The limited role of education for regional income convergence: Evidence from 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

Increased education level and geographic concentration of educated people in cities are expected to generate urban income growth. We investigate the dynamic process of accumulating tertiary education and regional income growth in Norway since 1970. The increasing education level is a combination of catching up in the periphery and growth in the cities. Income levels are shown to converge in distribution analysis using kernel functions and first order Markov chains. The movements in the income distribution are unrelated to education. The hypothesis of equal income transition probabilities across subgroups of regions with different levels and changes in education cannot be rejected. We conclude that human capital has not been important for the pattern of income growth. Catching up from low income is not driven by education and income growth has not taken off in cities with high education level.

JEL codes: C14, I21, O15, R11, R12

Key words: Human capital, knowledge, urban growth, income convergence, education convergence, Markov chain, kernel density function

Date: August 18, 2011

* The project is funded by the Demosreg program of the Norwegian Research Council. We appreciate data handling by Jo Jakobsen, discussions at the economics department seminar in Trondheim and the NUPI workshop on 'Barriers to Entry, Internationalisation and Technology', and comments from Hans Bonesrønning, Fredrik Carlsen, Torberg Falch and Bjarne Strøm.

1. Introduction

Highly educated people concentrate in urban areas and generate ‘skilled cities’ or ‘smart cities’ (see Glaeser and Saiz, 2004, Shapiro, 2006, and Winters, 2011). The urban evolution is biased towards skill. Glaeser and Saiz find that areas with large initial concentrations of college-educated adults have experienced above-average population growth. And many of the growing cities have increasing share of highly educated inhabitants. Highly educated people move to cities to enjoy well-paid jobs and urban amenities. Winters argue that smart cities are growing primarily by inflow of young people pursuing education. The institutions of higher learning are located in cities. This broad pattern can be observed in most industrialized countries.

Only a limited literature looks at the distribution of human capital across regions within countries and the evidence is mixed. Studies of US regions conclude that human capital measured by population shares with college education find divergence during the last decades (Berry and Glaeser, 2005, Wheeler, 2006, Hammond and Thompson, 2010). A German study looking at shares of high-skilled workers conclude that the regional skill composition has converged over time (Sudekum, 2008). But the methodology in reaching these conclusions is different. When the dynamic process is studied in terms of changes in college/skill shares in the US studies, education divergence comes out. Highly educated cities increase their college/skill shares more in percentage points than periphery regions with low education level. When growth rates of college/skill shares are studied in the German analysis, education convergence results. The low education periphery regions have higher growth rates in college/skill shares.

We study the level and accumulation of education across 89 Norwegian NUTS-4 regions during the past four decades. The overall education level in Norway is high, and primary and secondary education are compulsory, but tertiary education varies significantly between regions. The education level is measured by the share of the grown up population with tertiary education (at the bachelor level or more) and is observed in 1970, as well as every year during the period 1980-2008. The share with tertiary education increases from 7% in 1970 to 25.5% in 2008. We observe concentration of highly educated in cities and also strong growth in the share with higher education in the periphery. Our distribution analysis using relative education levels indicates convergence. We look at this as the long run tendency of the data. It is of interest to pursue the background of the regional education pattern, but here we concentrate on the relationship with regional income growth.

Regions with high education level are expected to have higher income. The highly educated have human capital that adds to production factor inputs. Also, human capital is assumed to raise the innovative capacity and the capacity to imitate and absorb technology innovation elsewhere, and with spillover effects. The positive association between education and income levels is well established across countries and is also observed across regions within countries. Using individual data for France, Combes et al. (2008) find that the related skill sorting explains 40-50% of wage disparities between regions. The number is consistent with the finding of Glaeser and Gottlieb (2008) that differences in human capital account for about 50% of the variance in metropolitan-area wage levels in the US.

The dynamics of the income generation process is more complicated. The literature typically views smart cities as engines of growth. Early empirical evidence for the US is provided by Glaeser et al. (1995), Jaffe et al. (1993), and Rauch (1993). Initial level of schooling is an important determinant of economic growth and human capital come out with significant importance in growth regressions. The mechanisms are discussed by Black and Henderson (1999) in a model of urban growth with human capital externalities including both agglomeration economies and localized knowledge spillovers. Recent empirical evidence confirms the relationship between cities, skills and income growth as shown in overviews by Glaeser and Resseger (2010), Gennaioli et al. (2011), and Henderson (2007). The obvious prediction is that income growth is related to high and increasing education level in urban areas.

Econometric analyses of education effects face serious methodological challenges of endogeneity and unobservables. It is hard to find long-term exogenous determinants of education levels that are independent of income levels. We limit ourselves here to an analysis of the co-movements of education and income among regions using distribution analysis. The analysis concentrates on patterns of income transitions and relations to the accumulation of education levels. The investigation of systematic patterns between tertiary education and income transitions offers information of the importance of education for the regional income growth pattern.

The dynamic process of accumulating education and income generation is studied for 89 NUTS-4 regions. We are able to establish a measure of regional income per capita for the period 1972-2008 based on personal income from tax data. The income concept basically covers wage income. The distribution analysis shows convincing income convergence. The kernel density function of income levels is narrowing over time and first order Markov chains have ergodic distribution with single peak. The data do not show systematic differences in income transitions with respect to accumulation of education. The transitions in the income distribution are independent of whether regions are moving up or down in the distribution of relative education levels. Rising educational attainment is as common in regions catching up as in regions falling behind, and whether a low income region increases the relative education level or not does not affect the chances of catching up. The results of the analysis of accumulation of human capital are hardly consistent with education as a driving factor of income convergence. We also investigate the role of the education stock, but find only a weak relationship between education level and income growth. Given the evidence of income convergence this is not surprising. Overall we conclude that the high education level of cities has not contributed to an overall pattern of income divergence. The result is consistent with the analysis of lacking agglomeration effects by Rattsø and Stokke (2011).

The methodological approach is addressed in section 2, and the distribution analysis is outlined in section 3. Section 4 analyzes the accumulation of the education level across regions. Section 5 shows the evidence of regional income convergence in Norway. The relationship between income convergence and rising educational attainment is investigated in section 6. Section 7 studies income growth and education level. Concluding remarks are offered in section 8.

2. The methodological approach

The analysis of the role of education for income growth must deal with other determinants of regional development. The income growth in urban regions may result from local resources as the source of high productivity growth. Highly educated people move to cities because they are more productive and offer higher wages – the cities are not necessarily more productive because they have more educated 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. Econometric studies address these issues of endogeneity and sorting and the likely effect of background unobservables. The overview article of Moretti (2004) discusses a variety of ways to handle the methodological challenges, and he concludes that the evidence for human capital externalities in cities is not convincing.

We are interested in the long run growth process and the dynamics of the accumulation of education. The desired dataset would reflect a natural experiment with large shifts in education levels over time independent of local income generation. But finding instruments to represent some exogenous part of the long run shift in the education pattern is difficult. We want to take benefit of a long time series of education and income growth and it follows that we cannot easily separate between the consequences of education for income growth, the importance of income growth for migration of highly educated persons, and other sources of increasing productivity and income. Our ambition is more modest, we investigate whether our time series observations are consistent with a relationship between the patterns of education and income growth.

Our starting point is that there must be systematic relationships between transitions in the distribution of education and incomes across regions. When education is important, we expect positive association between level and change of education and income growth. It should be noticed that many other adjustment mechanisms can be involved. Most important is the separation of education as a labor demand versus labor supply effect. When the labor demand effect is dominant, we expect high income growth in low-income regions catching up in education level and in high-income regions with high education level. The labor supply effect implies that income (wage) growth in low-income regions with increasing education level and high-income high education level regions is held back. In general the relationship between education and income can be correlated with a third unobserved variable. Young people can be more educated, but have lower wages, because they have less experience. Regions with rich local resources may have high income growth even when the education level is low. Also endogenous mechanism may be at work, such as regional income stagnation stimulating young people to take more education. We discuss the interpretation of our results in the concluding section. Broadly we assume that the lack of a relationship between level and change of education and income growth among regions indicates that education is of limited importance for the regional income development.

We apply distribution analysis to capture heterogeneous processes with different growth paths from different starting points. In particular, we can study the two ends of the distribution of per capita incomes – relative low income and relative high income regions. We estimate the transition probabilities of the Markov chains by the maximum likelihood method to facilitate tests of time stationarity, as well as tests of how education levels and changes are related to income transitions. There is a large literature applying Markov chain transition probability matrices to study income convergence across regions and countries. 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).

The analysis is based on data for taxable income of each of 89 economic regions, as defined by the European Union standard of NUTS-4 regions. This level of aggregation captures functional regions understood as common labor markets. The data cover all years during the period 1972-2008 and it follows that we have 3.293 observations of per capita incomes. Personal income measured in the tax statistic basically reflects wage income, and capital income is hard to locate at this level of disaggregation. No regional GDP measure is available. The education level across regions is quite similar for primary and secondary education, since both are compulsory. The interesting variation relates to tertiary education. We measure the level of education in each region as the share of the grown up population with tertiary education, including both short higher education (college level, up to 4 years in duration) and long higher education (university level, more than 4 years in duration).[1] The data cover the single year 1970 and all years during the period 1980-2008. In the analysis the income and education levels are measured relative to the average levels of income and education across all regions in each year. The discussion of the method below relates to the distribution of income, but we also apply the method to investigate the distribution of education levels across regions.

3. Distribution analysis

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

[1] Data source: Statistics Norway, Table 06983; Number of persons above 16 years according to the level of education.

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[2], 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 income levels and is estimated based on the Markov chain matrix under the assumption that the transition dynamics remain unchanged.

To check whether the estimated transition probabilities are constant over time we test for time stationarity. 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 *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 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 *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):

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

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

To statistically test for the relationship between income growth and education (both level and change), we apply Pearson and Likelihood Ratio tests in similar ways as for tests of time stationarity. The test investigates whether the income transition probabilities are independent of the level of education and the increase in the education level. The test divides the entire sample of regions into *M* mutually exclusive and exhaustive subsamples according to the degree of education change/level and compares the transition matrices under each of the *M* subsamples to the entire sample. The Pearson and Likelihood Ratio test statistics are calculated from equations (4) and (5).

4. Education level convergence

The convergence in education levels has been observed across countries in OECD data. Wolff (2000) finds convergence in schooling levels using dispersion measures and observe that it corresponds to convergence in labor productivity levels. Cuaresma (2006) estimates kernel density functions for educational attainment, but different data sets provide contradictory conclusions. A German study by Sudekum (2008) concludes that there has been convergence of regional skill composition over time, also for high-skilled with university degree. Studies of US regions find divergence in education levels. Berry and Glaeser (2005) analyze cities (metropolitan areas) using the standard growth regression framework. They draw the conclusion that human capital levels measured by college education diverge during the last three decades. Regions with higher levels of human capital have attracted more educated people. Wheeler (2006) and Hammond and Thompson (2010) add to the US evidence, the last study using kernel density estimates.

As noted in the introduction, the conflicting results regarding education convergence is related to methodological issues. The US studies look at changes in college/skill shares, and highly educated cities increase their college/skill shares more in percentage points than other regions. When growth rates of college/skill shares are studied, like the German analysis, education convergence results. The low education periphery regions have higher growth rates in college/skill shares. Our dataset looks similar, the cities have higher increases in shares with high education in percentage points, but periphery regions have higher growth rate of the tertiary share.

We apply distribution analysis to study education convergence across regions in Norway below, and the distribution is analyzed by the relative education level. The level of education has increased significantly during the period studied. The development of the distribution of the education level in the regions is first described by estimated kernel density functions for the first year 1970 and the last year 2008, as shown in Figure 1. The horizontal axis represents the share of the grown up population with tertiary education relative to the average share across regions, while the vertical axis gives the density of regions at different relative education levels.[3] Both distributions have a single-peak around the average educational level, but over time, the distribution becomes narrower and the peak more pronounced, indicating convergence with respect to the level of education.

Figure 1 about here

The robustness of this result is investigated by estimating the kernel using absolute shares of tertiary educated, similar to Hammond and Thompson (2010). The log shares are presented in Figure 2 and they show the positive shifts in the education share since 1970 and that the distribution has narrowed over time. The standard deviation of the log distribution is reduced over time. We have calculated the variance coefficients of the absolute shares and they also are reduced. Different from the US studies it seems to us that education divergence is inconsistent with our data.

Figure 2 about here

Furthermore, we investigate the education dynamics through a Markov chain transition matrix for the period 1970-2008. Since education levels change slowly we focus on decade transitions, and apply the relative education levels in 1970, 1980, 1990, 2000, and 2008. Four transitions and 89 regions imply that transition probabilities are estimated based on 356 observations. We follow the convention of discretization based on a uniform initial

[3] 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 0.1914 and 0.1014 for 1970 and 2008, respectively.

distribution of relative educational levels across education groups, which gives the following five groups: 1) less than 79% of the average educational level, 2) between 79% and 90%, 3) between 90% and 100%, 4) between 100% and 111%, and 5) more than 111% of the average level of education. The Markov matrix with respect to the level of education is given in Table 1. As seen from the binomial standard errors given in parentheses, the estimated transition probabilities are significant. The only exception is the probability of moving two education groups during a decade (from group 2 to group 4).[4]

The transition matrix is consistent with the findings from the kernel functions with convergence in educational levels across regions. Regions located in the lowest education group (below 79% of the average level) have 25% chance of moving up the distribution during a 10-year period. Regions in education group 2 are more likely to move towards the middle of the distribution than towards the lower end. The probability of moving upwards from education group 2 is more than 20%, compared to 6% chance of moving downwards. In education groups 3 and 4 the probabilities of upward and downward movements are about the same. Regions in the highest group (educational level at least 11% higher than the average) have good chances of remaining in this group (89%). These dynamics imply a movement towards the top end of the distribution, and education group 5 dominates the long-run distribution with more than 30% of the regions (given in the last row of the matrix). The distribution of educational levels goes from a uniform distribution initially towards a single-peaked distribution, consistent with convergence in education levels, and shows no tendencies of a bimodal twin peaked distribution. But the transition towards the long run distribution is slow, and based on the second eigenvalue of the matrix we find that it takes more than 80 years to reach half way to the steady state. The understanding of the peak at the top end of the distribution is that when regions have reached the highest education group, they are not likely to leave it again. Kremer et al. (2001) document similar dynamics for the world income distribution.

Table 1 about here

---

[4] The binomial standard error for the probability of moving from group i to group j is defined as $\sqrt{\frac{\hat{p}_{ij}(100-\hat{p}_{ij})}{n_i}}$, where $\hat{p}_{ij}$ is the estimated probability and n_i is the total number of observations starting from group i.

To sum up, the density functions and the Markov matrix both identify a clear pattern of convergence with respect to educational level among Norwegian regions during 1970-2008. The growth of the share of the population with tertiary education has been higher in the periphery, and in growth rate terms the emergence of smart cities has been dominated by catching up from behind. To further investigate the possible existence of smart cities we take a closer look at the dynamics within the top education group. Based on the transition matrix above, we divide group 5 into two equal subgroups: a lower group between 111% and 136% of the average education level, and an upper group with at least 36% higher education level than the average. The disaggregation does not reveal a small group of highly educated smart cities taking off relative to the rest. The probability of moving down the distribution from the lower group is 22%, compared to only 8% probability of moving up. Even regions in the upper group have 20% probability of moving down the distribution. The share of regions in this group decreases from 10% initially to 9% in the long run, while the lower group expands and eventually accounts for 18% of the regions. This implies that the expansion of the top education group in Table 1 is accounted for by more regions in the lower end of this group (between 111% and 136% of the average education level), and not by the growth of highly educated smart cities. A closer look at the data shows that 15 out of 89 regions are in education group 5 in at least four of the five transition periods. Among these, the broad picture is not education take off, but movement towards the average. Only 2 of the 15 regions have higher relative level of education in 2008 compared to 1970, while the rest have lower or about the same relative education level. On average the relative level of education among the top 15 group decreases from 1.65 to 1.39 during 1970-2008.

5. Income level convergence

An enormous literature has analyzed income convergence between regions and countries. 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. An influential summary of the recent literature using econometric analysis of panel data has been given by de la Fuente (2002), including his own analysis 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 use of distribution analysis came as a reaction to these shortcomings and they have given more support to income divergence. The study of 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.

To sort out the relationship between education and income growth, we need to address the income generation process. Income convergence is analyzed in the same manner as education convergence by using kernel density functions as well as Markov chain transition matrices. To examine how the distribution of regional income per capita develops over time, we compare the estimated kernel density functions for the first year 1972 and the last year 2008, as shown in Figure 3.[5] The horizontal axis represents income per capita relative to the average level across regions, while the vertical axis gives the density of regions at different relative income levels. Both functions have a single-peak distribution with the majority of regions located close to the average level of income per capita. The estimated distributions show a clear pattern of convergence over time. The distribution is narrower and more concentrated around the peak in 2008 than in 1972. Compared to the distribution of education levels, the variations in income per capita are smaller across regions.

Figure 3 about here

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 focus on 4-year transitions in the analysis 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 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 91% of the average, 2) between 91% and 96%, 3) between 96% and 101%, 4) between 101% and 107%, and 5) more than 107% of the average income across regions. The transition probability

[5] Consistent with Silverman's rule of thumb the bandwidth is set to 0.075 and 0.0474 for 1972 and 2008, respectively.

matrix is shown in Table 2, and is estimated based on nine 4-year transitions and a total of 801 observations. Most of the estimated transition probabilities are significant. As seen from the binomial standard errors, all the diagonal and immediately off-diagonal transition probabilities are significant, while the estimates of the probability of moving two income groups during a 4-year period are typically insignificant.

Table 2 about here

The Markov matrix shows income convergence across regions. The distribution of per capita incomes is tending towards a point mass, rather than towards a two-point distribution. Regions in the lowest income group (income level relative to the average below 0.91) have 24% probability of catching up during a 4-year period, and the high income regions have 19% chance of moving down the distribution. Regions in income groups 2 and 4 have higher probability of moving towards the middle of the distribution than towards the respective ends. The probability of moving from group 4 to the high income group is 11%, compared to about 25% 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. Regional incomes go from a uniform distribution initially to a normal distribution in the long run. The peak is located around the average income level, with income groups 2-4 (relative income between 91% and 107% of the average), accounting for about 70% of the regions in the long run. The lowest and the highest income groups are reduced from 20% initially to about 15%. The distribution tends to accumulate in the middle, combined with thinning of both the lower and the higher tail, consistent with income convergence. The stationarity of the distribution is addressed by looking at the second eigenvalue of the transition matrix, which is found to equal 0.89. This implies that the distribution converges to a steady state, and the asymptotic half life of the process (the time it takes to reach half way to the long run distribution) equals 5.9 transition periods, which corresponds to about 24 years.

We investigate the time stationarity (constant transition probabilities over time) of the income transition matrix by applying the Pearson (Q) and the Likelihood Ratio (LR) tests, as described in section 3. The transition probabilities are estimated based on nine 4-year transitions during 1972-2008. The sample period is divided into three subperiods, each containing three 4-year transitions. The transition matrices for each subperiod are then compared to the full period matrix (given in Table 2). With 28 degrees of freedom the critical value at 5% significance level equals 41.3. The test statistics are calculated to Q = 39.4 (prob = 0.07) and LR = 40.1 (prob = 0.06), which means that the null hypothesis of constant transition probabilities over time cannot be rejected. The detailed contributions to the Pearson test statistic from each transition in the four subperiods are shown in Table 3. The differences in transition probabilities are minor. Of the 57 comparisons (19 probabilities, 3 subperiods) about half the probabilities contribute with less than 0.4 to the test statistic, and more than ¾ have error terms lower than 1. Transition probabilities seem to be constant over time.[6]

Table 3 about here

6. Relationship between change of education and income transitions

The data analyzed above show large regional heterogeneity with respect to education and income. Since high education level in cities does not dominate and lead to education divergence and income divergence is not observed, the big picture is not consistent with smart cities driving the income growth pattern. However, the expanding education level in periphery regions is consistent with convergence in the level of income. In this section we investigate the relationship between the income transitions and the changes in the relative education levels in the regions. Is there a systematic pattern of rising educational attainment in regions moving upward in the income distribution? Referring to the discussion at the start of section 4 we also look at differences between regions with respect to changes in the level of education in percentage points.

Direct observation of the distribution across regions in the early 1970s confirms the expected relationship between education level and income level. Small regions in the periphery have low income and education level, while the large cities have high share of the grown up population with tertiary education and high income level. But the correlation between education level and income level is decreasing over time (raw correlation coefficient 0.80 falling to 0.71), and the correlation coefficient between change in education and change in

[6] We have also divided the sample period into two subperiods, and the null hypothesis of time stationarity is still not rejected, and even holds at 15% significance level.

income is only 0.43. This is our first indication that change in education level has not been of much importance for the income level.

Furthermore, we investigate the role of education for income growth by calculating income transition probability matrices conditioned on the movement in the distribution of education. We rank the 89 regions according to the change in their relative level of education, and divide the sample into two equal subsamples (top 50% and bottom 50%). Among the top 45 regions, the relative level of education on average increased by 0.13 (from 0.79 to 0.92). Hence, this subsample reflects regions with below average level of education that is gradually moving up in the education distribution. In the other end, the 44 regions with the largest decrease in relative education are dominated by regions with above average level of education that are gradually moving towards the middle of the education distribution (decreasing by 0.13 from 1.21 to 1.08, on average). The estimated Markov matrices for the two subsamples of regions are shown in panels a and b of Table 4. The relationship between income level and change in the relative education level is reflected in the number of observations for the different income groups in the two education subsamples. The subsample with large decreases in the relative educational level is dominated by regions in the upper half of the income distribution, while movements up the education distribution are more common in the bottom half of the income distribution.

Table 4 about here

The broad picture is that the development in the relative level of education does not matter much for income transition probabilities. If increased education level is important for upward income transitions, we expect the numbers above the diagonal to be higher in the matrix in panel a than in panel b. But the transition probabilities are roughly similar. Whether a region move up or down in the distribution of relative educational levels does not affect its chances of catching up with respect to income. If anything, upward transitions seem to be more likely for the subsample of regions with large decrease in the relative level of education. When it comes to the lowest income group, almost all the regions belong to the subsample with large increase in relative education, and the matrices do not give much information about the role of increased education for income catch-up from this group. But a closer look at the data shows that the average increase in the relative level of education is actually higher among regions that remain in the lowest income group than in regions that are able to catch-up (0.16 vs.

0.12). We further consider the dynamics within the lowest income group by dividing the group into two equal subgroups: a lower group with income p.c. below 86% of the average, and an upper group with relative income between 86% and 91%. Again, the average increase in the relative level of education is higher among regions that remain in the lower end of income group 1 than in regions moving from the lower to the upper end of the group. And regions moving downwards within group 1 have larger increase in relative education than regions remaining in the upper end of the group or moving up the income ladder. Rising educational attainment is observed in regions both catching up and falling behind, and cannot explain the income convergence seen in the data.

The transition probabilities in the upper end of the distribution are consistent with this view. Regions with increasing relative level of education that are in the third income group have 22% chance of catching up. For regions with decreasing relative educational levels the same probability equals 25%. Similar, the probability of catching up from the fourth income group is largely independent of the development in the relative level of tertiary education. Income growth has not taken off in high income regions with rising relative education level. And for regions in the subsample with decreasing relative level of education, the decrease is much larger among regions that remain in the top income group than in regions moving down the distribution.

To statistically test for the importance of changes in the relative educational level for the convergence process, we apply Pearson and Likelihood Ratio tests, as explained in section 3. In this way, we can investigate whether the transition probabilities are independent of the change in the relative education level. Comparing the matrices in Table 4 to the matrix for the entire sample of 89 regions (given in Table 2) simultaneously results in test statistic equal to about 22 for both tests.[7] With 13 degrees of freedom, the 5% critical value equals 22.4. This implies that the null hypothesis of equal transition probabilities across different developments in relative education cannot be rejected. The contributions to the Pearson test statistic from each transition in the two subsamples are given in Table 5. Differences in transition probabilities are minor, and most error terms are well below 0.5. The relatively large test

[7] When comparing the matrices, we exclude transitions with six or less observations in the full sample matrix if these are unevenly allocated across subsamples. Transitions with only one observation in the full sample matrix are excluded if the single observation is located in a subsample with less than 40 observations in the same income group (giving an unreasonably high probability and large error term). The degrees of freedom are adjusted accordingly. In this case, the transition from group 5 to group 3 is excluded.

statistic comes from the inaccurate estimates of transition probabilities in the lowest income group in panel b of Table 4, which is based on only 7 observations. To sum up, rising educational attainment cannot explain the income catch-up of low income regions, and income growth has not taken off in high income regions with increasing education level.[8]

Table 5 about here

In the introduction we have discussed the importance of looking at growth rates of education shares and change of education shares. Instead of focusing on the movement in the distribution of education (the development in a region's level of education relative to the average) we consider the change in the absolute level of education, measured as the percentage point increase in the tertiary education share of the grown up population during 1970-2008. Given this classification of regions, we perform the same analyses and tests as above, and our main results remain. The null hypothesis of equal transition probabilities across subsamples of regions cannot be rejected and even holds at 15% significance level. Rising educational attainment cannot explain the income convergence found in the data.

As seen from Table 4, the probability of moving down from the second income group to the lowest group is in fact higher for regions that are catching up with respect to education. Similarly, the probability of remaining in the top income group is highest among regions with large decreases in the relative level of education. These counterintuitive findings might indicate an importance for the level of education in income growth. The group of regions moving up the education distribution typically starts from below average level of education, while regions moving down the distribution start with above average level of education. We investigate the relationship between education level and income growth in the next section. But given the convincing evidence of income convergence among Norwegian regions (section 5) together with high correlation between income and education, we do not expect a strong association between education level and income growth.

[8] When dividing the regions into three (rather than two) subsamples according to the change in the relative level of education, the main conclusions still hold.

7. Relationship between level of education and income transitions

The role of the education level for income growth is analyzed using the same method as above for the accumulation of educational attainment. We estimate income transition probability matrices conditioned on the average level of education during the past four decades. The regions are ranked according to the tertiary education share of the grown up population, and the sample is divided into two equal subsamples (top 50% and bottom 50%). Among the regions with high level of education the tertiary share ranges from 12% to 33% and equals 15% on average, while the regions with low educational level has an tertiary share in the range 7-12% and an average of about 10%. The estimated Markov matrices are shown in panels a and b of Table 6. Tables 4 and 6 are similar, but the regions in the two subsamples are different, now based on the education level. The relationship between income level and education level is reflected in the number of observations for the different income groups in the two education subsamples. More than half the observations of high education level regions are in the two top income groups. Similar, low education level regions are overrepresented in the lower end of the income distribution.

By comparing the estimated transition probabilities we observe a positive relationship between education level and income growth, but the size of the effect is small, and not sufficient to generate income divergence, as one would expect. Among low income regions with low educational level the probability of moving up the income ladder equals 23%, compared to 28% for low income regions with high level of education. A closer look at the data shows that the average level of education is about the same among regions that remain in the lowest income group and among regions that are able to catch-up. The positive role of education level is mainly seen for income group 2, where regions with high level of education have 29% chance of moving up the income distribution and 9% chance of falling behind, compared to 19% and 21%, respectively, for regions with low level of education. In the other end of the distribution, the probability of moving upwards from group 4 is actually higher for low education level regions, but high level of education increases the probability of remaining in the top income group. Among regions staying in the high income group the share of the adult population with tertiary education equals 17% on average, while in regions moving downwards from the top group the average tertiary share equals 13%. This is again only a small trace of smart cities, to be pursued below.

Table 6 about here

The limited role of education level for income growth is confirmed by the implied ergodic distributions of the two sub-matrices. If the level of education is important to generate growth, low income regions with low educational level should remain in the lowest income group, while high income regions with high educational level should take off and increase the income gap. The opposite is happening. Among regions with low education, one third is initially in the lowest income group, 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 group contains about a quarter of the regions. Similarly, among regions with high education, the top income group is not taking off, but is significantly reduced from 32% of the regions initially to 19% in the long run.

The test for the importance of the education level for the income growth follows the design above using Pearson (Q) and Likelihood Ratio (LR) tests. The transition probabilities in the two sub-matrices of Table 6 are compared to the full sample matrix.[9] With 12 degrees of freedom, the 5% critical value equals 21. The test statistics are calculated to Q = 18.3 (prob = 0.11) and LR = 19 (prob = 0.09). The null hypothesis of equal transition probabilities across different levels of education is not rejected, and the Pearson test even holds at 10% significance level. The contributions to the Pearson test statistic from each transition in the two subsamples are given in Table 7.

Table 7 about here

The top 20 smart regions in Norway are presented in Table 8. Most of the regions have an average tertiary education share in the range 14-20% and with Oslo and neighbor Bærum/Asker at the top with more than 25%. While the education level has increased significantly, the relative education has been reduced among most of the top 20. Bærum/Asker had an education level four times higher than the average in 1970, but only about twice as high in 2008. This is consistent with education level convergence. And the top 20 smart regions have not had income growth take off. Rather, most of them have reduced relative income in 2008 compared to 1972, and some are even below average income in 2008.

[9] Using the same arguments as described in footnote 7, we exclude the transition from group 4 to group 2 and the transition from group 5 to group 3 in this comparison.

The main exception is the oil region Stavanger/Sandnes. The numbers are consistent with income convergence. Three regions in the northern part of Norway (Tromsø, Bodø and Harstad) had income and education levels close to the average in the early 1970s. They have increased their relative level of education significantly during the past four decades, but the income level is still around the average. Despite large education increases, they are not able to catch-up with respect to income.

Table 8 about here

Based on the analyses in sections 6 and 7, we conclude that changes in the relative level of education cannot explain the observed income transitions, and that the relationship between education level and income transitions is positive, but weak. Large increases in the relative educational level cannot account for the income catch-up of low income regions, and high income regions with high level of education do not take off and generate income divergence. As a robustness check we have done the same analysis at the municipal level. Norway is characterized by many small municipalities (average population size is about 10.000 in 2008), and we therefore focus on the subsample of 176 municipalities with more than 5.000 inhabitants. Based on kernel density functions and Markov chain transition matrices the findings of education and income convergence are confirmed. The initial uniform distribution of relative income levels goes towards a normal distribution in the long run. The majority of municipalities gather around the average income level, while the two ends of the distribution lose significance. We then rank the municipalities both according to the average level of education and the change in the relative level of education during 1970-2008, and estimate Markov matrices for the different subsamples. The broad picture is that the transition probabilities are not much affected by the education level or the accumulation of education.

8. Concluding remarks

We have investigated the relationship between education and income growth at the regional level. The background literature on smart cities indicates that highly educated people concentrate in cities and that the associated increase in human capital contributes to income growth. Our dataset for Norwegian regions confirms the broad expansion of the education level. The regional pattern of the development of the shares of people with tertiary education reflects both catching up in periphery regions and increasing education level in the cities.

Regions with increasing education level are expected to have higher income growth. The Norwegian data offer some contrarian observations worth contemplating. Using distribution analysis we find that the regional income level is converging, but the movements in the income distribution are unrelated to the accumulation of education. The overall conclusion of the analysis is that education has a limited role in explaining the income growth pattern among Norwegian regions.

The Norwegian economy is characterized by large movements of population and economic activity from the periphery to urban centers. Interestingly, this urbanization and structural change is combined with income convergence and the education level has not diverged. The emergence of smart cities with high education level and income growth is not dominating the pattern of income growth.

It can be argued that our result follows from heterogeneity and variation in quality of tertiary education. But compared to international studies, the quality differences across tertiary education in Norway probably are small. The quality differences existing may sort themselves geographically so that the high quality competence ends up in the urban centers. If this is true, we expect urban centers to have even higher income growth. It is obvious that this mechanism cannot contribute to income convergence. The expanding service sectors in the advanced urban economy seem not to be that tertiary education intensive, as they also invite inflow of unskilled labor. And they don’t contribute much to income level growth in our data.

Income growth following higher education level must result from both supply and demand effects at the market for human capital. The increased education level measured in this analysis shows that the supply side has delivered in quantity. The income catching up in the periphery and overall income convergence is consistent with dominating labor demand effect for highly educated in the periphery and dominating labor supply effect of highly educated in the cities. Most of the new candidates from higher education in Norway end up in the public sector. In particular the relative size of the public sector is expanding in the periphery with many college-educated nurses and teachers in local services. They help keep up the income level in the many small regions, but the expanded public administration does not contribute much to overall income growth. The return to tertiary education is low, in particular in the public sector. At the same time limited demand for tertiary education in the expanding private service sector in urban areas may hold back income growth in the cities. This structural background of the limited role of education is worth further analysis.

References

Berry, C. and E. Glaeser (2005), The divergence of human capital levels across cities, *Papers in Regional Science* 84, 407-444.

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 J.V. Henderson (1999), A theory of urban growth, *Journal of Political Economy* 107, 252-284.

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

Cuaresma, J. (2006), Convergence of educational attainment levels in the OECD: More data, more problems? *Economics of Education Review* 25, 173-178.

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

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

Glaeser, E.L. and M.G. Resseger (2010), The complementarity between cities and skills, *Journal of Regional Science* 50, 1, 221-244.

Glaeser E.L. and A, Saiz (2004), The rise of the skilled city, *Brookings-Wharton Papers on Urban Affairs* 5, 47-94.

Glaeser E. L., J. Scheinkman and A. Shleifer (1995), Economic growth in a cross section of cities, *Journal of Monetary Economics* 36, 117-143.

Gennaioli, N., R. La Porta, F. Lopez-de-Silanes and A. Shleifer (2011), Human capital and regional development, NBER Working Paper 17158.

Hammond, G. W. and E. Thompson (2010), Divergence and mobility in college attainment across U.S. labor market areas: 1970-2000, *International Regional Science Review* 33, 4, 397-420.

Henderson, J. V. (2007), Understanding knowledge spillovers, *Regional Science and Urban Economics* 37, 497-508.

Jaffe, A.B., M. Trajtenberg and J.V. Henderson (1993), Geographic localization of knowledge spillover as evidenced by patent citations, *Quarterly Journal of Economics* 108, 577-598.

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.

Moretti. E. (2004), Human capital externalities, in V. Henderson and J.-F. Thisse (eds.), *Handbook of Urban and Regional Economics*, vol. 4, Amsterdam: Elsevier North-Holland.

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.

Rauch, J. (1993), Productivity gains and geographic concentration of human capital: Evidence from the cities, *Journal of Urban Economics* 34, 380-400.

Rattsø, J. and H. Stokke (2011), Migration and dynamic agglomeration economies: Regional income growth in Norway, Working Paper No. 02/2011, Department of Economics, Norwegian University of Science and Technology.

Shapiro, J. (2006), Smart cities – quality of life, productivity, and the growth effects of human capital, *Review of Economics and Statistics* 88, 324-335.

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

Sudekum, J. (2008), Convergence of the skill composition across German regions, *Regional Science and Urban Economics* 38, 148-159.

Wheeler, C. H. (2006), Human capital growth in a cross-section of US metropolitan areas, *Federal Reserve Bank of St. Louis Review* 88, 113-132.

Winters J. (2011), Why are smart cities growing? Who moves and who stays, *Journal of Regional Science* 51, 2, 253-270.

Wolff, E. (2000), Human capital investment and economic growth: exploring the cross country evidence, *Structural Change and Economic Dynamics* 11, 433-472.

Figure 1: Kernel density estimates, relative level of education, 89 regions, 1970 and 2008.

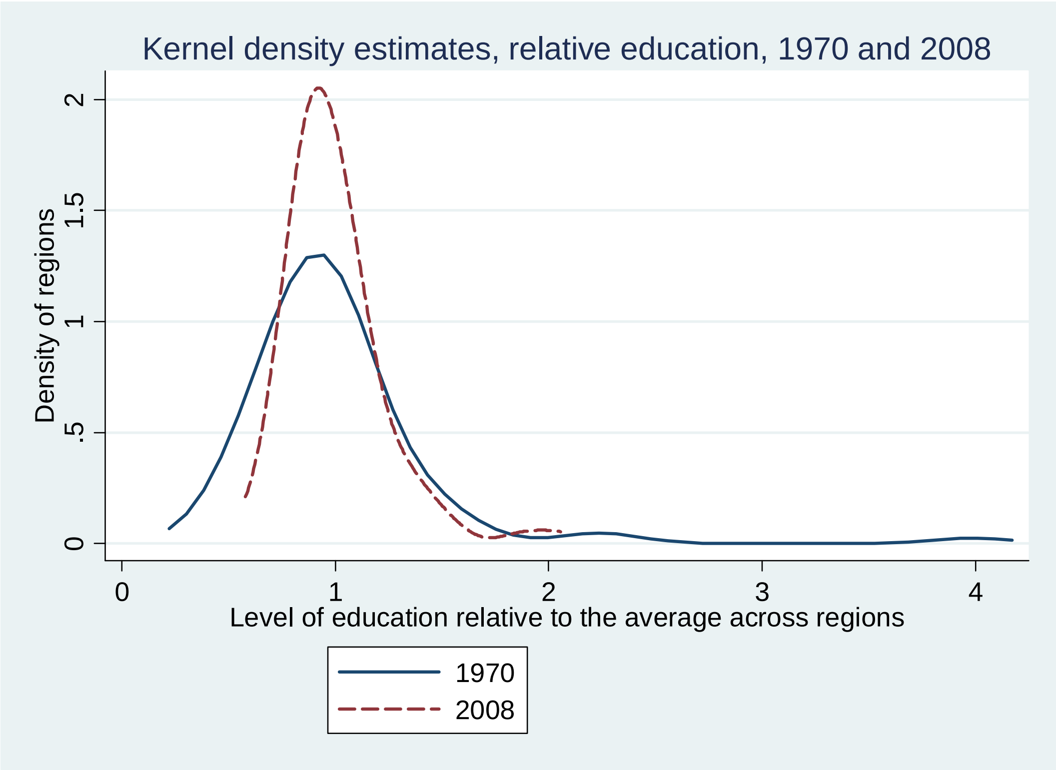

Figure 2: Kernel density estimates, level of education (log), 89 regions, 1970-2008.

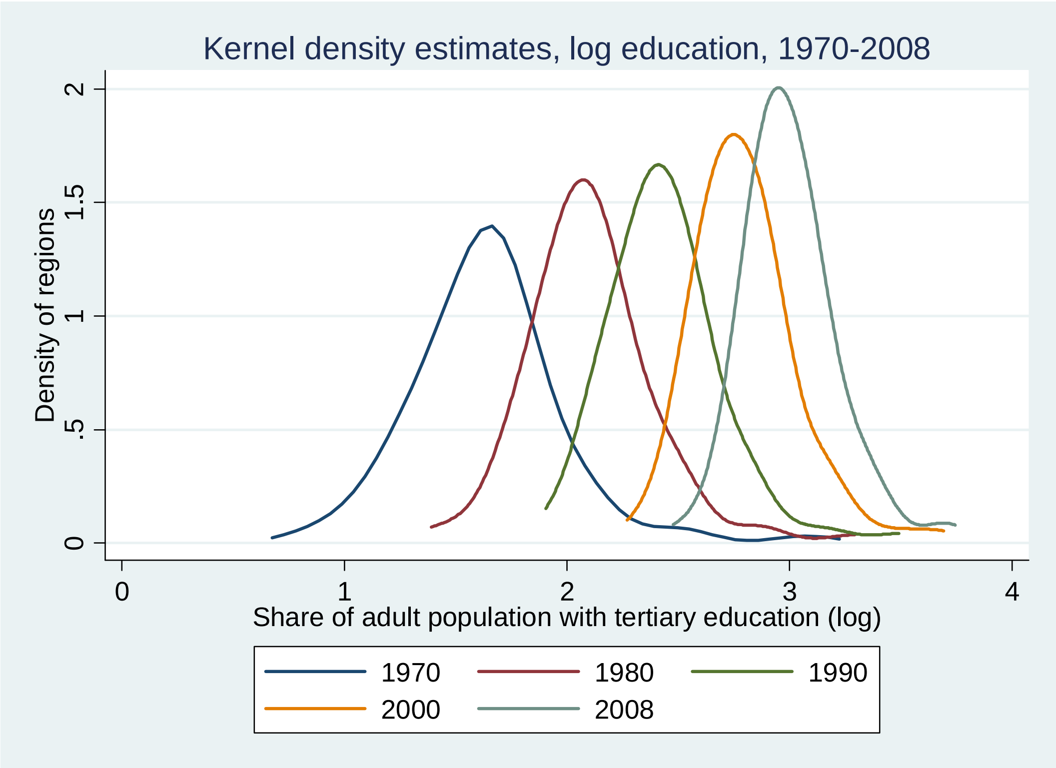

Figure 3: Kernel density estimates, relative income per capita, 89 regions, 1972 and 2008.

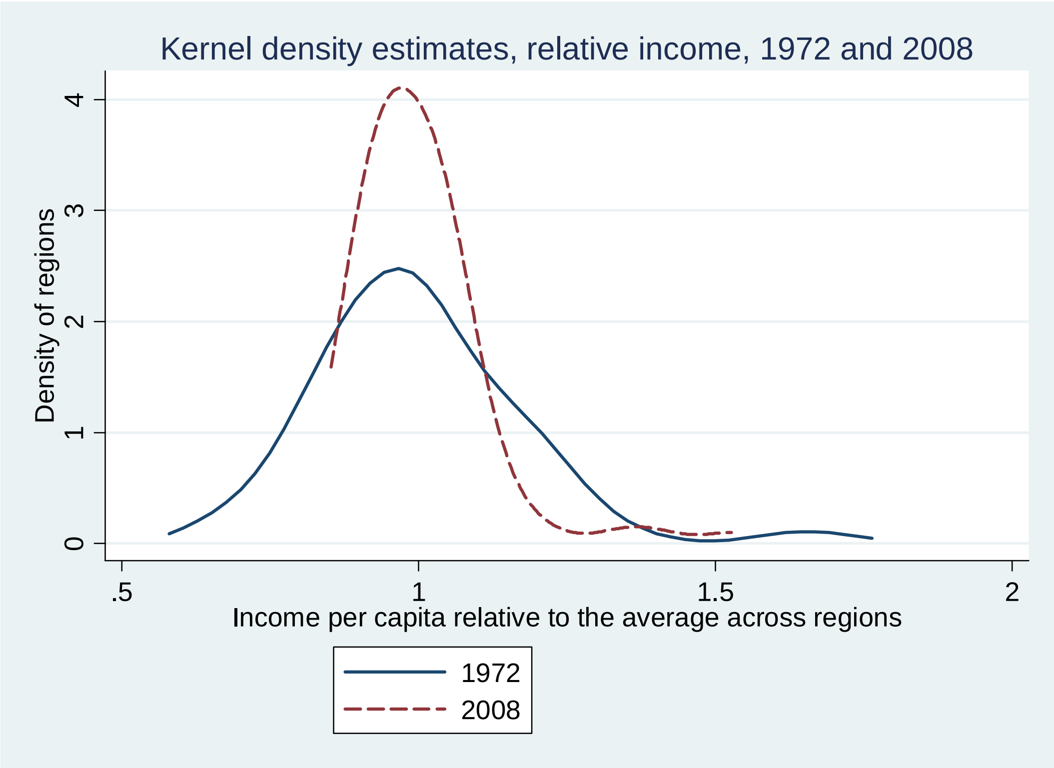

Table 1: Markov chain transition probability matrix, level of education, decade transitions, 1970-2008, 356 observations (binomial standard errors in parentheses).

Education groups	1 ≤ 0.79	2 ≤ 0.9	3 ≤ 1.0	4 ≤ 1.11	5 > 1.11	Obs.
1	**74.7** **(5.2)**	25.3 (5.2)				71
2	5.6 (2.7)	**73.2** **(5.3)**	19.7 (4.7)	1.4 (1.4)		71
3		13.9 (4.1)	**73.6** **(5.2)**	12.5 (3.9)		72
4			14.1 (4.1)	**70.4** **(5.4)**	15.5 (4.3)	71
5				11.3 (3.8)	**88.7** **(3.8)**	71
Initial distribution	20.0	20.0	20.0	20.0	20.0	
Ergodic distribution	3.6	16.2	24.6	23.4	32.2	

Table 2: Markov chain transition probability matrix, income per capita, 4-year transitions, 1972-2008, 801 observations (binomial standard errors in parentheses).

Income groups	1 ≤ 0.91	2 ≤ 0.96	3 ≤ 1.01	4 ≤ 1.07	5 > 1.07	Obs.
1	**76.3**	23.1	0.6			160
	(3.4)	(3.3)	(0.6)			
2	15.6	**61.3**	21.3	1.9		160
	(2.9)	**(3.9)**	(3.2)	(1.1)		
3	0.6	19.4	**56.3**	21.1	0.6	160
	(0.6)	(3.1)	**(3.9)**	(3.2)	(0.6)	
4		1.3	24.4	**63.1**	11.3	160
		(0.9)	(3.4)	**(3.8)**	(2.5)	
5			0.6	18.0	**81.4**	161
			(0.6)	(3.0)	**(3.1)**	
Initial distribution	20.0	20.0	20.0	20.0	20.0	
Ergodic distribution	15.0	21.9	24.3	23.7	15.1	

Table 3: Test of time stationarity of the Markov matrix in Table 2. Contributions of single subperiods to the Pearson test statistics.

Subperiods	Income groups	Number of obs.	1 ≤ 0.91	2 ≤ 0.96	3 ≤ 1.01	4 ≤ 1.07	5 > 1.07	Sum
1972-1984	1	64	0.55	1.56	0.40			2.51
	2	46	0.00	0.28	0.79	0.02		1.09
	3	43	0.27	1.33	0.96	0.09	0.27	2.92
	4	45		0.34	0.80	0.24	0.00	1.38
	5	69			0.43	2.37	0.61	3.41
	Sum							**11.31**
1984-1996	1	48	0.86	3.13	0.30			4.29
	2	58	0.13	0.01	0.04	0.01		0.19
	3	50	1.49	0.75	0.16	1.70	0.32	4.42
	4	61		0.07	2.53	0.32	1.19	4.11
	5	50			0.31	1.00	0.18	1.49
	Sum							**14.50**
1996-2008	1	48	0.00	0.11	1.60			1.71
	2	56	0.18	0.32	0.37	0.00		0.87
	3	67	0.42	2.79	0.19	0.79	0.79	4.98
	4	54		0.68	0.76	0.02	1.41	2.87
	5	42			2.10	0.78	0.30	3.18
	Sum							**13.61**
Pearson test statistic								**39.4**
Critical value at 5% significance level (28 degrees of freedom)								**41.3**

Table 4: Markov chain transition probability matrix, income per capita, 4-year transitions, conditioning on the change in the relative educational level during 1970-2008 (binomial standard errors in parentheses).

Panel a: Top 50% with large increase in the relative educational level (405 observations)

Income groups	1 ≤ 0.91	2 ≤ 0.96	3 ≤ 1.01	4 ≤ 1.07	5 > 1.07	Obs.
1	**78.4**	20.9	0.7			153
	(3.3)	(3.3)	(0.7)			
2	21.9	**57.3**	18.7	2.1		96
	(4.2)	**(5.0)**	(4.0)	(1.5)		
3	1.4	21.9	**54.8**	20.5	1.4	73
	(1.4)	(4.8)	**(5.8)**	(4.7)	(1.4)	
4		1.7	27.1	**59.3**	11.9	59
		(1.7)	(5.8)	**(6.4)**	(4.2)	
5			4.2	25.0	**70.8**	24
			(4.1)	(8.8)	**(9.3)**	
Initial distribution	37.8	23.7	18.0	14.6	5.9	
Ergodic distribution	27.0	25.3	22.1	17.5	8.1	

Panel b: Bottom 50% with large decrease in the relative educational level (396 observations)

Income groups	1 ≤ 0.91	2 ≤ 0.96	3 ≤ 1.01	4 ≤ 1.07	5 > 1.07	Obs.
1	**28.6**	71.4				7
	(17.1)	(17.1)				
2	6.2	**67.2**	25.0	1.6		64
	(3.0)	**(5.9)**	(5.4)	(1.6)		
3		17.2	**57.5**	25.3		87
		(4.0)	**(5.3)**	(4.7)		
4		1.0	22.8	**65.3**	10.9	101
		(1.0)	(4.2)	**(4.7)**	(3.1)	
5				16.8	**83.2**	137
				(3.2)	**(3.2)**	
Initial distribution	1.8	16.2	22.0	25.5	34.6	
Ergodic distribution	1.7	19.3	27.9	31.0	20.1	

Table 5: Test of whether the change in the relative educational level affects transition probabilities. Contributions of single subsamples to the Pearson test statistics.

Change in relative educational level	Income groups	Number of obs.	1 ≤ 0.91	2 ≤ 0.96	3 ≤ 1.01	4 ≤ 1.07	5 > 1.07	Sum
Top 50% large increase (405 obs)	1	153	0.10	0.32	0.00			0.42
	2	96	2.40	0.25	0.28	0.02		2.95
	3	73	0.63	0.24	0.03	0.21	0.63	1.74
	4	59		0.09	0.18	0.14	0.02	0.43
	5	24				0.62	0.31	0.93
	Sum							**6.47**
Bottom 50% large decrease (396 obs)	1	7	2.09	7.06	0.04			9.19
	2	64	3.60	0.37	0.42	0.03		4.42
	3	87	0.55	0.21	0.02	0.18	0.55	1.51
	4	101		0.05	0.11	0.08	0.01	0.25
	5	137				0.11	0.06	0.17
	Sum							**15.54**
Pearson test statistic								**22.0**
Critical value at 5% significance level (13 degrees of freedom)								**22.4**

Table 6: Markov chain transition probability matrix, income per capita, 4-year transitions, conditioning on the average level of education (binomial standard errors in parentheses).

Panel a: Top 50% with high level of education (405 observations)

Income groups	1 ≤ 0.91	2 ≤ 0.96	3 ≤ 1.01	4 ≤ 1.07	5 > 1.07	Obs.
1	**72.4** **(8.3)**	27.6 (8.3)				29
2	8.7 (3.4)	**62.3** **(5.8)**	26.1 (5.3)	2.9 (2.0)		69
3		17.9 (4.3)	**59.0** **(5.6)**	23.1 (4.8)		78
4		2.0 (1.4)	20.0 (4.0)	**69.0** **(4.6)**	9.0 (2.9)	100
5				14.7 (3.1)	**85.3** **(3.1)**	129
Initial distribution	7.2	17.0	19.3	24.7	31.8	
Ergodic distribution	5.9	18.6	26.6	30.4	18.6	

Panel b: Bottom 50% with low level of education (396 observations)

Income groups	1 ≤ 0.91	2 ≤ 0.96	3 ≤ 1.01	4 ≤ 1.07	5 > 1.07	Obs.
1	**77.1** **(3.7)**	22.1 (3.6)	0.8 (0.8)			131
2	20.9 (4.3)	**60.4** **(5.1)**	17.6 (4.0)	1.1 (1.1)		91
3	1.2 (1.2)	20.7 (4.5)	**53.7** **(5.5)**	23.2 (4.7)	1.2 (1.2)	82
4			31.7 (6.0)	**53.3** **(6.4)**	15.0 (4.6)	60
5			3.1 (3.1)	31.3 (8.2)	**65.6** **(8.4)**	32
Initial distribution	33.1	23.0	20.7	15.1	8.1	
Ergodic distribution	24.9	25.9	22.9	17.7	8.5	

Table 7: Test of whether the average level of education affects transition probabilities. Contributions of single subsamples to the Pearson test statistics.

Average education level	Income groups	Number of obs.	1 ≤ 0.9	2 ≤ 0.95	3 ≤ 1.0	4 ≤ 1.09	5 > 1.09	Sum
Top 50% high level (405 obs.)	1	29	0.06	0.25	0.18			0.49
	2	69	2.12	0.01	0.76	0.38		3.27
	3	78	0.49	0.08	0.10	0.00	0.49	1.16
	4	100			0.77	0.53	0.44	1.74
	5	129				0.77	0.24	1.01
	Sum							**7.67**
Bottom 50% low level (396 obs.)	1	131	0.01	0.06	0.04			0.11
	2	91	1.60	0.01	0.58	0.29		2.48
	3	82	0.45	0.08	0.10	0.00	0.45	1.08
	4	60			1.31	0.91	0.75	2.97
	5	32				3.02	0.94	3.96
	Sum							**10.60**
Pearson test statistic								**18.3**
Critical value at 5% significance level (12 degrees of freedom)								**21.0**

Table 8: Top 20 smart regions in Norway

Region	Average level of education 1970-2008	Relative education 1970	Relative education 2008	Relative income 1972	Relative income 2008	Population 2008
Bærum/Asker	32.9	3.98	2.06	1.62	1.53	161 066
Oslo	25.4	2.24	1.88	1.69	1.40	560 484
Follo	22.2	2.24	1.54	1.35	1.21	111 942
Trondheim	19.1	1.60	1.49	1.17	1.05	221 058
Bergen	17.8	1.51	1.36	1.17	1.12	378 818
Tromsø	17.6	1.07	1.42	1.02	0.99	78 320
Stavanger/Sandnes	17.4	1.31	1.34	1.18	1.34	244 118
Lillehammer	17.3	1.32	1.33	1.04	1.01	36 834
Kongsberg	17.3	1.44	1.31	1.15	1.16	30 471
Kristiansand	16.8	1.57	1.26	1.18	1.06	107 473
Tønsberg/Horten	16.2	1.30	1.25	1.24	1.05	112 606
Ørsta/Volda	15.5	1.26	1.21	0.91	0.97	18 569
Lillestrøm	14.9	1.43	1.08	1.28	1.13	185 471
Arendal	14.8	1.29	1.14	1.09	1.02	74 957
Lillesand	14.5	1.16	1.16	1.03	1.06	13 741
Bodø	14.5	1.07	1.16	1.00	0.99	77 782
Levanger/Verdalsøra	14.2	1.13	1.13	0.88	0.89	34 915
Moss	14.0	1.22	1.06	1.23	1.04	54 080
Drammen	14.0	1.10	1.09	1.23	1.10	164 542
Harstad	13.9	1.02	1.13	0.97	0.97	31 026

APPENDIX: Correlation matrix

	Edu level 1970	Edu level 2008	Inc level 1972	Inc level 2008	Edu change	Inc change	Average edu level
Edu level 1970	1						
Edu level 2008	0.887	1					
Inc level 1972	0.803	0.727	1				
Inc level 2008	0.795	0.715	0.806	1			
Edu change	-0.908	-0.611	-0.715	-0.713	1		
Inc change	-0.487	-0.447	-0.795	-0.283	0.428	1	
Average edu level	0.953	0.983	0.778	0.778	-0.740	-0.465	1

Note: Edu level = Share of adult population with tertiary education, Inc level = Income p.c. constant prices, Edu change = Difference between the relative level of education in 2008 and 1970, Inc change = Difference between the relative income level in 2008 and 1972, Average edu level = Average adult tertiary share during 1970-2008.