistic from each transition in the four subperiods are shown in Table 4. The differences in transition probabilities across subperiods are minor. Of the 64 comparisons (16 probabilities, 4 subperiods) more than half the probabilities contribute with less than 0.25 to the test statistic. As much as 43 contribute with less than 0.5. It is evident from Table 4 that the relatively high value of the test statistic is driven by a few large contributions. For instance, in the third subperiod (1988/91-1996/99) the probability of going from quintile 1 to quintile 3 during a 4-year period is much higher than for the entire sample period (4.35 vs. 0.4), which results in a test statistic contribution of 8.22. However, this is due to a single observation (out of 2807 observations) from quintile 1 to quintile 3 in this subperiod, while the other periods have no transitions of this kind. Excluding

this outlier observation lowers the Pearson test statistic (and reduces the degrees of freedom to 30) and the null hypothesis holds at 20% significance level.³

Table 4 about here

The stationarity of the distribution also is addressed by looking at the second eigenvalue of the transition matrix. For the 4-year Markov matrix at the municipal level the second eigenvalue equals 0.856, which implies that the distribution converges to a steady state. As explained in section 2, the asymptotic half life can be calculated based on the second eigenvalue. With an eigenvalue of 0.856 the time it takes to reach half way to the long run distribution equals 4.45 transition periods, which corresponds to about 18 years.

For the time stationary process in panel b of Table 1, we test the first order Markov assumption against both lower and higher order alternatives. Markovity of order 0 (previous quintile does not matter) is clearly rejected (test statistics exceeding 4 000). To test Markovity of order 1 versus order 2, the sample is divided into five sub-samples depending on the municipality's income quintile in period $t-2$. The estimated transition matrices for each subsample are compared to the full sample matrix, and corresponding test statistics are calculated, as documented in Table 5. The null hypothesis of a first order Markov process is rejected for both the Pearson and the Likelihood Ratio test. The calculated test statistics are $Q = 161.7$ and $LR = 156.8$, while the 5% critical value equals 37.65 (with 25 degrees of freedom). The result indicates that the dynamics is more complicated than an ordinary First-order Markov chain. It is of interest to analyze further the dynamic properties of time dependent Markov chains, and the results above about the ergodic distributions may change with a richer dynamic formulation.

Table 5 about here

³ Similarly, there are in total six transitions from quintile 2 to quintile 4, giving a transition probability of 0.9 for the full sample period. Four of these transitions occur during 1972/75-1980/83. The estimated probability is therefore much higher for this subperiod, reflected in a test statistic contribution of 3.43 [given in cell (2, 4) of the top matrix of Table 4].

5. Migration and convergence

The per capita income distribution among regions investigated here must be understood at the background of massive and persistent migration from the periphery to urban areas. As discussed in the introduction, migration has an ambiguous effect on income convergence within the neoclassical growth framework, while the recent literature on economic geography and agglomeration predicts income divergence as a result of migration. The high income regions may take off because of scale effects strengthening their economic expansion. We address the role of migration for convergence/divergence by looking at relationships between transitions in the income distribution and the migration pattern. Migration flows are measured as the percentage change in the population size of municipalities during 1972-2003. Since the demographics of fertility and mortality are similar across Norway, the changes in population size primarily reflect migration. The aggregate population increases by 16% during this period, representing an annual population growth rate of about 0.5%. The municipalities have dramatic different development of population size. The unweighted average increase in population among the 401 municipalities is 8.7%, and the standard deviation is as high as 30.7%. Almost half the municipalities had a decrease in the absolute population size over the 31-year period. Table 6 shows how this measure of migration varies across different income transitions.⁴ The calculation is based on the 4-year transition matrix given in panel b of Table 1, which means that each municipality has seven income transitions, giving a total of 2807 observations.

If migration from low income to high income municipalities explains the convergence pattern found in the data, municipalities with outmigration are relatively more likely to catch-up, while municipalities with immigration are relatively more likely to fall behind. This would be the case if the average change in the population size is relatively lower in municipalities catching up. On the other hand, if the migration pattern contributes to income divergence, the population change is higher in municipalities moving up the distribution than in municipalities falling behind. As seen from the last column of Table 6, the population development varies significantly between income quintiles (vertical variation). The average change in the population size clearly increases with the income level. Outmigration is more common in low income municipalities, and similarly immigration more common in high

⁴ Blank cells indicate either no transitions or less than 10 transitions in total in the respective income mobility group.

income municipalities. We also observe variations in the population change between movements up and down within a given quintile (horizontal variation), in particular in the upper end of the distribution. Municipalities catching-up from quintile 4 to the high income quintile on average have an increase in the population size of 27.8%, while municipalities moving down the distribution from quintile 4 have an increase of 12.6%. These findings indicate that migration into high income regions holds back the convergence process. However, among municipalities starting in the lowest income quintile at the beginning of a 4-year period, the degree of outmigration is about the same whether they catch-up or not (average population decline of 14.3% vs. 18.3%). Strong outmigration in low income municipalities does not affect catching up.

Table 6 about here

To analyze the role of migration in the convergence process we calculate income transition probability matrices for subgroups of the municipalities based on the migration pattern. The municipalities are ranked according to the percentage change in the population during 1972-2003. We split the sample of 401 municipalities in three. The 25% of the sample with the largest population decline is defined as municipalities with strong outmigration, and similarly, the 25% with the largest increase in the population size is defined as municipalities with strong immigration. The mid 50% represents municipalities with more stable development of the population size. On average (unweighted), the population in the 100 municipalities with strong outmigration decreased by 24% during 1972-2003, while the 100 municipalities with strong immigration had an increase in the population size of 51%. The matrices are shown in Table 7 below. The number of observations in different income quintiles reflects that outmigration is dominated by municipalities in the lower half of the distribution, while immigration is most common in the upper half of the distribution.

Panel a shows that among low income municipalities with strong outmigration the probability of moving up the income ladder equals 35.4% (34.8% + 0.6%). For the sample as a whole, the probability of catching-up from the lowest income quintile equals 40.5% [given as the sum of cells (1, 2) and (1, 3) in panel b of Table 1]. Similar, municipalities with strong outmigration that are in income quintile 2 at the beginning of a 4-year period have 16.5% (16.1% + 0.4%) chance of catching-up, compared to 22.5% among all municipalities [given as the sum of cells (2, 3) and (2, 4) in panel b of Table 1]. In other words, the fact that a low income municipality

experiences strong outmigration does not affect the chances of catching up very much. The transition probabilities in municipalities with high immigration given in panel c of Table 7 are not very different. With immigration the probability of remaining in the high income quintile increases from 80.1% [given in cell (5, 5) in panel b of Table 1] to 83.8%. Similar, for municipalities with strong immigration that are in income quintile 4 at the beginning of a 4-year period the probability of catching-up equals 7%, compared to 4.8% for the full sample [given in cell (4, 5) in panel b of Table 1].

Table 7 about here

Based on the estimated transition probabilities, migration from low income to high income municipalities cannot explain the income convergence seen in the data. If anything, the migration pattern acts to hold back the convergence process, but as seen from the implied ergodic distributions of the three submatrices, the magnitude of this effect is limited. If the migration pattern contributes to divergence, low income municipalities with strong outmigration should remain in the lowest income quintile, while high income municipalities with strong immigration should take off and increase the income gap. The opposite is happening. Among municipalities with strong outmigration, 23% is initially in the lowest income quintile, but instead of being stuck in a poverty trap, they are able to catch-up. The estimated transition probabilities imply that in the long-run ergodic distribution, the lowest income quintile is significantly reduced, and contains 5% of the municipalities. Similarly, among municipalities with strong immigration, the top income quintile is not taking off, but is decreasing from 35.3% of the municipalities initially to 18% in the long run.

To statistically test for the importance of the migration pattern to the convergence process, we apply Pearson (Q) and Likelihood Ratio (LR) tests in similar ways as for tests of time stationarity. The sample of 401 municipalities is divided into three subsamples according to the migration pattern, as defined above. Comparing the matrix for each of these subsamples to the matrix for the entire sample simultaneously results in test statistics of $Q = 51.2$ and $LR = 52.4$. With 22 degrees of freedom, the 5% and 1% critical values equal 33.9 and 40.3, respectively. This implies that the null hypothesis of equal transition probabilities across migration patterns is rejected.

Table 8 about here

To investigate this rejection further we show the contributions to the Pearson test statistic from each transition in the three subsamples in Table 8. Almost half of the 48 comparisons (16 probabilities, 3 subsamples) contribute with less than 0.5 to the test statistics and 2/3 contributes with less than 1. Some large error terms among the final 1/3 of the probabilities are the main source of the large test statistic leading to rejection. Among municipalities with strong outmigration it is mainly the transition probabilities from income quintile 2 that deviate from the full sample probabilities.⁵ The matrix based on municipalities with strong immigration, deviates the most from the full sample matrix, in particular due to the probabilities from income quintiles 2 and 4.⁶ It follows that municipalities at the ends of the distribution have similar transition probabilities and are not much affected of whether they have large outmigration or large immigration.

The analysis above is inconsistent with a divergence model where low income municipalities are left behind because of outmigration and high income municipalities take off because of immigration. Possible scale effects cannot be large and widespread given these movements in the distribution of incomes. We have done the same analysis at the level of the 89 NUTS-4 economic regions, and the main conclusions still hold.

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

⁵ The probability of moving from quintile 2 to quintile 3 is 21.6% in the original matrix (given in panel b of Table 1), while it is only 16.1% among municipalities with strong outmigration (given in panel a of Table 7). The difference does not seem to be very large, but the high number of observations in this income quintile (261) results in a considerable contribution to the test statistics (3.68). Similarly, the probability of moving down the income distribution from quintile 2 to quintile 1 increases from 3.7% in the full sample matrix to 5.4% with strong outmigration, giving a test statistics contribution of more than 2.

⁶ The probability of catching-up from quintile 2 to quintile 3 increases from 21.6% in the full sample to 38.3% in the immigration subsample (given in panel c of Table 7). The contribution to the test statistics is high (more than 6). The other two large error terms for this matrix is the probability of moving up or down from income quintile 4, which increases from 4.8% to 7% and decreases from 17.7% to 12.1%, respectively. In total, these two probability deviations have a test statistics contribution of 7.22.

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 20 counties in Norway include both periphery with large outmigration and urban areas with immigration.

6. Geographic patterns of convergence

As discussed in the introduction, the Norwegian development is characterized by large population flows and stark centralization in the settlement pattern. The large shift in population and economic activity from the periphery to urban areas is associated with convergence of per capita incomes between municipalities and economic regions. Here we examine the importance of geography and clustering to explain convergence using the methodological approach similar to Rey and Le Gallo (2009). Spatial Markov chains show how the performance of a region is affected by the performance of its neighbors. To which extent are the regions that are catching-up (or falling behind) geographically concentrated?

To take this into account, we estimate a spatial transition matrix for the Norwegian municipalities during 1972-2003. The sample of 401 municipalities is divided into 89 labor market regions. Based on regional income per capita relative to the average, the 89 regions are classified as income quintile 1-5, as earlier defined for municipal incomes. The regional income quintile classification is updated for each 4-year period. The spatial Markov chain consists of transition matrices for subsets of the municipalities depending on the income quintile of the greater labor market region to which they belong. The top matrix of Table 9 consists of 4-year transitions of municipalities located in regions with income per capita below 82.7% of the average (the lowest income quintile) at the beginning of a 4-year transition. The second matrix represents municipalities located in regions in income quintile 2, and so on.

As seen from Table 9, the impact of the spatial context is evident in both ends of the income distribution. On average, the probability of the low income municipalities to move up one income quintile equals 40.1% [given in cell (1, 2) in panel b of Table 1]. If these municipalities are surrounded by other low income municipalities, the probability of catching-up declines to 25.6% [given in cell (1, 2) of the top matrix of Table 9]. On the other hand, if their neighbors are high income, the probability of moving up the income ladder increases. A

low income municipality located in a labor market region with income per capita between 92.4% and 101% of the average (income quintile 3), has a 44.6% chance of catching up. Similar effects apply to high income municipalities. On average, the probability of the high income municipalities to move down one income quintile equals 19.9% [given in cell (5, 4) in panel b of Table 1]. For high income municipalities located in high income regions the chance of moving from income quintile 5 to income quintile 4 is only 10.9% [given in cell (5, 4) of the bottom matrix of Table 9]. The probability of moving down one income quintile increases as the income level of its neighbors decreases. If a municipality from the high income quintile is surrounded by municipalities from income quintile 4, the probability of moving down the income distribution increases to 39.7%.

Table 9 about here

The importance of the geographical environment in the convergence process is further investigated by looking at the relationship between the direction of a 4-year transition (down, none or up) and the income level of your neighbors. The probability of catching up increases with the income level of surrounding municipalities, and low income neighbors increase the chance of moving down the income distribution. The probability of catching-up is only 4.5% when the neighbors have lower income, but increases to 21.6% when the neighbors have higher income. Equivalently, the probability of going down in the distribution is high (28.1%) when neighbors are low income and low (6.5%) when neighbors are high income.

The overall conclusion of income convergence holds, but there is a distinct geographic pattern in the speed of transition towards convergence. To statistically test for the importance of the geographical context, we apply Pearson (Q) and Likelihood Ratio (LR) tests in similar ways as for tests of time independence, as explained in section 2. Comparing the five sub-matrices in Table 9 to the matrix for the entire sample simultaneously results in $Q = 116.8$ and $LR = 100.7$. With 40 degrees of freedom the critical value at 1% significance level equals 63.7, and, consequently, the null hypothesis of spatial independence is clearly rejected. The transition behavior of a municipality depends on the income level of its neighbors.

7. Concluding remarks

The data for regions in Norway, an integrated market with open flows of labor, capital, knowledge and technology, show income convergence. The large population flows from the periphery to the urban centers do not affect the convergence process much. In particular, catching up of low income regions has no systematic relationship to migration. Immigration in high income regions has a weak positive effect on income growth, and this holds back convergence. But overall the results are hardly consistent with strong agglomeration and scale effects and high income regions taking off.

The income development shows a distinct geographic pattern, reflecting historical differences in income levels and slow speed of convergence. The geographic pattern is consistent with the role of fixed endowments for the income growth path.

Our interpretation of the results is that homogenous institutions in an open integrated market are conducive to convergence and that brain drain/gain and agglomeration effects are limited. Possibly agglomeration effects are of less importance when knowledge and technology are available for all regions. It is of interest to build up more evidence about convergence and agglomeration in integrated markets. Future research may add more microeconomic evidence about labor migration and productivity spillovers among firms linked to regional growth experiences.

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.
- Baldwin R. and P. Martin (2004), Agglomeration and regional growth, in V. Henderson and J-F. Thisse (eds.), *Handbook of Regional and Urban Economics Cities and Geography*, Amsterdam: North Holland.
- Barro R. and X. Sala-i-Martin (1992a), Convergence, *Journal of Political Economy* 100, 2, 223-251.
- Barro R. and X. Sala-I-Martin (1992b), Regional Growth and Migration: A Japan-United States Comparison. *Journal of the Japanese and International Economies*, 6: 312-346.

- Bickenbach, F. and E. Bode (2003), Evaluating the Markov property in studies of economic convergence. *International Regional Science Review* 26, 363-392.
- 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*, forthcoming.
- Fingleton B. and M. Fischer (2008), Neoclassical Theory versus New Economic Geography; Competing Explanations of Cross-Regional Variation in Economic Development. SSRN, New York.
- 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.
- Ozgen C., P. Nijkamp and J. Poot (2009), The effect of migration on income convergence: Meta-analytic evidence, Tinbergen Institute Discussion Paper 2009-022/3.
- Persson J. (1997), Convergence across the Swedish counties, 1911-1993. *European Economic Review*, 41: 1835-1852.
- 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. (2004), Why are population flows so persistent? *Journal of Urban Economics* 56, 554-580.
- 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.
- Reichlin P. and A. Rustichini (1998), Diverging Patterns with Endogenous Labour Migration. *Journal of Economic Dynamics and Control*, 22: 703-728.

Rey, S. and J. Le Gallo (2009), Spatial analysis of economic convergence. In T. Mills and K. Patterson (Eds.), *Palgrave Handbook of Econometrics, Volume 2, Applied Econometrics*, Palgrave Macmillan.

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, 401 municipalities, relative income per capita, average values during 1972-1975 and 2000-2003.

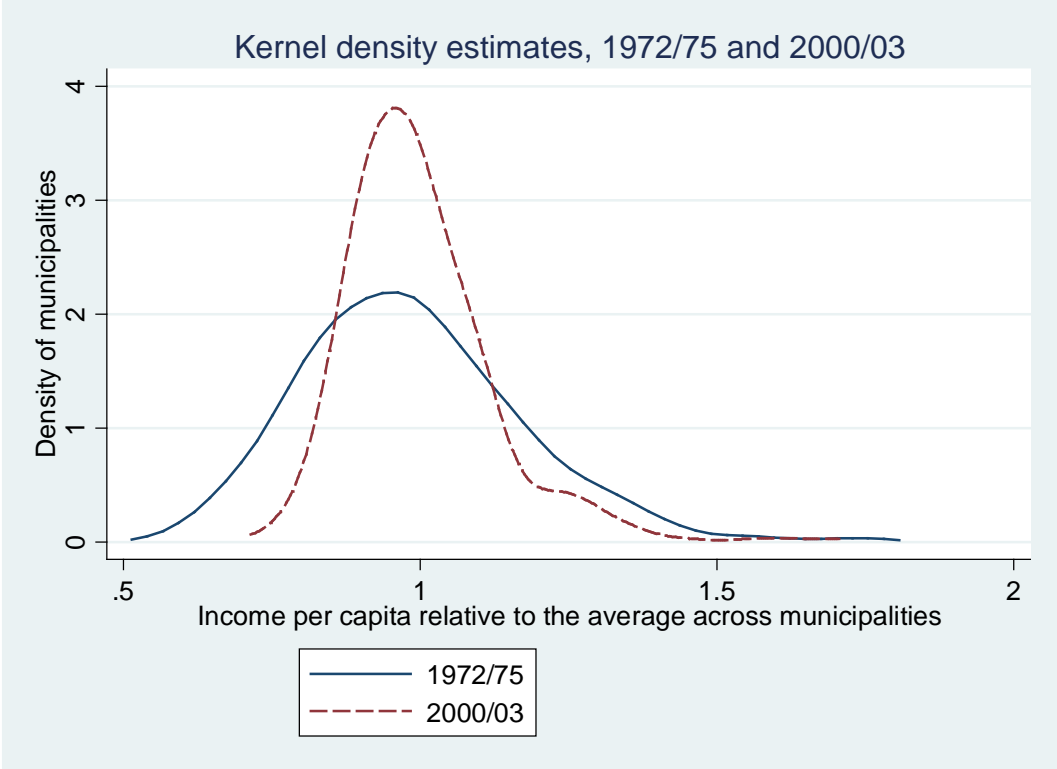


Figure 2: Population weighted Kernel density estimates, 401 municipalities, relative income per capita, average values during 1972-1975 and 2000-2003.

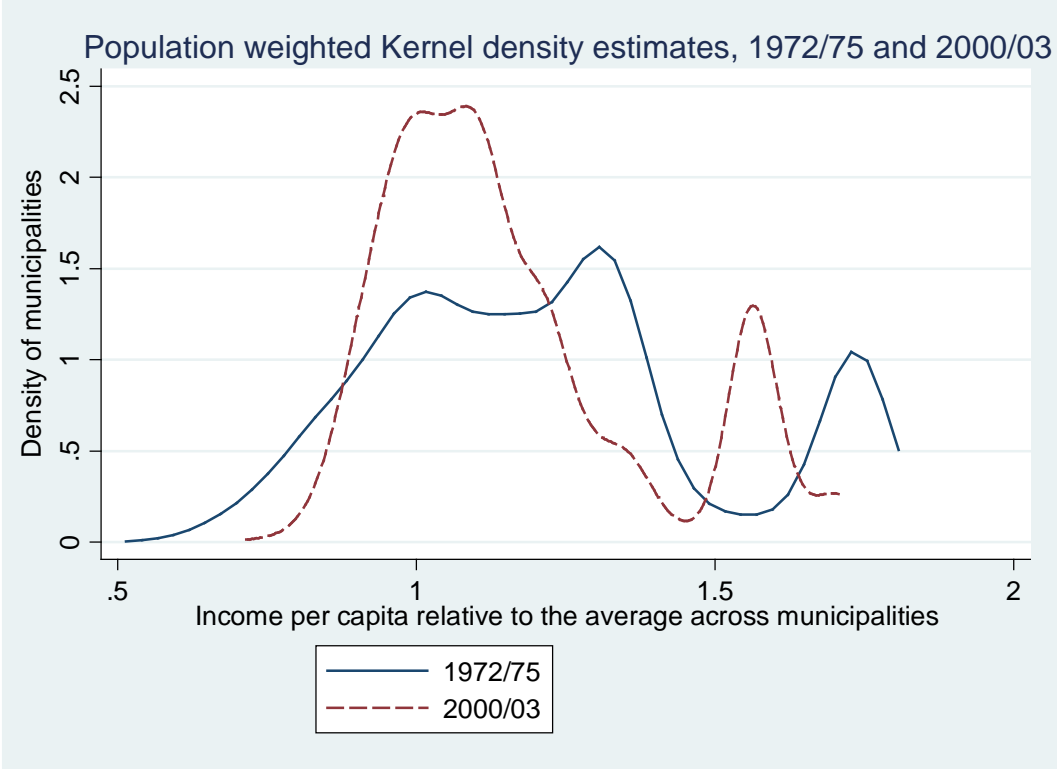


Figure 3: Population weighted Kernel density estimates, 400 municipalities (excluding the capital Oslo), relative income per capita, average values during 1972-1975 and 2000-2003.

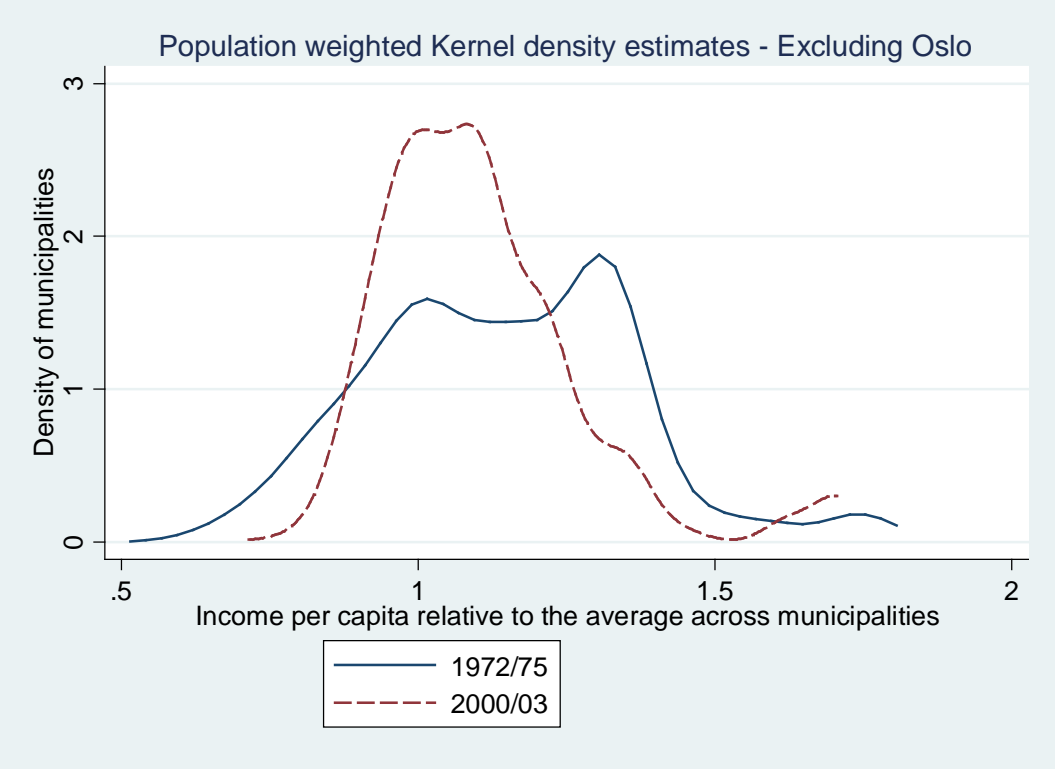


Table 1: Markov chain transition probability matrix, municipal level, 1972/75-2000/03.

Panel a: Full period 28-year transition (401 observations)

Income quintiles	1	2	3	4	5	Obs.
	≤ 0.827	≤ 0.924	≤ 1.01	≤ 1.135	> 1.135	
1	13.6	56.8	22.2	7.4		81
2		42.0	44.4	12.4	1.2	81
3		25.0	57.5	12.5	5.0	80
4		7.6	34.2	51.9	6.3	79
5			3.8	58.8	37.5	80
Initial distribution	20.0	20.0	20.0	20.0	20.0	
After 1 transition	2.7	26.4	32.4	28.4	10.0	
Ergodic distribution	0.0	22.7	45.0	25.6	6.6	

Panel b: 4-year transitions (2807 observations)

Income quintiles	1	2	3	4	5	Obs.
	≤ 0.827	≤ 0.924	≤ 1.01	≤ 1.135	> 1.135	
1	59.5	40.1	0.4			232
2	3.7	73.8	21.6	0.9		657
3		12.8	73.0	13.7	0.5	812
4			17.7	77.6	4.8	714
5				19.9	80.1	392
Initial distribution	8.3	23.4	28.9	25.4	14.0	
After 1 transition	5.8	24.3	30.7	26.7	12.5	
Ergodic distribution	1.9	21.2	37.4	31.1	8.4	

Table 2: Population weighted Markov chain transition probability matrix, municipal level, 1972/75-2000/03, 4-year transitions (2807 observations).

Income quintiles	1	2	3	4	5	Weighted obs.
	≤ 0.827	≤ 0.924	≤ 1.01	≤ 1.135	> 1.135	
1	55.7	44.0	0.3			0.2
2	2.3	72.2	24.9	0.6		0.6
3		9.8	76.8	13.1	0.3	1.2
4			12.0	84.6	3.4	1.6
5				7.6	92.4	3.4
Initial distribution	2.9	8.6	17.1	22.8	48.6	
After 1 transition	1.5	8.8	18.2	25.7	45.8	
Ergodic distribution	0.6	11.9	31.5	36.7	19.3	

Table 3: Markov chain transition probability matrix, regional level, 1972/75-2000/03.

Panel a: Full period 28-year transition (89 observations)

Income quintiles	1 ≤ 0.9	2 ≤ 1.01	3 ≤ 1.067	4 ≤ 1.2	5 > 1.2	Obs.
1	16.7	83.3				18
2	5.9	70.6	23.5			17
3		55.6	38.9	5.6		18
4		16.7	55.6	27.8		18
5			16.7	50.0	33.3	18
Initial distribution	20.0	20.0	20.0	20.0	20.0	
After 1 transition	4.5	44.9	27.0	16.9	6.7	
Ergodic distribution	4.7	65.9	27.3	2.1	0.0	

Panel b: 4-year transitions (623 observations)

Income quintiles	1 ≤ 0.9	2 ≤ 1.01	3 ≤ 1.067	4 ≤ 1.2	5 > 1.2	Obs.
1	68.3	31.8				63
2	2.9	87.3	9.3	0.5		205
3		19.0	71.9	9.2		153
4			20.3	79.1	0.7	148
5				24.1	75.9	54
Initial distribution	10.1	32.9	24.5	23.8	8.7	
After 1 transition	7.9	36.6	25.5	23.3	6.7	
Ergodic distribution	4.9	53.5	27.5	13.7	0.4	

Table 4: Test of time stationarity, 401 municipalities, 4-year transitions, 1972/75-2000/03. Contributions of single subperiods to the Pearson test statistics.

Subperiods	Income quintiles	Number of obs.	1 ≤ 0.827	2 ≤ 0.924	3 ≤ 1.01	4 ≤ 1.135	5 > 1.135	Sum
1972/75- 1980/83	1	128	0.01	0.00	0.55			0.56
	2	179	1.90	0.06	0.07	3.43		5.46
	3	177		1.96	1.07	1.12	0.02	4.17
	4	173			0.21	0.17	0.61	0.99
	5	145				0.12	0.03	0.15
	Sum							11.33
1980/83- 1988/91	1	72	0.08	0.16	0.31			0.55
	2	194	2.17	0.88	1.55	0.32		4.92
	3	223		1.09	0.01	1.85	1.09	4.04
	4	190			0.07	0.02	0.00	0.09
	5	123				0.26	0.06	0.32
	Sum							9.92
1988/91- 1996/99	1	23	0.01	0.16	8.22			8.39
	2	180	0.37	0.01	0.25	0.24		0.87
	3	282		3.90	0.94	0.05	0.28	5.17
	4	233			0.84	0.33	0.33	1.50
	5	84				0.01	0.00	0.01
	Sum							15.94
1996/99- 2000/03	1	9	0.08	0.10	0.04			0.22
	2	104	0.38	1.11	3.99	0.95		6.43
	3	130		0.01	0.10	0.80	0.21	1.12
	4	118			0.16	0.02	0.03	0.21
	5	40				0.11	0.03	0.14
	Sum							8.12
Pearson test statistic								45.31

Table 5: Test of Markovity of Order 1, 401 municipalities, 4-year transitions, 1976/79-2000/03.

Income quintile		Number of obs.	Transition probabilities $t-1$ to t					Test statistics	
$t-2$	$t-1$		1	2	3	4	5	Q	LR
All	1	151	60.9	38.4	0.7			25.0	22.6
	2	576	4.0	74.5	21.2	0.4		20.2	19.4
	3	732		13.3	72.7	13.5	0.6	37.5	36.4
	4	635			17.3	77.8	4.9	54.7	44.5
	5	312				19.2	80.8	24.3	33.9
	Sum							161.7	156.8
1	1	132	65.2	34.1	0.8			1.0	1.1
	2	90	12.2	80.0	7.8			23.6	20.9
	3	1			100.0			0.4	0.6
	4	0							
	5	0							
	Sum							25.0	22.6
2	1	19	31.6	68.4				7.3	7.1
	2	399	3.0	74.4	22.1	0.5		1.4	1.5
	3	129		20.2	70.5	7.8	1.6	10.2	9.4
	4	6			33.3	66.7		1.3	1.4
	5	0							
	Sum							20.2	19.4
3	1	0							
	2	87		69.0	31.0			8.1	11.4
	3	495		12.9	75.2	11.5	0.4	2.1	2.2
	4	97			37.1	58.8	4.1	26.6	21.5
	5	3					100.0	0.7	1.3
	Sum							37.5	36.4
4	1	0							
	2	0							
	3	107		6.5	63.6	29.9		26.7	22.7
	4	461			15.6	80.5	3.9	2.1	2.2
	5	28				57.1	42.9	25.9	19.6
	Sum							54.7	44.5
5	1	0							
	2	0							
	3	0							
	4	71				87.3	12.7	22.0	31.5
	5	281				15.7	83.3	2.3	2.4
	Sum							24.3	33.9

Table 6: Average change (%) in the population level during 1972-2003 in different income transitions. Based on 4-year transitions shown in panel b of Table 1.

Income quintile	1 ≤ 0.827	2 ≤ 0.924	3 ≤ 1.01	4 ≤ 1.135	5 > 1.135	Quintile average
1	-18.3	-14.3				-16.7
2	-13.1	-7.2	-0.9			-6.0
3		1.0	7.8	11.1		7.3
4			12.6	18.3	27.8	17.7
5				24.9	37.0	34.6

Table 7: Markov chain transition probability matrix according to migration pattern during 1972-2003, 4-year transitions 1972/75-2000/03.

Panel a: Municipalities with strong outmigration (700 observations)

Income quintiles	1 ≤ 0.827	2 ≤ 0.924	3 ≤ 1.01	4 ≤ 1.135	5 > 1.135	Obs.
1	64.6	34.8	0.6			161
2	5.4	78.2	16.1	0.4		261
3		16.7	68.6	14.7		156
4			21.4	73.0	5.6	89
5				21.2	78.8	33
Initial distribution	23.0	37.3	22.3	12.7	4.7	
After 1 transition	16.9	40.9	24.1	13.7	4.4	
Ergodic distribution	5.0	32.9	32.7	23.2	6.1	

Panel b: Municipalities with relatively stable population (1407 observations)

Income quintiles	1 ≤ 0.827	2 ≤ 0.924	3 ≤ 1.01	4 ≤ 1.135	5 > 1.135	Obs.
1	47.0	53.0				66
2	2.6	72.5	23.5	1.4		349
3		11.7	75.0	12.5	0.8	511
4			20.6	76.4	3.0	369
5				27.7	72.3	112
Initial distribution	4.7	24.8	36.3	26.2	8.0	
After 1 transition	2.8	24.7	38.5	27.2	6.8	
Ergodic distribution	1.0	20.7	44.0	29.9	4.4	

Panel c: Municipalities with strong immigration (700 observations)

Income quintiles	1 ≤ 0.827	2 ≤ 0.924	3 ≤ 1.01	4 ≤ 1.135	5 > 1.135	Obs.
1	60.0	40.0				5
2	2.1	59.6	38.3			47
3		12.4	71.0	16.6		145
4			12.1	80.9	7.0	256
5				16.2	83.8	247
Initial distribution	0.7	6.7	20.7	36.6	35.3	
After 1 transition	0.6	6.9	21.7	38.7	32.1	
Ergodic distribution	0.5	9.8	30.3	41.4	18.0	

Table 8: Test of whether the migration pattern affects transition probabilities, 401 municipalities 1972-2003, 4-year transitions. Contributions of single subsamples to the Pearson test statistics.

Migration pattern	Income quintiles	Number of obs.	1 ≤ 0.827	2 ≤ 0.924	3 ≤ 1.01	4 ≤ 1.135	5 > 1.135	Sum
Strong outmigration	1	161	0.71	1.13	0.14			1.98
	2	261	2.09	0.67	3.68	0.80		7.24
	3	156		1.81	0.42	0.13	0.76	3.12
	4	89			0.69	0.24	0.14	1.07
	5	33				0.03	0.01	0.04
	Sum							13.45
Relatively stable population	1	66	1.74	2.76	0.28			4.78
	2	349	1.09	0.08	0.58	1.04		2.79
	3	511		0.46	0.26	0.49	0.88	2.09
	4	369			1.82	0.06	2.46	4.34
	5	112				3.41	0.84	4.25
	Sum							18.25
Strong immigration	1	5	0.00	0.00	0.02			0.02
	2	47	0.30	1.29	6.06	0.43		8.08
	3	145		0.02	0.08	0.88	0.71	1.69
	4	256			4.45	0.36	2.77	7.58
	5	247				1.71	0.42	2.13
	Sum							19.50
Pearson test statistic								51.2

Table 9: Spatial Markov chain, 4-year transitions, conditioning on the income class of the greater labor market region at the beginning of each 4-year transition period.

Income quintile of neighbors	Income quintiles	1 ≤ 0.827	2 ≤ 0.924	3 ≤ 1.01	4 ≤ 1.135	5 > 1.135	Number of obs.
1	1	74.4	25.6				43
	2	12.5	87.5				8
	3			50.0	50.0		2
	4						0
	5						0
2	1	55.3	43.4	1.3			76
	2	4.1	73.4	21.9	0.6		169
	3		16.0	72.0	12.0		75
	4			33.3	66.7		12
	5						0
3	1	55.4	44.6				56
	2	4.5	75.7	19.3	0.4		243
	3		16.7	73.5	9.5	0.3	294
	4			25.0	70.5	4.5	112
	5				25.0	75.0	4
4	1	58.9	41.1				56
	2	2.5	73.1	23.9	0.5		201
	3		10.4	72.5	16.8	0.3	364
	4			16.7	78.3	5.0	419
	5				39.7	60.3	121
5	1		100.0				1
	2		63.9	27.8	8.3		36
	3		6.5	75.3	15.6	2.6	77
	4			14.0	81.3	4.7	171
	5				10.9	89.1	267