

 Open access • Journal Article • DOI:10.1353/JHR.2015.0000

Long-Term Intergenerational Persistence of Human Capital: An Empirical Analysis of Four Generations — [Source link](#)

Mikael Lindahl, Mårten Palme, Sofia Sandgren Massih, Anna Sjögren

Institutions: Uppsala University

Published on: 01 Jan 2015 - Journal of Human Resources (University of Wisconsin Press)

Topics: Human capital, Earnings, Social mobility and Educational attainment

Related papers:

- [Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States*](#)
- [A multigenerational view of inequality.](#)
- [The Son Also Rises: Surnames and the History of Social Mobility](#)
- [Social stratification across three generations : New evidence from the Wisconsin Longitudinal Study](#)
- [Human capital and the rise and fall of families](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/long-term-intergenerational-persistence-of-human-capital-an-30raudw86o>

Long-term intergenerational persistence of human capital: an empirical analysis of four generations*

by

Mikael Lindahl, Mårten Palme, Sofia Sandgren Massih and Anna Sjögren[§]

December 19, 2012

Abstract

Most previous studies of intergenerational transmission of human capital are restricted to two generations – parents and their children. In this study we use a Swedish data set which enables us link individual measures of lifetime earnings for three generations and data on educational attainments of four generations. We investigate to what extent estimates based on income data from two generations accurately predict earnings persistence beyond two generations. We also do a similar analysis for intergenerational persistence in educational attainments. We find two-generation studies to severely under-predict intergenerational persistence in earnings and educational attainment over three and four generations.

Keywords: Intergenerational income mobility, Human capital transmission, Multigenerational income mobility

JEL-codes: D31; J62

* We thank Anders Björklund, Sandra Black, Susan Dynarski, Peter Fredriksson, Caroline Hall and Björn Öckert as well as seminar participants at Uppsala, SOFI (Stockholm University), Trondheim, the CESifo 2011 meeting in Munich, the 2011 Nordic Summer Institute in Labor Economics at the Faroe Islands, Conference on the Economics of the Family (INED) in Paris 2011 and EALE Bonn 2012 for valuable comments on previous drafts and Eskil Forsell, Erika Karlenius and Arvid Olovsson for excellent research assistance. Special thanks to Adrian Adermon for help with the data construction and programming.

Mikael Lindahl is a Royal Swedish Academy of Sciences Research Fellow supported by a grant from the Torsten and Ragnar Söderberg Foundation, the Scientific Council of Sweden and the European Research Council [ERC starting grant 241161]. Mårten Palme gratefully acknowledges financial support from the Swedish Council of Social Research. Swedbank has provided financial support for the construction of the dataset.

[§] Mikael Lindahl: Department of Economics, Uppsala University, SE-751 20 Uppsala, Sweden, CESifo, IFAU, IZA and UCLS, E-mail: Mikael.Lindahl@nek.uu.se; Mårten Palme: Department of Economics, Stockholm University, SE-106 91 Stockholm, Sweden and IZA, E-mail: Marten.Palme@ne.su.se; Sofia Sandgren: Department of Economics, Uppsala University, SE-751 20 Uppsala, Sweden; and Anna Sjögren: IFAU, Box 513, SE-751 20 Uppsala, Sweden, UCLS, Uppsala, and SOFI Stockholm University, E-mail: Anna.Sjogren@ifau.uu.se.

1 Introduction

Families often maintain close connections with the grandparent generation and honour the memories of great-grandparents. Many individuals would also admit to strong influences and transmission of different resources beyond their parent generation. Yet, economic analysis of intergenerational links is almost exclusively concerned with the relation between the parent and child generations. Dynamic macroeconomic models of human and physical capital investments, fertility and inequality, as well as models of cultural transmission, focus on the link between two consecutive generations (Diamond, 1965, Becker, Murphy and Tamura, 1990, Galore and Zeira, 1993, Bisin and Verdier, 2000, Mulligan, 1997, and Saez-Marti and Sjögren, 2008). Moreover, empirical studies on intergenerational income mobility, as surveyed in Solon (1999) and Black and Devereux (2010), are with few exceptions restricted to two generations.¹ The Becker-Tomes model – the by far most important model for intergenerational transmission of human capital – relates financial and other resources of the parent generation to the outcome of the child generation.

The fact that generations beyond the parent generation influence individual outcomes has important implications for how we view income inequality at a given point in time, as well as for how we interpret intergenerational transmission of human capital. Income inequality in a mobile society is commonly regarded as more justifiable since an individual's relative economic position is to a larger extent the outcome of the individual's own choices and economic performance than an inheritance from previous generations. A frequently cited example, e.g. Borjas (2009), is based on an initial income difference of 20 percent between two families. If there is an intergenerational correlation of 0.3, we expect only 30 percent of this difference, or 6 percentage points, to remain in the second generation. In the third generation, the difference is almost entirely eliminated, since only 1.8 percent is expected to remain. However, this example relies critically on the assumption that the intergenerational transmission process of human capital has a memory of only one period. If this is not the case, income convergence will take longer.

Clark (2012) provides additional motivation for why only considering two generations may severely underestimate the *long-term* intergenerational persistence in socio-economic status. Clark views socio-economic outcomes across generations as a stochastic process where the outcome of each generation can be viewed as the sum of a systematic and a stochastic part.

¹ Examples of some studies that estimate the relationship between outcomes for grandparents and grandchildren are Behrman and Taubman (1985), Jaeger (2012), Lucas and Kerr (2012), Maurin (2002), Plug and Vijverberg (2005), Sacerdote (2005), Sauder (2006) and Warren and Hauser (1997).

The stochastic part is for instance related to luck on the labor market and variations in the genetic set up between generations. Since we are not able to disentangle the systematic part from the stochastic part in any data set, our estimate of the long-term persistence will be attenuated and thus biased towards zero. This argument resembles the Solon (1992) argument for why using annual rather than lifetime earnings will attenuate estimates of intergenerational earnings correlations. In Clarke's case, however, the solution to the estimation problem requires data from more than two consecutive generations.

Extensions of the empirical analysis of intergenerational transmission of human capital beyond two consecutive generations relate to at least two additional strands in the literature on equality of opportunity and socio-economic mobility across generations. First, as pointed out in Solon (1999) or Björklund et al. (2010), the "explained" variation in models based on siblings correlations is in general much higher than in models based on intergenerational correlations (around 0.3 compared to around 0.1). A plausible interpretation of this difference is that siblings share more characteristics than just parents. The potential influence of grandparents – and great-grandparents – is obviously one of these shared characteristics in addition to the influence of neighborhoods during adolescence, schools, and other environmental factors that siblings in most cases share, which may affect their economic position as adults. Second, extension of the analysis of intergenerational transmission beyond two generations relates to a recent literature which, following Roemer (1993), aims to measure the degree of equality of opportunity; see e.g. Aaberge et al., (2010) or Björklund et al. (2012). Generations beyond the parental generation constitute an obvious "circumstance" that may influence the economic position of the child generation in addition to the investment decisions and endowments of the parent generation, as suggested in the Becker-Tomes model.

The main purpose of this paper is to investigate to what extent conventional estimates of intergenerational mobility based on data from two consecutive generations can predict long-term intergenerational mobility in human capital. First, we simply calculate mobility measures for earnings and educational attainment from two-generation data and compare the predictions with actual outcomes for cohorts two and three generations apart, respectively. We find that mobility is significantly overestimated for education and marginally significantly overestimated for earnings. We also discuss calculations reported by Clark (2012) on parameters for two-generation mobility consistent with our long-term mobility measures. Second, we include outcomes for the grandfather generation in the mobility regressions. We find that they have an independent effect on the outcome of the grandchildren. Again, we find stronger effects for education.

We use an exceptional data set containing measures of lifetime earnings for three consecutive generations and data on educational attainments for four generations. The data set is based on a survey of all third graders in Sweden's third largest city, Malmö, and its suburbs, in 1938. This index generation has subsequently been followed until retirement and information on parents, spouses, children and grandchildren have been added. The first generation was, on average, born in the late nineteenth century and the fourth generation typically completed their education in the early twenty-first century. Altogether there are 901 complete families, i.e., families where education data are available on at least one individual in each of four consecutive generations.

The paper proceeds as follows. In Section 2 we introduce the data set, discuss the construction of variables and provide some descriptive statistics of the variables used in this study. In Section 3 we present descriptive estimations from associating outcomes of children with those of parents, grandparents (income and education) and great-grandparents (education). Section 4 concludes.

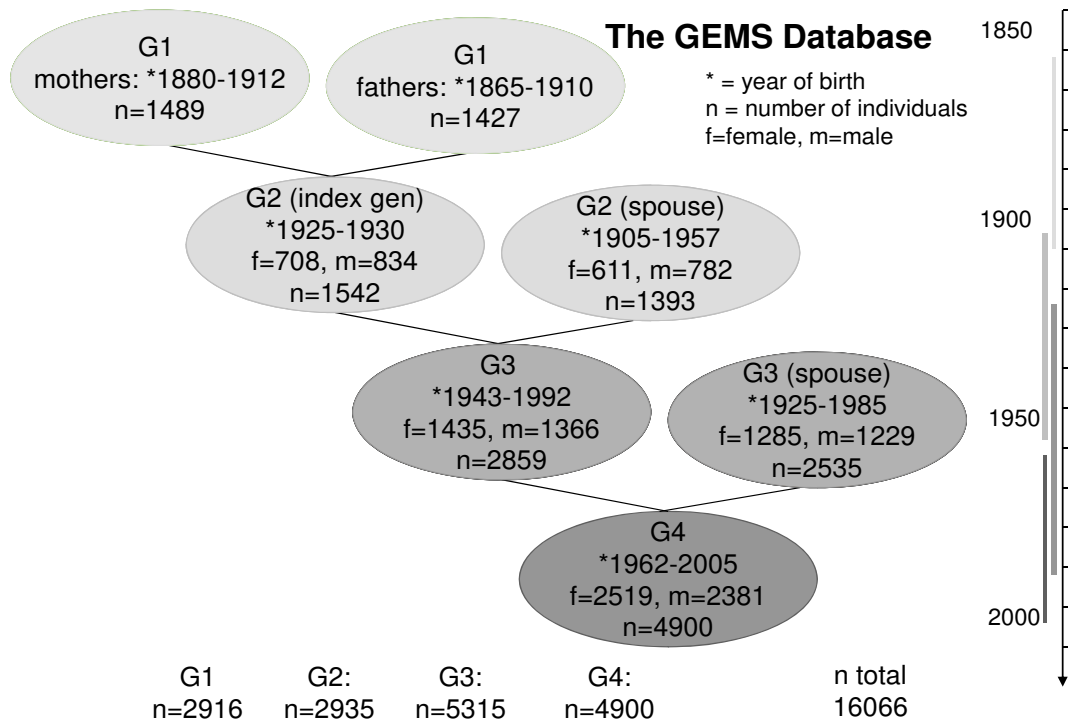
2 Data and descriptive statistics

Figure 1 shows an overview of the data set consisting of information on individuals from four generations of the same family. The data set originally stems from the so called Malmö Study, a survey initiated in 1938 by a team of Swedish educational researchers.² All pupils attending third grade (normally at age 10) in any school in the Malmö metropolitan area (n=1,542) were part of the original survey and constitute the index generation, which is the second generation included in the data set. The original purpose was to analyze the correlation between social surroundings and cognitive ability. Hence, a host of family background information was collected, including parental earnings for several years and father's education. Over the years, the Malmö Study has been extended with information from both several rounds of follow-up surveys and register data.³

² The material was originally collected by Siver Hallgren and developed by Torsten Husén.

³ In 1993, 38% of the third and fourth generations still lived in Malmö, an additional 31% lived elsewhere in the county of Skåne, which is where Malmö is situated, 8% lived in the county of Stockholm, and the rest were quite evenly spread out in the rest of Sweden.

Figure 1: Overview of the GEMS database.



We have extended the data in several ways. We have added parish-register information on date of birth and death of the parents of the index generation. These parents constitute the first generation and were born between 1865 and 1912. We have also added register information on the second generation's children and grandchildren, as well as information on the spouses of the index generation, i.e., the second parent of these children and of the grandchildren. The resulting data set consists of information on four generations of the same families. The average birth year of the first generation (G1) is 1898. The second generation (G2), i.e. the index generation, is on average born in 1928; the third generation (G3), the children of the index generation, in 1956; and, finally, the fourth generation (G4), the grandchildren of the index generation, in 1985.⁴

In the Appendix we provide a short historical account of the development in Malmö and Sweden, focusing on the evolvement of institutions of likely importance to intergenerational mobility and the welfare state in Sweden during the relevant time period.

⁴ In figure 1, the vertical line on the right-hand also shows the spread in birth year for all four generations.

2.1 Data on educational attainment

The measure of educational attainments for the first generation was constructed by educational scientists and based on occupational classification of fathers from a survey in 1938. For the second to fourth generations, we have obtained data on educational attainments from the national education register. We mainly use data from 1985 for the second generation and from 2009 for the third and fourth generations. We transform the educational level measure for all generations into years of schooling based on the required number of years that has to be completed for each level.⁵ In order to avoid the problem that some children in the youngest generation may still have been in school at the time of data collection, we restrict the analysis of years of education to individuals who were at least 25 years of age in 2009, hence excluding those born after 1984.

So as to further increase the sample size for the analysis of education transmission, we construct a measure of whether or not an individual has completed an academic track in high school. This is a strong predictor of whether or not the individual continues on to higher education. We are then able to include children born until 1990. This increases the sample by about 35 percent.

2.2 Measures of lifetime earnings

Detailed earnings information allows us to construct measures of lifetime earnings for men in the first three generations. The fourth generation is not included in the analysis of earnings transmission since a large fraction of these individuals are too young to allow the construction of meaningful measures of lifetime earnings. Although the amount of earnings information differs across generations, available data from local and national tax registers cover the most important years of working life for all generations.

As regards the first generation, born on average in 1896, we have annual income information from local tax registers for the years 1929, 1933, 1937, 1938 and 1942. This implies that income is typically observed between ages 33 and 46. The income measure is the sum of capital and labor income.

⁵ With detailed information on completed level of education, we construct years of schooling as follows: 7 for (old) primary school, 9 for (new) compulsory schooling, 9.5 for (old) post-primary school (*realskola*), 11 for short high school, 12 for long high school, 14 for short university, 15.5 for long university, and 19 for a PhD. For those few individuals in the second generation where registry information for 1985 is missing, we use survey information from 1964. The education information from 1964 is in 6 levels, and probably of lower quality than for 1985 or 2009. The conversion is done by imputing years of schooling by regressing the years of schooling variable in 1985 on indicators for 1964 using all individuals for whom educational information is available in both years. For individuals in the third generation with missing education data, we instead draw on registry information from 2005 and 1985.

The second generation, most of whom were born in 1928 (the original Malmö population) or around 1928 (the other parent of the Malmö children), is covered from age 20 by at least 15 observations of annual earnings. The first observations of labor earnings stem from 1948.⁶ From then on, there is information on earnings every third-fifth year until 1984. After 1984, we have annual observations of earnings.

As for the third generation, typically born in the mid-1950s, earnings data start in 1968. Like the second generation, information on earnings was collected every third-fifth year until 1984, after which there are annual observations.

We compute our earnings measure in two steps. First, using all earnings data available,⁷ we regress log-earnings on a cubic in birth year as well as year dummies, i.e.,⁸

$$\log(\text{earnings})_{it} = \alpha + \gamma_1 \text{birthyear}_i + \gamma_2 \text{birthyear}_i^2 + \gamma_3 \text{birthyear}_i^3 + \text{year}_t + \varepsilon_{it}. \quad (1)$$

Second, we obtain the residual for each individual-year cell it , and then compute the mean residual for each individual, i.e., the stable part of individual earnings, which is used as a measure of lifetime earnings.

2.3 Descriptive statistics

We have information on educational attainments for 901 complete families, i.e., with data available on at least one individual in each generation, for four consecutive generations.⁹ For earnings, there are 730 families with earnings information available for one male member of the family in three consecutive generations. The main reason for attrition of families is that the individual has no children. There are, however, some individuals with missing information on earnings and/or education. Since earnings data are less informative for women in the earlier years, we restrict the analysis of earnings associations to sons, fathers and grandfathers. Note that for roughly half of the earnings sample, the male family member in

⁶ Prior to 1968, information on earnings is from local tax registers. As of 1968, the earnings data are from national registers. For individuals in the second generation who were not part of the original sample, i.e. the other parent of the third generation individuals, we have earnings information from 1948 if they cohabited with the Malmö-parent and from 1968 if they did not.

⁷ We include all years for which we observe positive earnings, but exclude the observations when the individual was very young: 19 years of age for the first generation, 23 for the second and 27 for the third.

⁸ This is the approach taken in e.g. Haider and Solon (2006) and Böhlmark and Lindquist (2006). Life-cycle bias should hence not be an issue here, as we have access to reasonable lifetime income measures for both parents and children. See also Lee and Solon (2009).

⁹ We have 901 complete families with four generations when we include fourth generation children born until 1990. For this sample, the education measure used for the fourth generation is academic high-school track. In order to obtain a meaningful measure of years of education for the fourth generation, we restrict the analysis to children born before 1986, resulting in 673 complete families.

the second generation (the father) is not the biological son of the male member of the family in the first generation (the grandfather), but is instead the son-in-law. This almost doubles the earnings sample.¹⁰

Table 1 reports descriptive statistics by generation and gender for the samples used in this study. We show statistics corresponding to the individuals in our estimation sample for education (four generations separated by gender) and earnings (three generations of men). The first column shows means and standard deviations for the fathers of the children in the index generation (generation 2). These 905 fathers were on average born in 1896 and had 7.3 years of schooling. The next two columns show descriptive statistics for those in the index generation (first interviewed in 1938 and typically born in 1928) as well as mothers and fathers of the children in the third generation. For this second generation, typically born in 1928, there are 470 men who acquired 10.2 years of schooling and 435 women who acquired 9.5 years, on average.¹¹

The earnings figures for men in the second and third generations pertain to sons and grandsons of the first generation of men as well as the male spouse of the daughters and granddaughters belonging to the index and the next generations. The last two columns show descriptive statistics for the descendants of the three earlier generations who are old enough to be included in the regressions: 27 years of age in 2008 for earnings regressions; 25 years of age in 2009 for education estimations; and, finally, 19 years of age for the academic high-school track regressions. The average residual of log earnings, with means and standard deviations reported in the third row, summarizes the earnings measure actually used in the estimations.¹²

¹⁰ As a check, we also estimated transmission coefficients for education using these sample restrictions. The estimates are then very similar to those using only individuals who are biologically related across the four generations (which are the estimates reported in Table 2).

¹¹ On average, earnings increased from about SEK 86,000 (calculated in 1933) for the men in the first generation to SEK 311,000 (in 2000) for the men in the third generation, all expressed in 2010 prices.

¹² These numbers are based on averages across years and are negative because those with fewer years of earnings data have lower earnings.

Table 1: Descriptive statistics

Variable	Generation 1 (great-grandparents)		Generation 2 (grandparents)		Generation 3 (parents)		Generation 4 (children)	
	Great-grandfather (1)	Grandmother (3)	Grandfather (4)	Mother (5)	Father (6)	Daughter (7)	Son (8)	
Years of schooling	7.30 (1.60) [5,14]	9.53 (2.67) [7,19]	10.15 (2.96) [7,20]	12.05 (2.47) [7,20]	12.11 (2.59) [7,20]	12.95 (1.98) [7,20]	12.42 (2.13) [7,20]	
Academic high-school track						0.55 (0.50) [0,1]	0.44 (0.50) [0,1]	
Average residual log earnings	-0.047 (0.529) [-1.74,2.76]		-0.018 (0.637) [-2.71,2.26]		-0.121 (0.763) [-4.11,1.90]			
Year of birth (Education)	1896.12 (7.20) [1859,1910]	1927.91 (0.40) [1925,1930]	1927.87 (0.40) [1926,1929]	1954.67 (4.90) [1944,1970]	1954.53 (4.46) [1943,1969]	1981.45 (6.30) [1962,1990]	1981.49 (6.35) [1962,1990]	
Year of birth (Earnings)	1895.70 (7.48) [1865,1910]		1926.73 (3.27) [1888,1947]		1956.69 (5.54) [1943,1981]			
Number of observations (Education)	905	435	470	831	722	1,451	1,548	
Number of observations (Earnings)	803		1,174		1,174			

Notes: The first figure in each cell is the mean of the variable. The figure in parenthesis is the standard deviation and the figures in square brackets are minimum and maximum values, respectively. The education statistics are calculated for the observations used in Table 2 (column 1) and Table 3 (columns 1-2) and the earnings statistics are calculated for the observations used in Table 5. The statistics for year of schooling for generation 4 are calculated for those born before 1985 (887 daughters and 936 sons)

3 Results: Intergenerational persistence in educational attainments and earnings

3.1 Intergenerational persistence in educational attainments

The first set of results, the estimated transmission coefficients for education across the four generations under study, are shown in Table 2. All estimates are results from the bivariate regression model

$$y_t = a + by_{t-j} + u_t, \quad (2)$$

where $j \geq 1$, y_t is the outcome of the child and y_{t-j} is outcome of the parent ($j=1$), grandparent ($j=2$) or great-grandparent ($j=3$). Since many members of the last generation had not yet completed their education at the date of data collection, we use completion of an academic track in secondary school as a proxy for educational aspiration. The last row in Table 2 reports linear probability model estimates of the relation between the probability of having completed an academic high-school track and earlier generations' educational attainments measured in years of education. The estimates (standard errors) are outcomes from regressions using unstandardized variables. We report standardized estimates in brackets.

Table 2 reveals several interesting results. First, there is a statistically significant estimate for the association between great-grandfather's educational attainment and that of great-grandchildren. This result shows that there is a persistent correlation despite the fact that there are two generations, or on average 75 years, between the births of these generations. Second, the association between educational outcomes of the great-grandparent generation and the child generation, as well as between the great-grandparent generation and the parent generation is stronger than what would be expected if we were to predict these correlations based on the correlation between the adjacent generations involved.

The second result is easily obtained by multiplying the diagonal elements in Table 2. For example, multiplying the coefficient estimate between the first and second generations, 0.607, by that between the second and third, 0.281, yields a prediction for the association between the first and third generations of 0.171. By applying the delta method we obtain approximate bounds for the standard error of this prediction of between 0.023 and 0.033.¹³ These

¹³The approximation of the variance for the product of β_1 and β_2 , where β_1 is the estimate between generation one and two and β_2 is the estimate between generation two and three, is $\beta_2^2 \sigma_{\beta_1}^2 + \beta_1^2 \sigma_{\beta_2}^2 + 2\beta_1\beta_2\sigma_{\beta_1\beta_2}$. Since we

approximate bounds enable us to formally test and reject that the prediction obtained is equal to the coefficient between the first and the third generation, which was estimated to be 0.375, which is more than twice the magnitude of the predicted association.

Over the studied generations, the education distribution has changed dramatically as education has expanded from a mean of seven years of schooling in the first generation to almost thirteen in the fourth generation. In Table 2, we therefore also report standardized estimates in square brackets. Interestingly, these estimates suggest that much of the decline in the intergenerational education persistence between the first two and the second two generations from 0.607 to 0.281, is in fact, due to changes in the education distribution. The standardized estimates are instead rather stable over time. However, also if we use these standardized estimates we can reject that the predicted relation between the first and the third generation is the same as the estimated standardized coefficient.¹⁴

Table 2: Matrix of estimated transmission coefficients across generations: Education

	Years of Schooling – great grandparent (1)	Years of Schooling – grandparent (2)	Years of Schooling – parent (3)
Years of Schooling – grandparent	0.607*** (0.065) [0.334] N=905		
Years of Schooling – parent	0.375*** (0.043) [0.229] N=1553	0.281*** (0.024) [0.312] N=1553	
Years of Schooling – child	0.145*** (0.046) [0.123] N=1823	0.131*** (0.023) [0.202] N=1823	0.296*** (0.021) [0.412] N=1823
Academic HS track (=1) – child	0.032*** (0.007) [0.104] N=2999	0.028*** (0.004) [0.163] N=2999	0.066*** (0.004) [0.343] N=2999

Notes: Each reported estimate is from a separate regression of the education of members of one generation on the education of members of an older generation. All regressions control for a quadratic in the birth year of the member of both generations. The reported standard errors (in parentheses) are clustered on families. Standardized estimates are reported in brackets.

are not able to estimate the covariance term $\sigma_{\beta_1\beta_2}$, we instead use the estimates of σ_{β_1} , σ_{β_2} and the fact that the maximum correlation coefficient value is 1 to obtain an upper bound for $\sigma_{\beta_1\beta_2}$. The lower bound for $\sigma_{\beta_1\beta_2}$ is set to 0.

¹⁴ We can do the same exercise for the third and fourth generation. Also then can we reject that the predicted association which is almost 60% higher is the same as the estimated association, at least when we use the standardized estimates. With the level estimates we are able to reject the null under favorable covariance assumptions, but fail to reject the null under the most unfavorable covariance assumptions.

As we described in the introduction, there are at least two alternative interpretations of the result that we cannot predict the intergenerational association between the grandparent and the child generation with the correlations between the two adjacent generations. First, there may be a direct relation between the generations beyond the parents' and the child generation. Second, as described by Clark (2012), the outcome of each generation may be seen as an outcome of a stochastic process, which will bias the long-term intergenerational correlation downwards.

Let us begin by investigating the first interpretation. Table 3 reports intergenerational correlation estimates where we have extended the simple bivariate regressions with the educational outcome for the grandparent generation, i.e.,

$$y_t = a + b_1 y_{t-1} + b_2 y_{t-2} + u_t. \quad (3)$$

The first two columns show the results for years of schooling: the first column for generations 1 through 3 and the second column for generations 2 through 4. The third column shows the results from a linear probability model that members of the fourth generation acquire degree from an academic high school track. The intergenerational schooling estimates in the first row of Table 3 (which are conditional on grandparent's education) can be related to, respectively, the following unconditional estimates in Table 2: 0.281, 0.296 and 0.066. The results show unambiguously that the coefficient for the grandparent generation is significantly different from zero. Using results from consecutive generations thus underestimates the long-term intergenerational persistence.

Table 3 OLS regressions of children's education on parent's and grandparent's education

Dependent variable:	Years of Schooling		Academic high-school track
Dependent variable is for generation:	3	4	4
<i>Main equation:</i>	Education of child		
Schooling of parent	0.240*** (0.027) [0.267]	0.266*** (0.023) [0.371]	0.060*** (0.004) [0.315]
Schooling of grandparent	0.180*** (0.046) [0.110]	0.062*** (0.021) [0.096]	0.011*** (0.004) [0.069]
N	1,553	1,823	2,999
R2	0.164	0.205	0.143

Notes: All regressions control for a quadratic in the birth year for all three generations. The reported standard errors (in parentheses) are clustered on families. Standardized estimates are reported in brackets.

An alternative explanation to the results that the estimated correlation between consecutive generations fails to predict the association two generations apart is that these estimates are biased downwards, since the outcome for each generation can be seen as an outcome with a stochastic component (Clark, 2012). The argument is built on two observations: First, any measured outcome of, for instance education, is an imperfect measure of long-run educational attainment. Clark (2012) suggest the following relation, $y_{t+n} = b_n y_t + e_{t+n}$, where e represents the generation specific stochastic part of the outcome. Second, that the product of two attenuated estimates, both from intergeneration estimates using two adjacent generations (for instance generations 1 on 2, and 2 on 3), will be more attenuated than an intergenerational estimate using non-adjacent generations (for instance generation 1 on 3). Clark (2012) combines these observations to rationalize our results to calculate estimates of long-term social mobility. He finds that an inter-generational correlation between two generations on 0.7 rationalizes our long-term estimates. This estimate is indeed substantially larger than our estimates and other conventional estimates of intergenerational persistence in socio-economic outcomes.

In Tables 2 and 3 we pooled data on men and women. This would be problematic if intergenerational transmission of education was very different for daughters and sons and from mothers and fathers. In Table 4 we provide reassuring support for pooling. This table reports the results from estimations of the intergenerational transmission coefficients separately by gender of offspring and ancestor. The most striking feature of these estimates is that the intergenerational correlation in educational attainments seems to be independent of the gender of both ancestor and offspring. For example, the correlation between the first and third generations is almost the same for males and females in the third generation.¹⁵

¹⁵ If we compare the estimates from using parents and children (i.e., for the third and fourth generation) in this Malmö-sample to estimates from using a random sample of Swedish parents born 1943-1955 (Holmlund, Lindahl and Plug, 2011), the estimates in this paper are only slightly higher.

Table 4: Matrix of estimated transmission coefficients across generations: Years of education

	Great-grandfather	Grandmother	Grandfather	Mother	Father
Years of Schooling – grandmother	0.565*** (0.076) [0.311] N=435				
Years of Schooling – grandfather	0.661*** (0.118) [0.364] N=470				
Years of Schooling – mother	0.344*** (0.049) [0.210] N=831	0.287*** (0.047) [0.319] N=415	0.273*** (0.039) [0.303] N=416		
Years of Schooling – father	0.409*** (0.060) [0.250] N=722	0.322*** (0.057) [0.357] N=335	0.249*** (0.048) [0.277] N=387		
Years of Schooling – daughter	0.159*** (0.062) [0.135] N=887	0.135*** (0.043) [0.208] N=461	0.117*** (0.040) [0.181] N=426	0.305*** (0.039) [0.425] N=556	0.228*** (0.041) [0.318] N=331
Years of Schooling – son	0.133** (0.052) [0.113] N=936	0.118*** (0.041) [0.183] N=483	0.146*** (0.042) [0.226] N=453	0.306*** (0.042) [0.426] N=521	0.328*** (0.035) [0.458] N=886
Academic HS track – daughter	0.035*** (0.009) [0.112] N=1451	0.022*** (0.008) [0.129] N=713	0.030*** (0.007) [0.172] N=738	0.069*** (0.007) [0.358] N=815	0.055*** (0.008) [0.289] N=636
Academic HS track – son	0.029*** (0.010) [0.093] N=1548	0.030*** (0.008) [0.176] N=747	0.028*** (0.007) [0.160] N=801	0.066*** (0.007) [0.343] N=829	0.071*** (0.006) [0.368] N=719

Notes: Each reported estimate is from a separate regression of the education of members of one generation on the education of members of an older generation. All regressions control for a quadratic in the birth year of the member of both generations. The reported standard errors (in parentheses) are clustered on families. Standardized estimates are reported in brackets.

Changes in education distributions not only affect intergenerations transmission coefficients, but also alter the meaning of a particular number of years of education over time and possible non-linearities in the transmission process are not fully captured in the linearly estimated transmission coefficients. We therefore compute intergenerational transmission probabilities across education categories and corresponding odds ratios. The results are reported in Table 5a-5d. For each generation we define four levels of education, from compulsory to university education.

Transition probabilities and odds ratios confirm the main result from Table 2, namely that there is substantial persistence across generations in the education level attained. In particular, Table 5c shows that there is a higher probability that an individual has university education if the ancestor, three generations before had education above the compulsory level at the time. In addition, these transition probabilities indicate a presence of non-linearities: there is higher persistence in the upper end of the education distribution. Those with more than compulsory education in the first generation are on average between 49 and 67 percent more likely, compared to random assignment, to have university educated great-grandchildren, whereas those with only compulsory schooling are only 3 percent more likely than random assignment to have great grandchildren with compulsory schooling.

Table 5a: Education of children (generation 2) and education of parents (generation 1), transition probabilities and odds ratios

Education of parents (generation 1)		Education of children (generation 2)				All
		Compulsory	Post compulsory : short or vocational	High school	University	Pi. Obsi.
Compulsory	P_{1j}	0.50	0.32	0.14	0.04	0.85
	$P_{1j}/P_{.j}$	1.12	1.01	0.86	0.54	765
Post compulsory: Some vocational	P_{2j}	0.23	0.31	0.31	0.16	0.08
	$P_{2j}/P_{.j}$	0.50	0.99	1.85	2.19	75
Post compulsory: Academic (short)	P_{3j}	0.08	0.32	0.30	0.30	0.04
	$P_{3j}/P_{.j}$	0.18	1.04	1.79	4.08	37
High school/ University	P_{4j}	0.11	0.18	0.25	0.46	0.03
	$P_{4j}/P_{.j}$	0.24	0.58	1.51	6.37	28
All	$P_{.j}$	0.45	0.31	0.17	0.07	
	Obs _{.j}	408	281	150	66	905

Notes: Education generation 1: compulsory max 8 years, post-compulsory: vocational 9 years, post-compulsory: academic (*Realskola*) 10 years, high school or university: min 12 years

Education generation 2: compulsory max 9 years, post-compulsory: short academic or vocational high-school track (*Realskola* or short high-school track) 10-11 years, academic high-school track 12-14 years, university: min 15 years.

Table 5b: Education of grandchildren (generation 3) and education of grandparents (generation 1), transition probabilities and odds ratios

Education of grandparents (generation 1)		Education of grandchildren (generation 3)				All
		Compulsory	Post compulsory : short or vocational	High school	University	P_i Obs _i
Compulsory	P_{1j}	0.20	0.40	0.26	0.15	0.85
	$P_{1j}/P_{.j}$	1.08	1.09	0.98	0.79	1317
Post compulsory: Some vocational	P_{2j}	0.13	0.23	0.30	0.34	0.08
	$P_{2j}/P_{.j}$	0.69	0.64	1.11	1.84	128
Post compulsory: Academic (short)	P_{3j}	0.10	0.15	0.33	0.42	0.04
	$P_{3j}/P_{.j}$	0.55	0.41	1.24	2.23	60
High school/ University	P_{4j}	0.02	0.13	0.29	0.56	0.03
	$P_{4j}/P_{.j}$	0.12	0.34	1.09	3.01	48
All	$P_{.j}$	0.18	0.37	0.27	0.19	
	Obs _j	280	567	416	290	1553

Notes: Education generation 1: compulsory max 8 years, post-compulsory: vocational 9 years, post-compulsory: theoretical (*Realskola*) 10 years, high school or university: min 12 years.

Education generation 3: compulsory max 9 years, post-compulsory: short academic or vocational high-school track (*Realskola* or short high-school) 10-11 years, academic high-school track, 12-14 years, university: min 15 years.

Table 5c: Education of great-grandchildren (generation 4) and education of great-grandparents (generation 1), transition probabilities and odds ratios. (families with 4th generation born before 1985)

Education of great-grandparents (generation 1)		Education of great-grandchildren (generation 4)				All
		Compulsory	Post compulsory : short or vocational.	High school	University	P_i Obs _i
Compulsory	P_{1j}	0.10	0.16	0.50	0.24	0.89
	$P_{1j}/P_{.j}$	1.03	1.05	1.01	0.93	1620
Post compulsory: Some vocational	P_{2j}	0.09	0.07	0.46	0.38	0.07
	$P_{2j}/P_{.j}$	0.93	0.43	0.94	1.49	121
Post compulsory: Academic (short)	P_{3j}	0.04	0.13	0.40	0.43	0.03
	$P_{3j}/P_{.j}$	0.43	0.82	0.82	1.67	47
High school/ University	P_{4j}	0.06	0.11	0.43	0.40	0.02
	$P_{4j}/P_{.j}$	0.58	0.74	0.87	1.57	35
All	$P_{.j}$	0.10	0.16	0.49	0.25	
	Obs _j	179	283	897	464	1823

Notes: Education generation 1: compulsory max 8 years, post-compulsory: vocational 9 years, post-compulsory: academic (*Realskola*) 10 years, high school or university: min 12 years.

Education generation 4: compulsory max 9 years, post-compulsory: short academic or vocational high-school track (*Realskola* or short high-school) 10-11 years, Academic high-school track 12-14 years, university: min 15 years.

Table 5d: Education of great-grandchildren (generation 4) and education of great-grandparents (generation 1), transition probabilities and odds ratios. (families with 4th generation born before 1990)

Education of great-grandparents (generation 1)	Education of great-grandchildren (generation 4)			All
		Compulsory or vocational high-school track	Academic high-school track	P_i Obs _i
Compulsory	P_{1j}	0.53	0.47	0.86
	$P_{1j}/P_{.j}$	1.04	0.95	2567
Post compulsory: Some vocational	P_{2j}	0.40	0.60	0.08
	$P_{2j}/P_{.j}$	0.79	1.22	238
Post compulsory: Theoretical (short)	P_{3j}	0.38	0.62	0.04
	$P_{3j}/P_{.j}$	0.75	1.26	111
High school/ University	P_{4j}	0.29	0.71	0.03
	$P_{4j}/P_{.j}$	0.57	1.44	83
All	$P_{.j}$	0.51	0.49	
	Obs _j	1521	1478	2999
		0.53	0.47	0.86

Notes: Education generation 1: compulsory max 8 years, post-compulsory: vocational 9 years, post-compulsory: theoretical (*Realskola*) 10 years, high school or university: min 12 years
Education generation 4: Compulsory or vocational high-school track, academic track measured at earliest age 19

3.2 Intergenerational persistence in earnings

Table 6 shows the estimates of intergenerational earnings mobility between the first and second generations, the second and third generations as well as between the first and third generations, respectively. Although Swedish society has undergone extensive and important changes in different dimensions between the most active period of the first generation born around 1900 and the third generation mostly born in the 1950s and 1960s, the elasticities in earnings between consecutive generations seem to be quite stable: 0.356 between the first and second generations and 0.303 between the second and third. The latter elasticity is only slightly larger compared to estimates in previous studies for Sweden for children born in similar years (see e.g. Björklund, Lindahl and Plug, 2006).

The results in Table allow us to predict the earnings mobility between the first and third generations from the two two-generation mobility measures. This gives us a prediction of 0.108, which is substantially lower than the estimate of 0.184 obtained from data. Again applying the bounding exercise for the delta method (as explained in footnote 13) gives an estimate of the standard error ranging from 0.020 to 0.027. A *t*-test of equality between the predicted and the estimated three-generation mobility

measure gives a *t*-statistic between 1.47 and 1.58, i.e., indicating a marginally significant difference.

Table 6: Matrix of estimated transition coefficients across generations: log earnings of male offspring regressed on log earnings of male ancestor

Offspring	Grandparent Generation 2	Ancestor	Parent Generation 3
Log(Earnings) – parent Generation 3	0.356*** (0.040) [0.307] N=803		
Log(Earnings) – child Generation 4	0.184*** (0.044) [0.141] N=1174		0.303*** (0.043) [0.268] N=1174

Notes: Each reported estimate is from a separate regression of the son's residual log earnings on residual log earnings of the ancestor. The earnings measures are average residual log-earnings from a regression of log earnings on a cubic in birth year and year dummies (see section 2). The reported standard errors (in parentheses) are clustered on families. Standardized estimates are reported in brackets.

Another way to test the influence of generations beyond the parents is, as we did for the education outcome, to include the outcome of the grandfather generation directly into the regression. The result from this exercise shown in Table 7 suggests that there is a positive effect of grandparent generation earnings and the result is significant on the 10 percent level. Although only marginally significant, this result suggests that persistence is underestimated in studies only including earnings from two generation.

Table 7: OLS regressions of son's log earnings on the log earnings of father's and grandfather's education

Dependent variable: Dependent variable is for generation:	Log(Earnings)-child 3
Log(Earnings) – parent	0.281*** (0.045) [0.249]
Log(Earnings) – grandparent	0.084* (0.044) [0.064]
N	1,174
R2	0.067

Notes: The estimates are from a regression of the son's residual log earnings on the residual log earnings of the father and the grandfather. The earnings measures are average residual log-earnings from a regression of log earnings on a cubic in birth year and year dummies (see section 2). The reported standard errors (in parentheses) are clustered on families. Standardized estimates are reported in brackets.

Parallel to the calculations for persistence in educational outcomes, Clark (2012) also calculated the long-term intergenerational correlation that would generate the results shown in Table 6. The point estimate for this parameter is 0.49, although the standard errors are so large that an estimate on 0.7 could not be rejected.

As in the case of education, it is interesting to explore whether our linear summary measure hides non-linearities in the transmission of earnings across generations. We examine this by means of transition matrices. Table 8 shows transition matrices for income quintiles across generations. The first panel reports the transition probabilities between the first and second generations; the second panel the corresponding figures for the second and third generations; finally, the third panel shows the transitions between the first and third generations.

There is one result of particular interest revealed in Table 8. The persistence across two consecutive generations is higher at the higher end of the income distribution. The highest persistence in all of three panels is found for the fifth quintile, i.e. the top 20 percent of the earnings distribution. As many as 34 percent of the grandchildren of those in the fifth quintile remain at the very top of the income distribution. Interestingly, the persistence in this cell is almost as high when we compare grandfathers and grandsons (first and third generations) as when the grandsons are instead compared to their fathers (second and third generations).

Table 8: Transition matrices: offspring earnings quintile conditional on ancestor's earnings quintile.

Earnings quintile of ancestor		Earnings quintile of offspring				
		Fathers – Generation 3				
Grandfathers Generation 2	Q1	Q2	Q3	Q4	Q5	
Q1	0.30	0.29	0.21	0.11	0.10	
Q2	0.25	0.20	0.20	0.23	0.11	
Q3	0.16	0.20	0.26	0.22	0.17	
Q4	0.16	0.18	0.21	0.27	0.18	
Q5	0.14	0.14	0.11	0.18	0.44	
		Sons- Generation 4				
Fathers Generation 3	Q1	Q2	Q3	Q4	Q5	
Q1	0.31	0.26	0.17	0.18	0.09	
Q2	0.20	0.24	0.19	0.20	0.18	
Q3	0.21	0.18	0.28	0.18	0.15	
Q4	0.15	0.18	0.22	0.22	0.23	
Q5	0.14	0.15	0.14	0.23	0.35	
		Sons Generation 4				
Grandfathers Generation 2	Q1	Q2	Q3	Q4	Q5	
Q1	0.19	0.23	0.21	0.24	0.14	
Q2	0.23	0.22	0.23	0.17	0.14	
Q3	0.25	0.19	0.20	0.20	0.16	
Q4	0.17	0.20	0.21	0.20	0.22	
Q5	0.16	0.16	0.16	0.18	0.34	

Notes: Fathers and sons; 774 families

If we briefly summarize the results from our descriptive estimations, they point toward a surprisingly strong association between grandparental education and earnings and education and earnings of grandchildren, and between great-grandparental education and education of great-grandchildren. Hence, regression toward the mean takes longer time in Sweden than suggested by the comparatively low estimates of intergenerational persistence found for two consecutive generations. In addition, transition matrices reveal that there is higher persistence at the upper end of the education and income distributions. We also find that simply taking the square of the intergenerational elasticity does not give an accurate picture of what we find using children and grandparents, suggesting that the basic assumption that intergenerational transmission follows an AR(1) process does not hold.

4 Conclusions

We have explored intergenerational transmission of economic status across adjacent and distant generations over the span of a century. Our data enable us to link great-grandparents born at the end of the nineteenth century to great-grandchildren who finished their education in the early twenty-first century. We estimate intergenerational correlations in educational attainments between these generations and income correlations between the first generation and their grandchildren.

Our most important finding is that the persistence of inequality across generations is stronger than we would expect from the numerous studies on mobility in earnings and educational attainments based on only two generations. This result has implications for how we should interpret results from previous studies on intergenerational mobility. First, our results show that intergenerational mean reversion takes longer time than we previously knew. Second, simple extrapolations from correlations between adjacent generations to more distant generations, as suggested in elementary text books on labor economics, such as Borjas (2009), are shown to be misleading.

We reach our finding in three different ways. First, we compare predictions based on correlation between consecutive generations with actual correlations estimates on data for generations further apart. We always find that our predictions using adjacent generations underestimate the persistence we find using regressions across three and four generations. Second, we simply add outcomes for the grandfather generation to the regression. For both education and earnings we find that outcomes for grandparents enter significantly in the regression, conditional on the same outcomes for parents. Third, we report the estimates obtained by Clarke (2012) on the long-term intergenerational persistence based on our results reported in a previous WP version of this paper. For education the results are highly significant and the magnitudes of the implied downward bias in the estimated intergenerational persistence from conventional two-generation studies are economically important. The estimated long-term intergenerational correlation is twice as large, 0.7, compared to average on 0.35 for the estimates on consecutive generations. For earnings our estimates are less precise and results only marginally significant, although the point estimates for the implied importance of the downward bias is quite important.

The result that estimates of intergenerational persistence based on siblings correlations are typically substantially higher than those originating from actually correlating the outcomes from two adjacent generation, has been debated as a puzzling fact in the recent literature on intergenerational mobility. (Björklund and Salvanes, 2011) The additional persistence in siblings correlations has been attributed to that siblings share not only parents, but also schools and environmental factors during adolescence that we know are important for socio-economic outcomes later in life. However, our results show that at least some of the discrepancy is attributed to that conventional estimates of intergenerational persistence based on two generations are biased downwards.

References

- Aaberge, R., M. Mogstad and V. Peragine (2010), “Measuring Long-term Inequality of Opportunity”, *Journal of Public Economics* 95, 193-204.
- Becker, G. S. and N. Tomes (1979), “An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility”, *Journal of Political Economy* 87(6), 1153-89.
- Becker, G. S. and N. Tomes (1986), “Human Capital and the Rise and Fall of Families” *Journal of Labor Economics* 4(3), S1-39.
- Becker, G. S., K. M. Murphy, and R. Tamura (1990), “Human Capital, Fertility, and Economic Growth”, *Journal of Political Economy* 98(5), S12-37.
- Behrman, J. R. and P. Taubman (1985), “Intergenerational Earnings Mobility in the US and a Test of Becker’s Intergenerational Endowments Model”, *Review of Economics and Statistics* 67, 144-151.
- Bentzel, R. (1952), *Inkomstfördelningen i Sverige* [The Income Distribution in Sweden], Stockholm: IUI.
- Bisin A. and T. Verdier (2000), “Beyond the Melting Pot: Cultural Transmission, Marriage and the Evolution of Ethnic and Religious Traits”, *Quarterly Journal of Economics* CXV, 955-88.
- Björklund, A., L. Lindahl and M. J. Lindquist (2010), “What More Than Parental Income, Education and Occupation? An Exploration of What Swedish Siblings Get from Their Parents”, *The B.E. Journal of Economic Analysis & Policy* 10(1), article 102.
- Björklund, A., M. Jäntti and R. E. Roemer (2012), “Equality of Opportunity and the Distribution of Long-Run Income in Sweden”, *Social Choice and Welfare*, forthcoming.
- Björklund, A., M. Lindahl and E. Plug (2006), “The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data”, *Quarterly Journal of Economics* 121(3), 999–1028.
- Björklund, A. and K. G. Salvanes (2011), “Education and Family Background: Mechanisms and Policies”, *Handbook in Economics of Education*, Amsterdam: Elsevier, 201-247.
- Black S., P. Devereux and K. G. Salvanes, (2005a), “The More the Merrier? The Effect of Family Size and Birth Order on Children’s Education”, *Quarterly Journal of Economics* 120(2), 669-700.
- Black, S. and P. Devereux (2010), “Recent Developments in Intergenerational Mobility”, in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Vol. 4B, Ch. 16, Amsterdam: Elsevier.

- Black, S., P. Devereux and K. G. Salvanes (2005b), “Why the Apple Doesn’t Fall Far: Understanding Intergenerational Transmission of Human Capital”, *American Economic Review* 95(1), 437-449.
- Böhlmark, A. and M. J. Lindquist (2006), “Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden”, *Journal of Labor Economics* 24(4), 879-900.
- Borjas, G. J. (2009), *Labor Economics*, 5th edition. New York: Irwin/McGraw-Hill.
- Clark, G. (2012) “What is the True Rate of Social Mobility in Sweden? A Surname Analysis, 1700-2012”. Mimeo, University of California, Davis.
- Diamond, P. A. (1965), “National Debt in a Neoclassical Growth Model”, *American Economic Review*, 55, 1126–1150.
- Galor, O. and J. Zeira (1993), “Income Distribution and Macroeconomics”, *Review of Economic Studies* 60(1), 35-52.
- Haider, S. and G. Solon (2006), “Life-Cycle Variation in the Association between Current and Lifetime Earnings”, *American Economic Review* 96(4), 1308-1320.
- Holmlund, H., M. Lindahl and E. Plug (2011), “The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods”, *Journal of Economic Literature* 49(3), 614-650.
- Jaeger, Mads Meier (2012), “The Extended Family and Children’s Educational Success”, *American Sociological Review* 77(6), 903-922.
- Lee, C. and G. Solon (2009), “Trends in Intergenerational Income Mobility”, *Review of Economics and Statistics* 91(4), 766-772.
- Lefgren, L., M Lindquist, och D. Sims (2011), “Rich Dad, Smart Dad: Decomposing the Intergenerational Transmission of Income,” Mimeo, SOFI, Stockholm University, October 2011.
- Lucas, R. E. B. and S. Pekkala Kerr (2012), Inter-Generational Income Immobility in Finland: Contrasting Roles for Parental Earnings and Family Income”, *Journal of Population Economics*
- Maurin, E. (2002), “The Impact of Parental Income on Early Schooling Transitions: A Re-examination Using Data over Three Generations”, *Journal of Public Economics* 85(3), 301-332.
- Mulligan, C. B. (1997), *Parental Priorities and Economic Inequality*, Chicago: University of Chicago Press.
- Plug, E. and W. Vijverberg (2005), “Does Family Income Matter for Schooling Outcomes? Using Adoptees as a Natural Experiment”, *Economic Journal* 115(506), 880-907.

- Roemer, J. E. (1993), "A Pragmatic Theory of Responsibility for the Egalitarian Planner", *Philosophy & Public Affairs* 22, 146-166.
- Sacerdote, B. (2005), "Slavery and the Intergenerational Transmission of Human Capital", *Review of Economics and Statistics* 87(2), 217-234.
- Saez-Marti M. and A. Sjögren (2008), "Peers and Culture", *Scandinavian Journal of Economics* 110(1), 73-92.
- Sauder, U. (2006), "Education Transmission across Three Generations - New Evidence from NCDS Data", Mimeo, University of Warwick.
- Solon, G. (1999), "Intergenerational Mobility in the Labor Market", in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 3, ch. 29, 1761-1800, Amsterdam: Elsevier.
- Warren, J. R. and R. M. Hauser (1997), "Social Stratification across Three Generations: New Evidence from the Wisconsin Longitudinal Survey", *American Sociological Review* 62(4), 561-572.

Appendix: Institutional background

The four generations studied in this paper span a century during which Swedish society was transformed from early industrialization to present day welfare society. While subsidized childcare, generous child allowances, free schooling through high school, generous grants and loans for higher education, social security, unemployment benefits, free health care and pensions constitute today's welfare system, Malmö in the beginning of the 20th century had some, but not all of these institutions in place, when the parents of the initially sampled index generation grew up.

Malmö is located in the southern part of Sweden. It was and is by population size Sweden's third city. At the beginning of the 20th century Malmö grew at a rapid pace and tripled its population from 61,000 to 192,000 between 1900 and 1950, compared to today's 300,000. Much of the population growth was a result of rapid urbanization. Malmö was early on one of the most industrialized cities in Sweden. When the original data collection of the Malmö study was initiated, in 1938, three large employers dominated.¹⁶ After 1960, an increasing fraction was employed within the public sector and by 1980, 20% of the men and 50% of the women held public sector jobs.

In the early 20th century, Swedish compulsory schooling was only six years, but a seventh year was introduced already in 1914 in Malmö. Yet, many children kept leaving school after six years. Seven years of schooling only become the norm around 1920 when a municipal grant was introduced to compensate poor families for the lost earnings during the seventh year of school. This grant existed until 1936 when compulsory schooling was extended to seven years throughout Sweden. In the late 1930's almost a third of all Malmö children continued beyond compulsory schooling. School enrolment was hence higher than in the rest of Sweden. Malmö was also the first large municipality to extend compulsory schooling to nine years in 1962. Arguably, basic educational infrastructure was well developed and accessible already to the index-generation studied here.

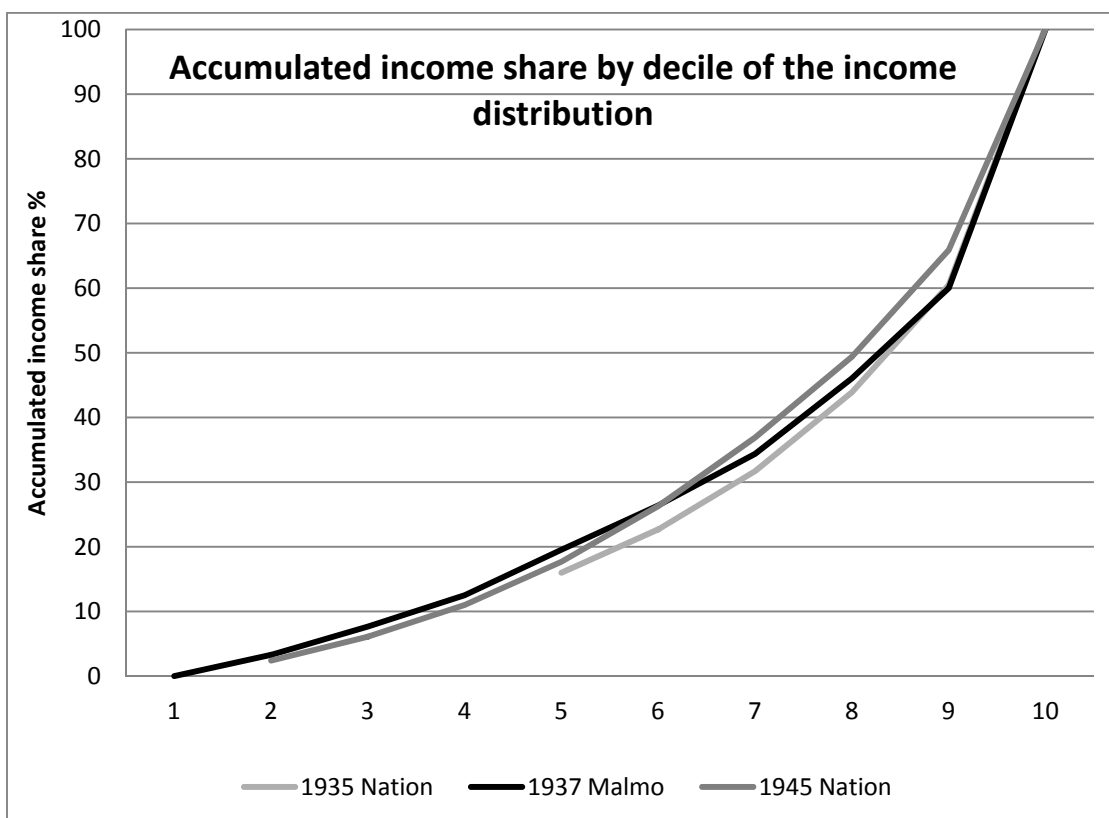
Since the 1920's, loans to help finance higher education were in principle available to the tiny fraction of young people qualified to studying at universities. In the late 1950's student loans were also made available for studies at the high school level. The

¹⁶ *Kockums*, a shipbuilding company and mechanical workshop, with 2,300 employees; *Skånska Cement*, a construction company, with almost 2,000 employees; and *Malmö strumpfabrik*, a stocking factory, with more than 1,000 employees.

present day generous grant and loans program for university students was introduced in 1964. Since then, credit constraints are arguably unlikely to play a role for higher education choices.

Although our sample is not a random sample from the Swedish population, Malmö was (and is) a fairly representative city in Sweden. This can be seen if we compare the earnings distribution for our first generation from Malmö (using our sample) with the earnings distribution for the entire country. To do this we use estimates of the earnings distribution obtained by Bentzel (1952), who used tax registers to construct measures of the Swedish income distribution. Figure 2 compares the earnings distribution of the first generation in our data in 1937 with those obtained by Bentzel for the years 1935 and 1945. It is interesting to note that the income distribution among the Malmö families does not deviate drastically from the national income distribution.

Figure 2: A comparison of earnings distribution for the first generation in the Malmö data for 1937 with those obtained by Bentzel (1952) for Sweden in 1935 and 1945.



Source: Own computation based on Malmö data and Bentzel (1952).