Life-Cycle Bias and the Returns to Schooling in Current and Lifetime Earnings*

by

Manudeep Bhuller
Research Departement, Statistics Norway
msi@ssb.no

Magne Mogstad
Research Department, Statistics Norway,
ESOP at University of Oslo, and IZA
magne.mogstad@gmail.com

Kjell G. Salvanes
Department of Economics, Norwegian School of Economics,
IZA, CES-ifo and CEE
kjell.salvanes@nhh.no

May 15, 2011

Abstract: This paper provides evidence on the returns to schooling in current and lifetime earnings. We use these results to assess the importance of life-cycle bias in earnings regressions using current earnings as proxy for lifetime earnings. To account for the endogeneity of schooling, we apply three commonly used identification strategies. Our estimates demonstrate a strong life-cycle bias, often exceeding the bias from assuming that schooling is exogenous. We also find that the cross-section estimates of the returns to schooling are highly sensitive to the age composition of the sample. They tend to increase with mean age, reflecting that higher educated workers experience more rapid earnings growth through most of the life-cycle. We further show that the returns to schooling in lifetime earnings are relatively low compared to what previous evidence based on cross-section data suggest.

* We thank Paul Bingley, Anders Björklund, Paul J. Devereux, Per-Anders Edin, Monique de Haan, Steven J. Haider, Colm Harmon, Matthew Lindquist, Philip Oreopolous, Oddbjørn Raaum, Gary Solon, Petra Todd, Ian Walker and participants at a number of seminars and conferences for helpful comments. Financial support from the Research Council of Norway (194339) is gratefully acknowledged. The project is part of the research activities at the ESOP center at the Department of Economics, University of Oslo. ESOP is supported by the Research Council of Norway. An earlier version of this paper circulated as NHH Discussion Papers 2011/4, dated February 29, 2011.

1. Introduction

Earnings regressions are at the heart of labor economics, and have been widely used to capture how the labor market rewards productivity attributes like schooling. The earnings regression can be derived from economic theory assuming that individuals choose schooling level to maximize their present value of lifetime earnings, taking as given the post-school earnings profile. Yet, empirical evidence on the returns to schooling usually comes from cross-section studies, regressing (log) current earnings on schooling conditional on age or (potential) experience.[1]

The common practice of using current earnings to proxy for lifetime earnings is due to the simple fact that researchers seldom have access to data on long-run or lifetime earnings. Unfortunately, this empirical simplification does not come without a price. Haider and Solon (2006) demonstrate that the association between current and lifetime earnings varies systematically over the life-cycle.[2] They further show that regression models, using current earnings as a proxy for lifetime earnings, will therefore produce inconsistent estimates (i.e. life-cycle bias) of the regression coefficients. Importantly, this misspecification leads to inconsistent estimates above and beyond the bias due to classical measurement error, and the inconsistency will occur even when the current earnings proxy is used as a dependent variable. Therefore, a critical element in identifying the returns to schooling is to assess the role of life-cycle bias in earnings regressions. That is the focus of this study.

Figure 1 illustrates the large amount of life-cycle bias that may be embedded in returns to schooling estimates based on current earnings. This figure plots the log-earnings age profiles for college and high-school educated Norwegian men born in the years 1948-1950.

[1] See Heckman et al. (2006, 2008) for a critical review of the large empirical literature on earnings regressions based on Mincer's (1974) seminal model of schooling choices.

[2] Haider and Solon (2006) use US data to demonstrate the strong life-cycle association between current and annual earnings. Their empirical analysis is replicated and extended for Sweden (Böhlmark and Lindquist, 2006) and Germany (Brenner, 2010). See also Björklund (1993) for an early study of the correlation between current and lifetime income.

Both earnings profiles display the familiar concave shape documented and analyzed by Mincer (1974), but the college educated workers experience more rapid earnings growth through most of the life-cycle. The horizontal lines depict the log of lifetime earnings, measured as the annuitized value of real earnings from age 20 to 58. The difference in the log of lifetime earnings between college and high-school educated workers is simply the vertical distance between the two horizontal lines. The life-cycle bias in the returns to schooling at a particular age depends on how well the difference in the log of current earnings approximates the difference in the log of lifetime earnings. The figure suggests that the current earnings gap between college and high school educated workers late (early) in their careers tends to overstate (understate) the lifetime earnings gap. Taken at face value and assuming that schooling is exogenous, this would mean that there is an upward (downward) life-cycle bias in the returns to schooling, when earnings are measured late (early) in the working lifespan.

The main objectives of this paper are (a) to estimate the returns to schooling in lifetime and current earnings, and (b) to assess the life-cycle bias in returns to schooling. Previous evidence on life-cycle bias in the returns to schooling comes from studies that have assumed that schooling is exogenous, and constructed synthetic cohort-based earnings profiles from short panels of earnings data spanning only a segment of the life-cycle.[3] We use a unique Norwegian data set with nearly career-long earnings histories for certain cohorts. Our analytic sample is restricted to males. To account for the endogeneity of schooling, we apply three different identification strategies that are currently in use in the literature: i) within-twin-pair estimation, ii) controls for ability test scores, and iii) compulsory schooling reform as instrument for schooling. It should be emphasized that our focus is not on the validity of these identification strategies: Our aims are to estimate the returns to schooling in

[3] See e.g. Heckman et al. (2003, 2006), who examine life-cycle variation in the returns to schooling, as well as several other important aspects of earnings regressions, such as functional form assumptions, the consequences of tuition and taxes, and uncertainty. However, these studies assume that schooling is exogenous, and they rely on synthetic cohort-based earnings profiles.

lifetime earnings, and to assess the importance of life-cycle bias in earnings regressions using current earnings as proxy for lifetime earnings, applying commonly used identification strategies.

Our returns to schooling estimates may be summarized with three important conclusions. First, we find evidence of substantial life-cycle bias in the returns to schooling, often exceeding the bias from assuming that schooling is exogenous. The life-cycle bias is minimized when individuals' earnings are measured in their early 30s, and there is large positive (negative) life-cycle bias with earnings measured after age 40 (before age 30). A possible remedy for cross-section estimates of the returns to schooling is to restrict the sample to individuals around age 32-33. Second, the common practice of using cross-section data when estimating the returns to schooling is shown to be highly sensitive to the age composition of the sample. They tend to increase with mean age, reflecting that higher educated workers experience more rapid earnings growth through most of the life-cycle. This means that it is necessary to pay close attention to differences in age composition when comparing estimates of the returns to schooling across countries, subgroups, or time. Third, the returns to schooling in lifetime earnings are relatively low compared to what previous studies using cross-section data have suggested. This means that we may need to reconsider how much the labor market actually rewards an additional year of schooling.

After assessing the life-cycle bias in cross-section estimates of the returns to schooling, we investigate whether it is likely to be merely an econometric peculiarity or a real cause for concern in empirical research. Using our Norwegian data, we first show that the large increase in the returns to schooling since the 1980s disappears once life-cycle bias is minimized by restricting the cross-section estimates to the sample of individuals aged 32-33. This raises the question of whether the rise in the returns to schooling observed in most developed countries over the last decades is an artifact of changes in life-cycle bias.[4] Next, we perform a meta-analysis of the studies reported in the review articles by Card (1999), Harmon, Oosterbeek and Walker (2003), Oreopolous (2006), and Devereux and Fan (2011). Consistent with a story of life-cycle bias, our analysis shows a strong positive correlation between the mean age in the sample and the estimated returns to schooling. Our meta-analysis also reveals that the sample mean age generally exceeds the age at which life-cycle bias in our estimates is minimized. This raises the concern that previous evidence may have overstated how much the labor market actually rewards schooling.

We conclude the empirical analysis with an examination of the usefulness of errors-in-variables models for analyzing and correcting for life-cycle bias in earnings regressions. Our findings echo the conclusion of Haider and Solon (2006), in that we need to exercise due caution in applying the generalized errors-in-variables model to address life-cycle bias in applied research.[5] On the one hand, the generalized errors-in-variables model predicts well the age at which life-cycle bias in the returns to schooling is minimized. On the other hand, the model appears to be less useful in correcting for life-cycle bias at other ages and in backing out the life-cycle profile in the returns to schooling. The main limitation of the generalized errors-in-variables model seems to be the assumption that the measurement error is uncorrelated with the determinants of earnings, and not that schooling is assumed to be uncorrelated with the error term. That said, the generalized errors-in-variables model is clearly a significant improvement over the textbook model, and highlights well the problems

[4] The rise in the returns to schooling and the associated increase in earnings inequality in almost all developed countries since the early 1980s is one of the most extensively researched topics in economics (see e.g. Lemieux, 2008). While there is substantial agreement about the facts, there is no consensus about the underlying causes. A number of explanations have been proposed and scrutinized, including skill-biased technical change, international trade and globalization, and changes in labor market institutions such as a decline in unionization and an erosion of the minimum wage.

5 The empirical analysis of the generalized errors-in-variables model by Haider and Solon has been replicated and extended for Sweden (Böhlmark and Lindquist, 2006), Germany (Brenner, 2010), and Norway (Nilsen, Vaage, Aakvik and Jacobsen, 2010). See also Stuhler (2010) and Nybom and Stuhler (2011) and for a critical assessment of life-cycle bias in intergenerational mobility estimation, with particular emphasis on the generalized errors-in-variables model.

due to life-cycle bias in a wide range of research that use current earnings variables as proxies for long-run earnings.

This paper unfolds as follows. Section 2 presents a theoretical framework that relates the returns to schooling in current and lifetime earnings and illustrates the possible role of life-cycle bias in earnings regressions. Section 3 describes our data. Section 4 presents the identification strategies and reports summary statistics. Section 5 provides the estimates of the returns to schooling in lifetime and current earnings, before assessing the life-cycle bias. Section 6 examines the usefulness of the generalized errors in-variables model in analyzing and correcting for life-cycle bias in the returns to schooling. Section 7 concludes.

2. Theoretical framework

This section uses a framework of compensating differences, originally proposed by Mincer (1958), to relate the returns to schooling in current and lifetime earnings and illustrate the possible role of life-cycle bias in earnings regressions.

Following Willis and Rosen (1979), suppose that individuals choose between two levels of schooling, labeled college (A) and high school (B), to maximize the present value of lifetime earnings. Assume that credit markets are perfect and the environment is perfectly certain, but occupations differ in the amount of schooling required. If an individual chooses college, his current stream is

$$y^{A}{}_{t} = \begin{cases} 0, & t \leq s' \\ \bar{y}_{A} e^{g_{A}(t-s)}, & t > s' \end{cases}$$

where *s'* is the number of years it takes to get a college degree, *t* represents age (measured as years since high school graduation), $\bar{y}_A$ is initial wage, and g_A is the growth rate in wages. If the individual chooses high school, his earnings stream is

$$y^{B}{}_{t} = \bar{y}_{B} e^{g_{B}(t)}, \quad t \geq 0 .$$

Additional schooling entails opportunity costs in the form of foregone earnings (but no direct cost such as tuition). Assume an infinite horizon, and an exogenously determined interest rate *r*, with $r > g_A > g_B > 0$. Then, the present value of earnings is

$$Y^A = \int_0^\infty e^{-rt} y_t^A dt = \frac{e^{-rs'} \bar{y}_A}{r - g_A}$$

if college is chosen, and

$$Y^B = \int_0^\infty e^{-rt} y_t^B dt = \frac{\bar{y}_B}{r - g_B} ,$$

if high school is chosen. To induce a worker to choose college, foregone earnings while in school must be compensated by higher future earnings, such that $Y^A > Y^B$. In long-run competitive equilibrium, the relationship between lifetime earnings and schooling is such that (i) the supply and demand for workers of each schooling level are equated, and (ii) no worker wishes to alter his schooling level.

In the basic framework of compensating differences, individuals are ex ante identical. In this case, equilibrium requires that individuals are indifferent between schooling levels such that the return to college in lifetime earnings is zero, $log\, Y^A = log\, Y^B$. However, the return to college in current earnings, $\log y_t^A - \log y_t^B$, will generally be non-zero in equilibrium. Moreover, the difference in the returns to college in lifetime and current earnings will vary as a function of the age (or experience level) at which current earnings are observed, $\partial(\log y_t^A - \log y_t^B)/\partial t = g_A - g_B > 0$. Following Haider and Solon (2006), we define $(\log y_t^A - \log y_t^B) - (\log Y^A - \log Y^B)$ as the *life-cycle bias* in using current earnings at age *t* as a proxy for lifetime earnings.

More realistic model of earnings allow for ex-ante heterogeneous individuals, such as in initial wages, growth rates, and the interest rate (see e.g. Willis and Rosen, 1979; Cameron and Taber, 2004, Heckman, Lochner and Todd, 2006). Yet, the crucial insight of the basic framework of compensating differences still applies: To induce a worker to undertake additional schooling, foregone earnings while in school must be compensated by higher future earnings. This may generate changes in earnings variation around the central tendency of earnings growth, causing life-cycle bias in earnings regressions using cross-section data.

To circumvent the issue of life-cycle bias, the data used to estimate the return to schooling would ideally consist of complete longitudinal life histories of earnings. Unfortunately, such ideal data are seldom available. Mincer (1974) therefore suggested two simple approaches to approximate the returns to schooling in lifetime earnings from cross-section data. In the remainder of the paper, we will estimate the returns to schooling in lifetime earnings from nearly career-long earnings histories, and assess how well the two approaches address the issue of life-cycle bias in earnings regressions based on cross-section data.

The first, much used approach assumes separability between schooling and age (or experience), in which case controlling for age (or experience) addresses the issue of life-cycle bias. In our case, the separability assumption would imply that $g_A = g_B$. Unfortunately, data do not support the separability assumption; moreover, it is at odds with more realistic models of earnings (see e.g. Heckman, Lochner and Todd, 2006).

The second approach relies on the so-called overtaking age (or experience level), at which $\log y_t^A - \log y_t^B$ equates $\log Y^A - \log Y^B$. In our case, the overtaking age is unique because the age-earnings profiles of the two schooling levels will not cross more than once, $t^* = \dfrac{s'(g_A - r) + \log(r - g_B) - \log(r - g_A)}{(g_A - g_B)}$. Knowledge of the overtaking age provides an empirically useful short cut method for estimating the returns to schooling in lifetime earnings directly from $\log y_{t*}^A - \log y_{t*}^B$. In particular, the generalized errors-in-variables model proposed by Haider and Solon (2006) can be used to identify the overtaking age under transparent assumptions.[6]

3. Data

Our empirical analysis utilizes several registry databases maintained by Statistics Norway. This allows us to construct a rich longitudinal data set containing records for every Norwegian from 1967 to 2008. The variables captured in this data set include individual demographic information (gender, birth year) and socio-economic data (annual earnings, years of schooling). Importantly, the data set includes personal identifiers, allowing us to link children to their parents and siblings. We can therefore merge the longitudinal data set with Census data from 1960 and 1970. This allows us to add family background variables, including family income (in quartiles), parental education, and childhood municipality of residence. Family income is obtained by summing the father's and the mother's incomes. The father's and the mother's educational attainment is represented by a dummy variable indicating whether or not they had college education. Detailed descriptions of all the variables used in the empirical analysis are given in Table A.1 in the Appendix.

Our measure of earnings is the sum of pretax market income (from wages and self-employment) and work-related cash transfers, such as unemployment benefits, sick benefits, and parental leave benefits. We define *current earnings* as the annual real earnings in a given year, adjusted for inflation and real wage growth. Following Haider and Solon (2006), our measure of *lifetime earnings* is the annuity value of the discounted sum of annual real earnings. To calculate the annuity value we use an interest rate of 2.3 percent, which

[6] A third approach is to assume a stationary economy, with zero aggregate productivity change and constant population growth rate, in which case cross-sectional earnings-age profiles can be used to construct synthetic cohort-based earnings-age profiles. However, recent analyses reveal that earnings patterns have changed dramatically across cohorts: As a result, constructing synthetic cohort-based earnings profiles from cross-section data will generate bias in earnings regressions (see e.g. Heckman, Lochner and Todd, 2006).

correspond to the average real interest rate on deposits and loans in Norway over the period 1967-2006 (Aaberge, Mogstad and Peragine, 2011).

The Norwegian earnings data has several advantages over those available in many other countries. First, there is no attrition from the original sample due to the need to ask permission from individuals to access their tax records. In Norway, these records are in the public domain. Second, our earnings data pertain to all individuals, and not only to jobs covered by Social Security. Third, we have nearly career-long earnings histories for certain cohorts, and do not need to extrapolate the earnings profiles to ages not observed in the data. And fourth, top-coding is only performed at very high earnings levels. In fact, less than 3 percent of the observations have right-censored earnings in any given year. Yet to make sure that top-coding is not driving our results, we have also estimated the returns to schooling using a Pareto distribution to simulate earnings above the top-coding threshold. Appendix B describes the results from this robustness check.

Our regressor of interest is the number of years of schooling. To ensure that virtually everyone has completed their education, we will throughout this paper measure schooling at age 40. Educational attainment is reported by the educational establishment directly to Statistics Norway, thereby minimizing any measurement error due to misreporting.

Our main results focus on the 1948-1950 cohorts, in order to ensure complete records on earnings from age 20 to 58. Our analytic sample is restricted to males, to minimize selection issues due to the low labor market participation rates for women in the early periods. We exclude immigrants as well as individuals with missing information on years of schooling, place of residence, or family background variables. Our key dependent variables are the log of the annuitized value of earnings from age 20 to 58, as well as the log of current earnings at every age 28-58. In order to ensure that our sample is the same for all dependent variables, we exclude individuals with zero earnings in one or more years between age 28 and

58. Applying these restrictions provided us with what we will refer to as the *full sample*, consisting of 56,832 individuals.

4. Identification strategies

In the absence of experimental evidence, it is difficult to know whether the higher earnings observed among high educated workers are caused by their additional schooling, or whether individuals with greater earning capacity have chosen to acquire more schooling. To address this concern for selection bias in earnings regressions, a number of identification strategies have been proposed and scrutinized. In this paper, we apply three different identification strategies that are currently in use in the literature.

Our earnings regressions are summarized by the following two equations:

(1) $$y_i = \alpha + \rho s_i + \chi' F_i + \delta^c + \delta^m + \varepsilon_i$$

(2) $$y_{it} = \alpha_t + \rho_t s_i + \chi'_t F_i + \mu^c + \mu^m + \varepsilon_{it}$$

In equations (1) and (2), s is the number of years schooling, F is a vector of control variables for family background, comprising family income and parental education. The only difference between the two earnings regressions is the specification of the dependent variable: equation (1) uses lifetime earnings, y, whereas equation (2) uses current earnings at age t, y_t. Both equations include a full set of indicators for childhood municipality of residence, δ^m and μ^m, and a full set of birth cohort indicators, δ^c and μ^c. The standard errors are always clustered at the municipality level and robust to heteroskedasticity.

Within-twin-pair estimation. Our first identification strategy is to use within-twin-pair estimation (see e.g. Griliches, 1979, Ashenfelter and Krueger, 1994). This strategy identifies the returns to schooling by comparing the difference in schooling of the twins in a pair with the difference in their earnings. The idea is that twins share genetics and the same family background environment, possibly reducing the extent of ability bias.[7]

Our *twin sample* consists of 702 individuals, amounting to around 1.3 percent of the full sample. Unfortunately, our data does not allow us to distinguish between monozygotic and dizygotic twins. This means that our within-twin-pair estimates might be confounded by unobserved heterogeneity in genetics. Since we only consider male twin pairs, we know from Weinberg's rule that about half of the twin sample is monozygotic.

Controls for ability. In the second identification strategy, we attempt to control for differences in ability (see e.g. Griliches, 1977), through information on IQ test scores from the Norwegian military records. In Norway, military service is compulsory for all able males. Before entering the service, their medical and psychological suitability is assessed: This occurs for the great majority around their eighteenth birthday. However, the IQ test scores are only available for cohorts born in 1950 or later. Our *IQ sample* therefore consists of 14,936 individuals who were born in 1950 and had non-missing IQ test scores.

The IQ measure is a composite score from three speeded IQ tests – arithmetic, word similarities, and figures.[8] The composite IQ test score is an unweighted mean of the three subtests. The IQ score is reported in stanine (Standard Nine) units, a method of standardizing

[7] Although much used, within-twin-pair estimation has been criticized. First, there could be other differences between the twins that are unobservable to the researcher and that affect both the schooling decision and earnings. Second, within-pair estimates will suffer from greater attenuation bias if measurement error is greater for schooling measured in differences than levels. However, we reduce the problem of measurement error in schooling, by measuring completed education at age 40 and by using administrative data rather then self-reported surveys. See e.g. Bound and Solon (1999) and Isacsson (2004) for a discussion of attenuation bias in within-twin-pair estimation.

[8] The arithmetic test is quite similar to the arithmetic test in the Wechsler Adult Intelligence Scale (WAIS) (Sundet et al. 2005; Cronbach 1964). The word test is similar to the vocabulary test in WAIS, and the figures test is similar to the Raven Progressive Matrix test (Cronbach 1964). See Sundet et al. (2004, 2005) and Thrane (1977) for details.

raw scores into a nine point standard scale with a normal distribution, a mean of 5, and a standard deviation of 2. We add a full set of test score indicators to the earnings regressions.

Instrumental variables strategy. Our third identification strategy follows Black, Devereux and Salvanes (2005) and Aakvik, Salvanes and Vaage (2010) in using the staged implementation of a Norwegian compulsory schooling law reform as a source to exogenous variation in schooling. The reform increased compulsory schooling from seven to nine years, and was implemented over a 12-year period from 1960 to 1971 in different municipalities (the lowest level of local administration) at different times. Thus, for more than a decade, Norwegian schools were divided into two separate systems, where the length of time of compulsory schooling depended on the year you were born and the municipality in which you lived.

We are able to successfully identify the year in which the reform was implemented for as many as 671 out of the 728 municipalities. In line with Black *et al*. (2005) and Aakvik *et al*. (2010), we drop individuals who were residing in a municipality to which we could not assign a reform indicator. Applying this sample restriction we get an *IV sample* consisting of 53,915 individuals, which is nearly 95 percent of the full sample.

Our instrumental variables (IV) strategy is summarized by the second stages expressed in equations (1) and (2), and the first stage:

$$(3) \qquad s_i = \gamma_0 + \gamma_1 R_i + \gamma_2 F_i + \gamma'_3 R_i F_i + \gamma'_4 R_i A_i + \pi_i^{\,c} + \pi_i^{\,m} + \eta_i$$

where R is the compulsory schooling reform dummy, being equal to 1 if the individual was exposed to the reformed schooling law and 0 otherwise. Following the baseline specification in Aakvik *et al.* (2010), we add interaction terms between the reform dummy and family background variables, RF, and between the reform dummy and variables indicating availability of different school types RA. The vector A includes indicator variables for the availability of upper secondary school, vocational college, regional college and university in the municipality that the individual grew up in. By adding the interaction terms, we allow the response to the compulsory schooling reform to vary with family background and availability of different school types. Since the availability of different schools at the municipality level is unchanged over this time period, the full set of municipality indicators, π^m, capture the direct effects of school availability on years of schooling. The full set of birth cohort indicators, π^c, allows for a (possibly nonlinear) secular trend in educational attainment.[9]

We refer to Black *et al.* (2005) and Aakvik *et al.* (2010) for detailed discussions of instrument validity and of relevant institutional details. For example, they show that there is no relationship between the timing of implementation of the schooling reform and municipality characteristics such as average earnings, education levels, average age, urban/rural status, industry or labor force composition, municipality unemployment rates, or the share of individuals who were members of the Labor Party (the most pro-reform of the dominant political parties).

Summary statistics. Table 1 reports summary statistics for each sample. There are common patterns in the summary statistics across the samples. First, average current earnings display the familiar concave shape over the life-cycle, increasing from age 28 to 48, and declining slightly afterwards. Second, average current earnings are most similar to average lifetime earnings when individuals are in their mid 30s. Third, the increase in average current earnings over the life-cycle is accompanied by an increase in the variance of current earnings. This is an important observation, since life-cycle bias is due to changes in earnings variation around

[9] Black *et al*. (2005) and Aakvik *et al.* (2010) also test for a municipality specific linear trend and find that this does not impact the results. The same holds true for our analysis.

the central tendency of earnings growth. The main difference across the samples is that twins have somewhat lower earnings and educational attainment, in line with the findings of previous studies (see *e.g.* Bound and Solon, 1999).

5. Returns to schooling estimates

We begin by reporting estimates of the returns to schooling in lifetime and current earnings from a cohort-based analysis, following individuals over their working life span. This allows us to assess the life-cycle profile in the returns to schooling, and identify the ages at which life-cycle bias is minimized. Next, we follow standard practice in the literature on earnings regressions and use cross-section data to estimate the returns to schooling. By comparing these results to those produced by the cohort-based analysis, we learn how well the cross-section estimates of the returns to schooling approximate the returns to lifetime earnings. Furthermore, by comparing the cross-section estimates from different years, we can examine the sensitivity of returns to schooling estimates to changes in the age composition of the sample. We conclude this section with a discussion of our finding, assessing whether life-cycle bias is likely to be merely an econometric peculiarity or a real cause for concern in empirical research.

5.1 Cohort-based analysis

Main analytical sample. As described above, our main analytical sample consists of the 1948-1950 cohorts, for which we have complete records on earnings from age 20 to 58. Table 2 shows the estimated returns to schooling in lifetime earnings and current earnings at different ages for these cohorts. The table reports results for the full sample (column 1), the IQ sample (columns 2-3), the twin sample (columns 4-5), and the IV sample (columns 5-6).[10] Each cell represents a separate regression. Figure 2 plots the estimated returns to schooling in lifetime earnings and current earnings, age 28-58.

There are clear patterns in our results, independent of identification strategy. We see that the returns to schooling increase over most of the life-cycle. The estimates start out negative when these men are young, reflecting that some individuals taking higher education are still in school, and that the low educated workers have considerably more work experience early in their careers. The returns to schooling rise quickly until individuals are in their late 30s, after which they increase modestly. The association between the returns to schooling in lifetime and current earnings is strongest when individuals are 32-33 years old, and there is positive (negative) life-cycle bias with earnings measured after age 40 (before age 30).

There are, however, some noticeable differences in the results across the identification strategies. These differences are unlikely to be due to the discrepancies in sample selection, as the OLS estimates are quite similar across the samples. Instead, they likely reflect population heterogeneity in the returns to schooling or omitted variables bias. The IV strategy produces the highest returns to schooling in lifetime earnings and the most pronounced life-cycle bias. A common interpretation of the relatively high IV estimates of the returns to schooling is that the effect of another year of schooling varies across individuals, and that the instruments used change the educational choice of a subgroup with relatively high returns.[11] An often-cited example is studies that measure the return to schooling among persons obliged to stay in school longer because of compulsory school laws. The argument is that compulsory schooling laws mostly affect the education decision of persons with poor family background, and that

[10] The first stage estimates are provided in Table A.2 in Appendix A. We can see that the first stages are strong with F-statistics on the excluded instruments exceeding 43, which means that we do not need to worry about problems due to weak instruments.

[11] An alternative explanation is that measurement error in schooling leads to a downward bias in the OLS estimates of the returns to schooling. Card (2001) concludes, however, that it is unlikely that so many studies would find large positive gaps between their IV and OLS estimates simply because of measurement error.

the return to more schooling in this subset of the population is relatively high. Along the same lines, an interpretation of our results is that the subgroup induced to take more schooling because of the compulsory schooling law reform not only achieve an increase in earnings levels, but also a more rapid earnings growth over the life-cycle.

Turning attention to the within-twin-pairs estimation and the strategy of controlling for test scores, we find that the OLS estimates generally exceed the within-twin returns to schooling estimates. A common interpretation of this finding is that endogeneity of schooling leads to upward bias in OLS estimates of the returns to schooling. However, our results suggest that this ability bias is fairly stable over the life-cycle.

*Extended analytical sample***.** Our estimates from the 1948-1950 cohorts may not necessarily extend to other cohorts, because of changes in skill prices or cohort quality. We therefore examine the external validity of our results by changing the cohorts included in the analytical sample. Specifically, we look separately at cohorts born 1951-1953 and 1954-1956, using the complete records of earnings from age 20 to 55 and age 20-52, respectively. The results are presented in the Appendix, in panel A of Table A.3 and in Figure A.1. We find no significant difference across cohorts in the estimated returns to schooling over the life-cycle: Our cross-section analysis discussed in the next subsection will therefore use cohorts born in the period 1948-1956.

For cohorts born after 1950, our data does not allow us to calculate the annuitized value of earnings from age 20 to 58. Instead, we construct an alternative measure of lifetime earnings, defined as the annuitized value of earnings from age 20 to 52. The results for the 1951-1956 cohorts are presented in panel B of Table A.3, whereas the results from the 1948-1950 cohorts are reported in panel B of Table 2. We find no significant difference across the cohorts in the estimated returns to schooling in this measure of lifetime earnings.

In an attempt to construct measures of lifetime earnings from age 20 to 58 for cohorts born 1950-1956, we impute earnings for cohorts born after 1950. Specifically, we use a nearest neighbor matching algorithm to impute the missing earnings history above age 55 for the 1951-1953 cohorts, and above age 52 for the 1954-1956 cohorts. The matching algorithm is described in detail in the Appendix, but to fix ideas consider an individual born in 1953. Conditional on the individual's level of schooling, family background characteristics, childhood county of residence, and a dummy variable for exposure to compulsory schooling reform, the matching algorithm identifies the best individual match from the 1948-1950 cohorts. The best individual match is defined as the one minimizing the Mahalanobis distance in annual real earnings from age 20 to age 55, between the individual and the potential matches. The missing earnings observations after age 55 are then imputed from the earnings record of the best individual match.[12] The results based on this alternative measure of lifetime earnings are reported in panel B of Table A.3. The key finding is that there is no significant difference in the estimated returns to schooling in lifetime earnings across the cohorts.

5.2 Cross-section analysis

Table 3 reports cross-section estimates of the returns to schooling for the years 1985, 1995, and 2005 (panel A), and estimates of the returns to schooling for two different measures of lifetime earnings (panel B). Each cell represents a separate regression. Both panels use the sample of males born during the period 1948-1956. The first lifetime earnings measure is based on complete records of earnings for all cohorts from age 20 to 52, whereas the second measure of lifetime earnings is also based on imputed earnings for some cohorts at ages 53-58. For each cross-section, the table reports the mean age of the sample.

[12] To test the matching method, we have performed an out-of-sample check for ages where we have complete earnings records for all cohorts. These out-of-sample results suggest that the matching method performs very well in predicting individuals' current earnings.

There are clear patterns in our results. The estimates of the returns to schooling from the 1985 cross-section are quite similar to the estimates using lifetime earnings as the dependent variable. The reason is that the individuals are in their early 30s, when the returns to schooling in current earnings are most similar to the returns to schooling in lifetime earnings. The cross-section estimates of the returns to schooling are much higher in 1995, mirroring that the returns to schooling in current earnings rise quickly until individuals are in their late 30s. From 1995 to 2005, we see a smaller increase in the cross-section estimates, consistent with the modest increase in the returns to schooling in current earnings after individuals turn 40. In fact, there is no increase in the IV estimates from 1995 to 2005, attributable to the fact that the IV estimates in the returns to schooling in current earnings change little after age 38.

5.3 Discussion

An important insight from our analysis is that cross-section estimates of the returns to schooling are highly sensitive to the age composition of the sample. In particular, they tend to increase with mean age, reflecting that high educated workers experience more rapid earnings growth through most of the life-cycle. This means that we need to pay close attention to differences in age composition when comparing cross-section estimates of the returns to schooling across countries, subgroups, or time. Below, we illustrate by two examples the possible implications of life-cycle bias for the conclusions drawn about the returns to schooling.

First, we use our Norwegian data to examine how changes in the age composition of the sample may affect the evolution of the returns to schooling from 1967 to 2008. In each year, we estimate the returns to schooling for males aged 16-64 and for the subsample of males aged 32-33. Figure 3 displays the result. We can see that the returns to schooling for males aged 16-64 increased over the 1980s and into the late 1990s. However, the large increase in the returns to schooling disappears once we minimize life-cycle bias by restricting the cross-section estimates to the sample of individuals aged 32-33. Although we cannot rule out that the differential time trends reflect differences in cohort quality, it raises the question of whether the increase in the returns to schooling for males aged 16-64 is an artefact of changes in life-cycle bias. In particular, since the 1980s the large baby boom cohorts have made their way along the earnings-age profile: We would therefore expect an increase in (upward) life-cycle bias in cross-section estimates of the returns to schooling.[13]

Next, we perform a meta-analysis of the cross-section studies reported in the review articles by Card (1999), Harmon et al. (2003), Oreopolous (2006), and Devereux and Fan (2011). We restrict the analysis to the studies from the Anglon-Saxon countries, which includes information about the mean age in the sample. Figure 4 plots the estimated returns to schooling and the sample mean age for the eleven cross-sections. Consistent with a story of life-cycle bias, the figure shows a strong positive association between the mean age in the sample and the estimated returns to schooling, with a correlation of .71. In fact, the positive association between mean age and the returns to schooling holds up even if we limit the comparison to cross-section estimates taken from the same year (1980 or 1993) in the US. We also see that the sample mean age generally exceed the age at which life-cycle bias in our estimates is minimized. This raises the concern that previous evidence may have overstated how much the labor market actually rewards an addition year of schooling.

6. Errors-in-variables models

[13] As in almost all developed countries, Norway experienced a large increase in the population growth rate following World War II, more familiarly called the baby boom. The baby boomers usually include children born from 1946 to about 1960. For example, The US Census Bureau considers a baby boomer to be someone born during the demographic birth boom between 1946 and 1964. Source: http://www.census.gov/population/www/socdemo/age/general-age.html.

This section examines the usefulness of errors-in-variables models for analyzing and correcting for life-cycle bias in earnings regressions. In our context, the textbook and generalized errors-in-variables model can be summarized by the following equations:

(4) $$y_{it} = \lambda_t y_i + v_{it}$$

(5) $$y_i = \rho s_i + \varepsilon_i$$

(6) $$y_{it} = \rho_t s_i + \varepsilon_{it}$$

where the error term ε is assumed to be uncorrelated with schooling *s*, and the measurement error v_t is assumed to be uncorrelated with each separate determinant of *y* (*s* and ε).[14] Under these assumptions, the widespread use of y_t as a proxy for *y* in equation (5) gives a probability limit of the slope coefficient equal to $\lambda_t \rho$. In the textbook case where $\lambda_t = 1$, ρ will be consistently estimated by OLS. Haider and Solon's (2006) generalized model relaxes this assumption, implying that ρ is biased by a factor of λ_t, and the inconsistency varies as a function of the age at which current earnings are observed.

The generalized-errors-in-variables model implies that knowing ρ_t and λ_t at any age *t* is sufficient to infer the returns to schooling in lifetime earnings, ρ. And *vice versa*, to infer the returns to schooling in current earnings at any age *t*, ρ_t, it is sufficient to know ρ and λ_t. Hence, if the generalized errors-in-variables assumptions hold, the model can be used to back out the life-cycle profile in the returns to schooling, and to correct for life-cycle bias in cross-section estimates of the returns to schooling.

[14] Throughout this section, we follow Haider and Solon in suppressing control variables as well as the intercepts by expressing all variables as deviations from their population means.

We begin by estimating equation (4). Figure 5 presents the estimates of λ_t for the full sample, the twin sample, the IQ sample, and the IV sample. We can see that the associations between current and lifetime earnings are generally different from one, and vary systematically over the life-cycle. Thus, our results confirm the findings of Haider and Solon in suggesting that the textbook errors-in-variables model provides an incorrect characterization of the association between current and lifetime earnings.[15]

Next, we use our estimates of ρ_t and λ_t at every *t* to construct a set of age-specific predictions for the returns to schooling in lifetime earnings, and compare them to the estimated returns to schooling in lifetime earnings. The difference between the predicted and the estimated returns to schooling in lifetime earnings tells us how well the generalized errors-in-variables model corrects for life-cycle bias at a given year *t*. Figure 6 displays the results for each sample. We see that the predicted returns to schooling in lifetime earnings are negative when these men are younger than 30. They rise quickly, crossing the estimated returns to schooling in lifetime earnings in the early 30s, after which they diverge. There is generally large positive (negative) bias in the predicted returns to schooling in lifetime earnings after age 35 (before age 30).

Finally, we use our estimates of ρ and λ_t to construct a set of predictions for the life-cycle profile in the returns to schooling. Figure 7 displays the results for each sample. We see that the predicted returns are not able to reveal the life-cycle profile in the estimated returns to schooling. The predicted returns start out positive and substantial when the estimated returns are negative. They coincide when individuals are in their early 30s, after which the predicted returns increase much less than the estimated returns.

[15] Haider and Solon's estimates of λ_t start out at .24 at age 19, increases steadily until it rises to 1 at age 32, and then declines somewhat in the later forties. Our estimates are quite similar until individuals are in their mid 30s, but we do not find evidence of any decline after age 40. In addition, our estimates of λ_t are much more precisely estimated, reflecting our relatively large sample size.

Our findings echo the conclusion of Haider and Solon, in that we need to exercise due caution in applying the generalized errors-in-variables model to address life-cycle bias in applied research. On the one hand, the generalized errors-in-variables model predicts well the age at which life-cycle bias in the returns to schooling is minimized. On the other hand, the model is not able to predict the life-cycle profile in the returns to schooling. Moreover, the model predictions of the returns to schooling in lifetime earnings are severely biased upward (downward), when current earnings are measured after age 35 (before age 30). The prediction errors occur independently of identification strategy: This suggests that the main limitation of the generalized errors-in-variables model in the returns to schooling application is the assumption that the measurement error is uncorrelated with the determinants of earnings, and not that schooling is assumed to be uncorrelated with the error term. That said, the generalized errors-in-variables model is clearly a significant improvement over the textbook model, and highlights well the problems due to life-cycle bias in a wide range of research that use current earnings variables as proxies for long-run earnings.

7. Conclusion

Research on the economic returns to schooling has a long history in economics. In particular, considerable effort has been directed towards examining the implicit assumption of the Mincer (1957, 1974) model that schooling is exogenous, and a number of identification strategies have been proposed and scrutinized. In contrast, much less attention has been devoted to the life-cycle bias that may arise from the widespread use of current earnings as a proxy for lifetime earnings.

This paper provides evidence on the returns to schooling in current and lifetime earnings. We use these results to assess the importance of life-cycle bias in earnings regressions using current earnings as proxy for lifetime earnings. To account for the endogeneity of schooling, we apply three different identification strategies that are currently in use in the literature: i) within-twin-pair estimation, ii) controls for ability test scores, and iii) compulsory schooling reform as instrument for schooling.

We find evidence of substantial life-cycle bias in the returns to schooling, often exceeding the bias from assuming that schooling is exogenous. The life-cycle bias is minimized when current earnings are measured in their early 30s, and there is large positive (negative) life-cycle bias with current earnings measured after age 40 (before age 30). A possible remedy for cross-section estimates of the returns to schooling is to restrict the sample to individuals aged 30-35. Another important finding is that the cross-section estimates of the returns to schooling are highly sensitive to the age composition of the sample. They tend to increase with mean age, reflecting that higher educated workers experience more rapid earnings growth through most of the life-cycle. This means that it is necessary to pay close attention to differences in age composition when comparing estimates of the returns to schooling across countries, subgroups, or time. Our study also shows that the returns to schooling in lifetime earnings are relatively low compared to what cross-section estimates typically suggest. This means that that we may need to reconsider how much the labor market actually rewards an additional year of schooling.

However, caution is in order. Since we use observational data, we cannot rule out that our estimates suffer from omitted variables bias. Nevertheless, it is reassuring that the main patterns in our results hold true across identification strategies. Another caveat is that the life-cycle bias and the returns to schooling in current and lifetime earnings for the Norwegian cohorts born in the late 1940 and the early 1950 may differ from those for other cohorts or other countries. In particular, we advise readers to exercise due caution in importing our estimates of life-cycle bias to other earnings data. The general lesson to be drawn from our paper is rather that more attention needs to be devoted to life-cycle bias, if we want to use earnings regressions to capture how the labor market rewards productivity attributes like schooling.

References:

Aaberge, R., M. Mogstad and V. Peragine (2011): Measuring long-term inequality of opportunity, *Journal of Public Economics*, 95, 193-204.

Aakvik, A., K. G. Salvanes and K. Vaage (2010): Measuring heterogeneity in the returns to education using an education reform, *European Economic Review*, 54, 483-500.

Ashenfelter, O. and A. Krueger (1994): Estimates of the Economic Return to Schooling From a New Sample of Twins. *American Economic Review*, 84, 1157-1173.

Atkinson, A. (2005): Top incomes in the UK over the 20th century, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 168, 325-343.

Björklund, A. (1993): A Comparison between Actual Distributions of Annual and Lifetime Income: Sweden 1951-1989, *Review of Income and Wealth*, 39, 377-386.

Black, S. E., P. J. Devereux and K. G. Salvanes (2005): Why the Apple Doesn’t Fall Far: Understanding Intergenerational Transmission of Human Capital, *American Economic Review*, 95, 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, 876-896.

Bound, J. and G. Solon (1999): Double Trouble: On the Value of Twins-Based Estimation of the Returns to Schooling, *Economics of Education Review*, 18, 169-182.

Brenner, J. (2010): Life-Cycle Variations in the Association between Current and Lifetime earnings: Evidence for German Natives and Guest Workers, *Labour Economics*, 17, 392-406.

Cameron, S. and C. Taber (2004): Estimation of Educational Borrowing Constrains Using Returns to Schooling, *Journal of Political Economy*, 112, 132-182.

Card, D. (1999): The causal effect of education on earnings, *Handbook of Labor Economics*, vol. 3, part 1, 1801-1863.

Card, D. (2001): Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems, *Econometrica*, 69, 1127-1160.

Cronbach, L. J. (1964): *Essentials of Psychological Testing, 2nd Edition*, London, UK: Harper and Row

Devereux, P. J. and W. Fan (2011): Earnings Returns to the British Education Expansion, mimeo.

Griliches, Z. (1977): Estimating the Returns to Schooling: Some Econometric Problems, *Econometrica*, 45, 1-22.

Griliches, Z. (1979): Sibling Models and Data in Economics: beginning of a Survey, *Journal of Political Economy*, 87, 2, 37-64.

Haider, S. and G. Solon (2006): Life-Cycle Variation in the Association between Current and Lifetime Earnings, *American Economic Review*, 96, 1308-1320.

Harmon, C., H. Oosterbeek and I. Walker (2003): The Returns to Education: Microeconomics, *Journal of Economic Surveys*, 17, 115-155.

Heckman, J. J., L. J. Lochner and P. E. Todd (2006): Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond, *Handbook of the Economics of Education*, 1, 307-458.

Heckman, J. J., L. J. Lochner and P. E. Todd (2008): Earnings Functions and Rates of Return, *Journal of Human Capital*, 2, 1-31.

Isacsson, G. (2004): Estimating the Economic Return to Educational Levels Using Data on Twins, *Journal of Applied Econometrics*, 19, 99-119.

Lemieux, T. (2008): The changing nature of wage inequality, *Journal of Population Economics*, 21, 21-48.

Lie, S.S. (1973): *The Norwegian comprehensive school reform. Strategies for implementation and complying with regulated social change. A diffusion study. Part 1 and II*. The American University

Lie, S.S. (1974): Regulated social change: a diffusion study of the Norwegian comprehensive school reform. *Acta Sociologica*, 16, 332–350.

Mincer, J. (1957): A Study of Personal Income Distribution. Unpublished Ph.D. dissertation, Columbia University.

Mincer, J. (1958): Investment in human capital and personal income distribution, *Journal of Political Economy*, 66 (4), 281-302.

Mincer J. (1974): *Schooling, Earnings and Experience*. Columbia University Press: New York.

Nilsen, Ø. A, K. Vaage, A. Aakvik, K. Jacobsen (2010): Intergenerational Earnings Mobility Revisited: Estimates Based on Lifetime Earnings, forthcoming *Scandinavian Journal of Economics.*

Nybom, M. and J. Stuhler (2011): Heterogeneous Income Profiles and Life-Cycle Bias in Intergenerational Mobility Estimation, IZA Discussion Paper, 5697.

Oreopolous, P. (2006): Estimating average and local average treatment effects of education when compulsory schooling laws really matter, *American Economic Review*, 96, 152-175.

Rosenbaum, P. R. and D. B. Rubin (1985): Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score, *The American Statistician*, 39, 33-38.

Stuhler, J. (2010): Empirical Strategies to Eliminate Life-Cycle Bias in the Intergenerational Elasticity of Earnings Literature, IZA Discussion Papers 5245.

Sundet, M. J., D. G. Barlaug, and T. M. Torjussen (2004): The End of the Flynn Effect? A Study of Secular Trends in Mean Intelligence Test Scores of Norwegian Conscripts During Half a Century, *Intelligence*, XXXII, 349-362.

Sundet, J. M., K. Tambs, J. R. Harris, P. Magnus, and T. M. Torjussen (2005): Resolving the Genetic and Environmental Sources of the Correlation Between Height and Intelligence: A Study of Nearly 2600 Norwegian Male Twin Pairs," *Twin Research and Human Genetics*, VII, 1-5.

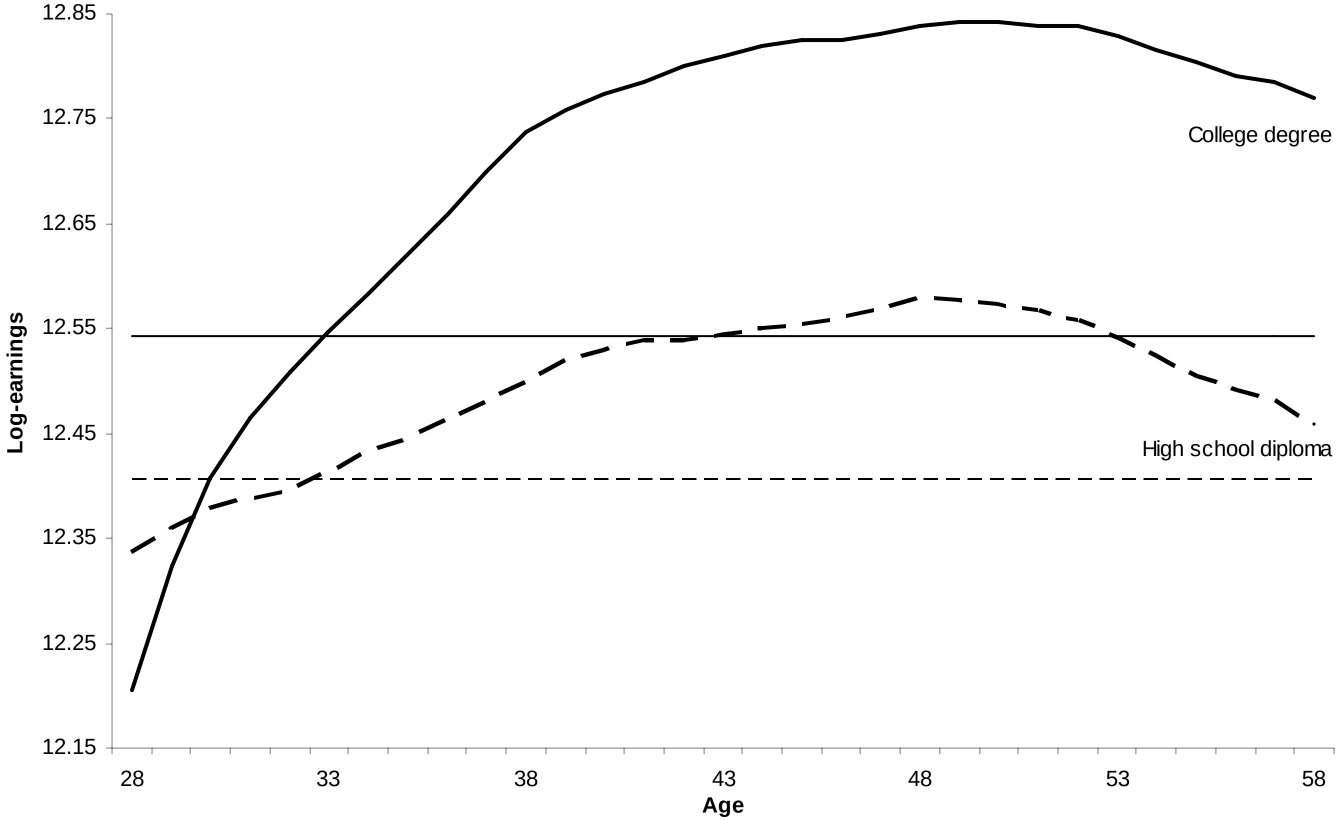
Telhaug, A.O. (1969): *Den 9 årige skolen og differensieringsproblemet. En oversikt over den historiske utvikling og den aktuelle debatt*. Lærerstudentenes Forlag, Oslo (in Norwegian)

Thrane, V. C. (1977): *Evneprøving av Utskrivingspliktige i Norge 1950-53,* Arbeidsrapport nr. 26, INAS (in Norwegian)

Willis, R. and S. Rosen (1979): Education and Self-Selection, *Journal of Political Economy*, 87, 7-36.

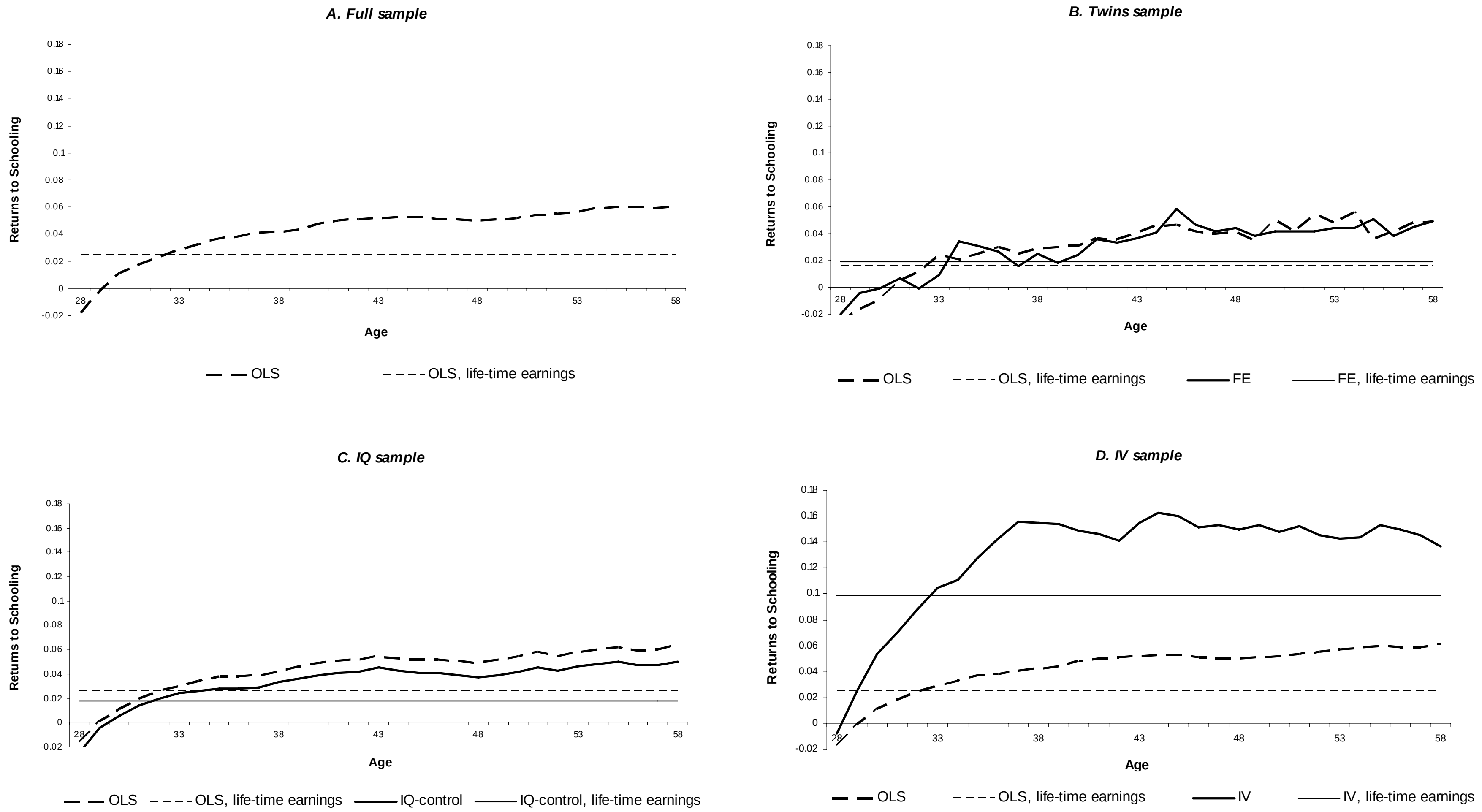
Figures

Figure 1 Log-earnings – age profiles



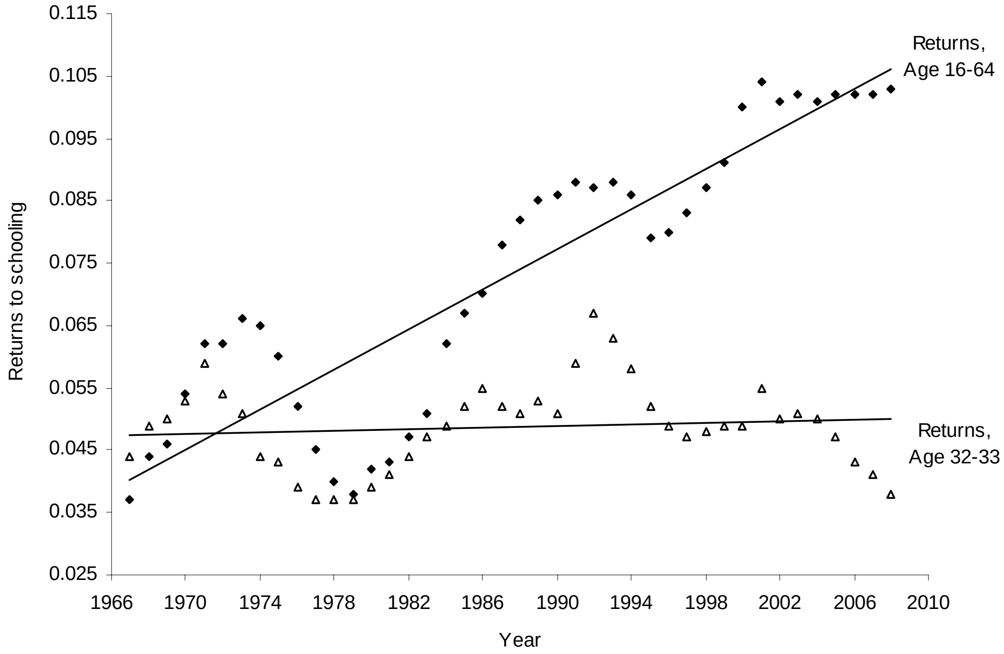
Note: The figure plots log current earnings and log lifetime earnings for males born 1948-1950, with either high school diploma or college/university degree as highest completed education at age 40. See Section 3 for details about sample selection and definition of earnings variables.

Figure 2 Returns to schooling in current and lifetime earnings



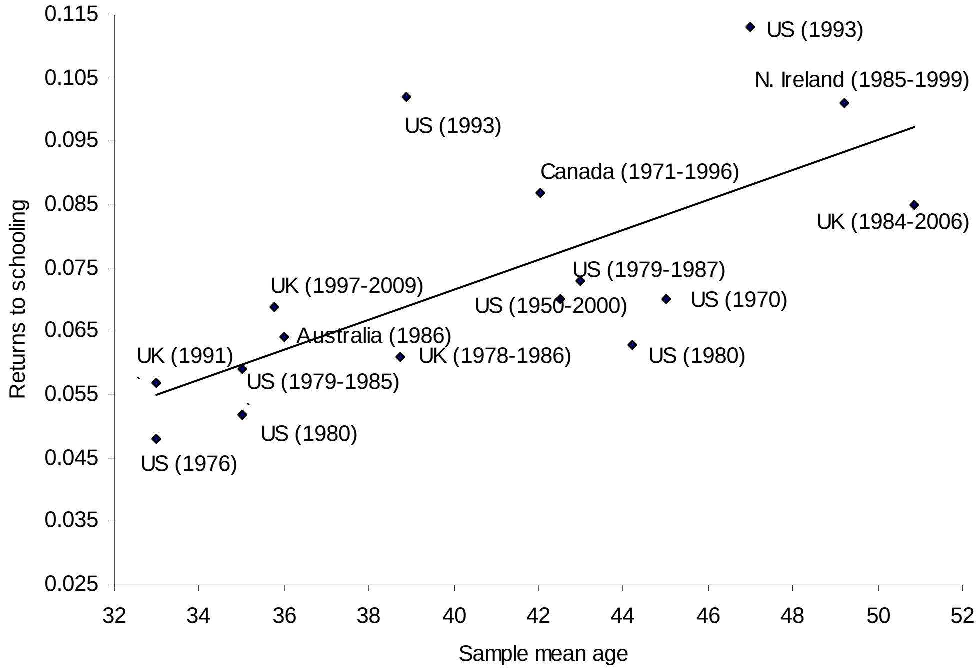
Note: This figure plots estimates of the returns to schooling in current and lifetime earning by identification strategy. Cohort dummies, municipality fixed effects and family background variables are included in all regressions. First stage IV estimation results are given in Table A.2. See notes in Table 1 for sample details.

Figure 3 Cross-sectional returns to schooling in Norway



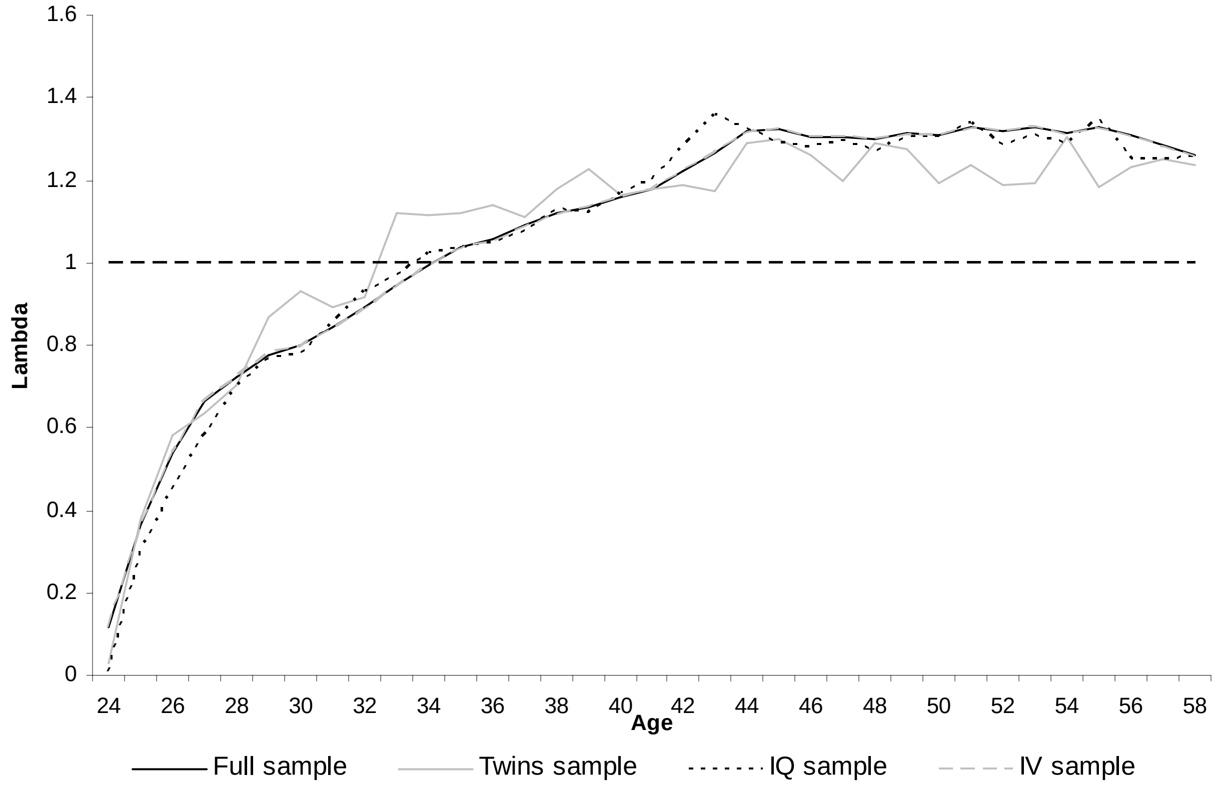
Note: This figure plots OLS estimates of the returns to schooling based on Norwegian cross-sections from 1967 to 2008. In each year, we estimate the returns to schooling separately for the sample of males aged 16-64 (with positive earnings) and for the subsample of males aged 32-33 (with positive earnings). The squared dots represent estimates of the returns to schooling for the samples aged 16-64. The triangular dots represent estimates of the returns to schooling for the samples aged 32-33. The figure also includes linear trends for the two sets of returns to schooling estimates.

Figure 4 Returns to schooling reported in commonly cited studies



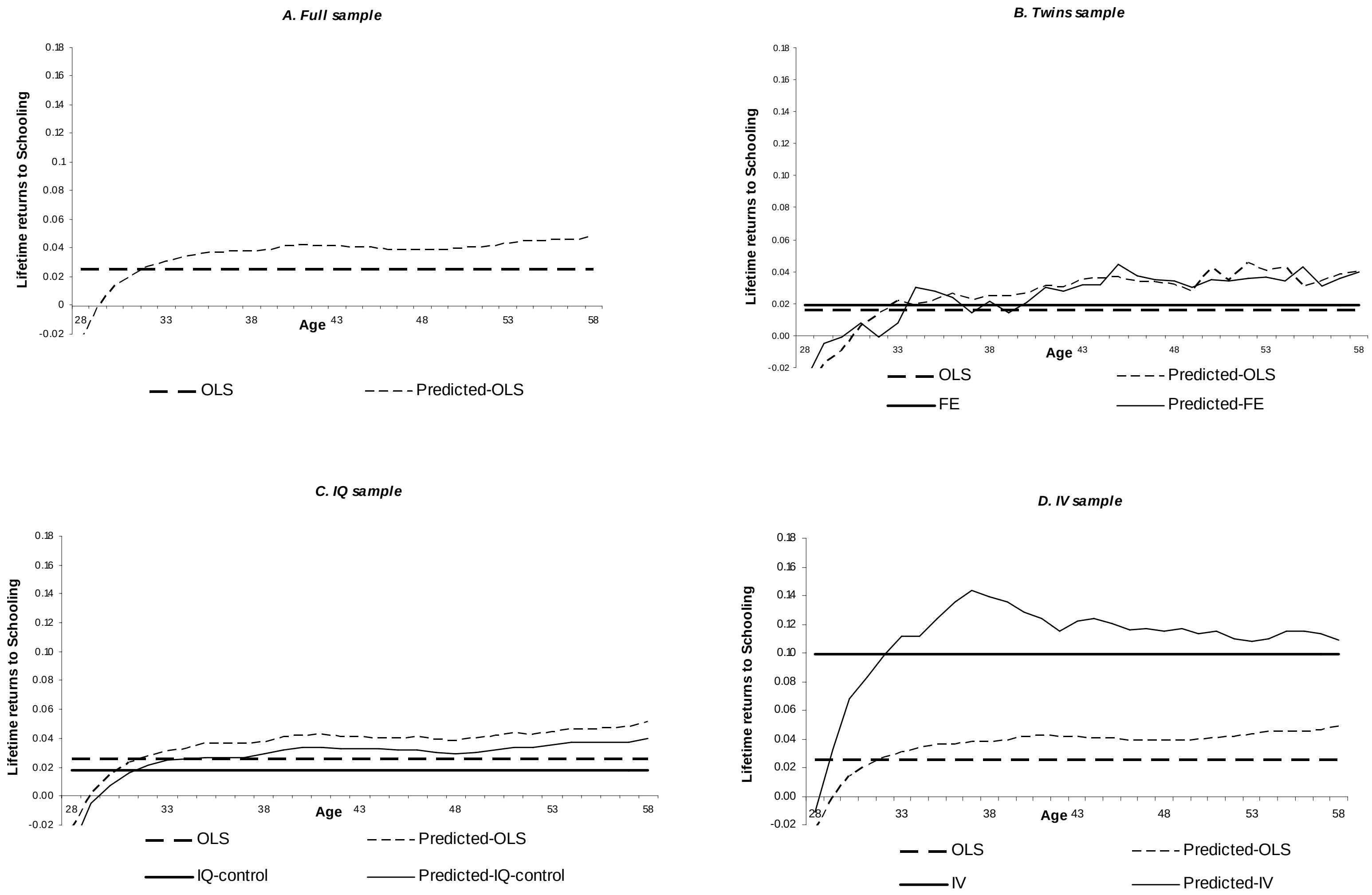
Note: This figure displays OLS estimates of the returns to schooling from the cross-section studies reported in the review articles by Card (1999), Harmon et al. (2003), Oreopolous (2006) and Devereux and Fan (2011). We only report estimates from the Anglo-Saxon countries, which includes information about the mean age in the sample. The years in which earnings and age are measured are reported in parentheses.

Figure 5 Association between current and lifetime earnings



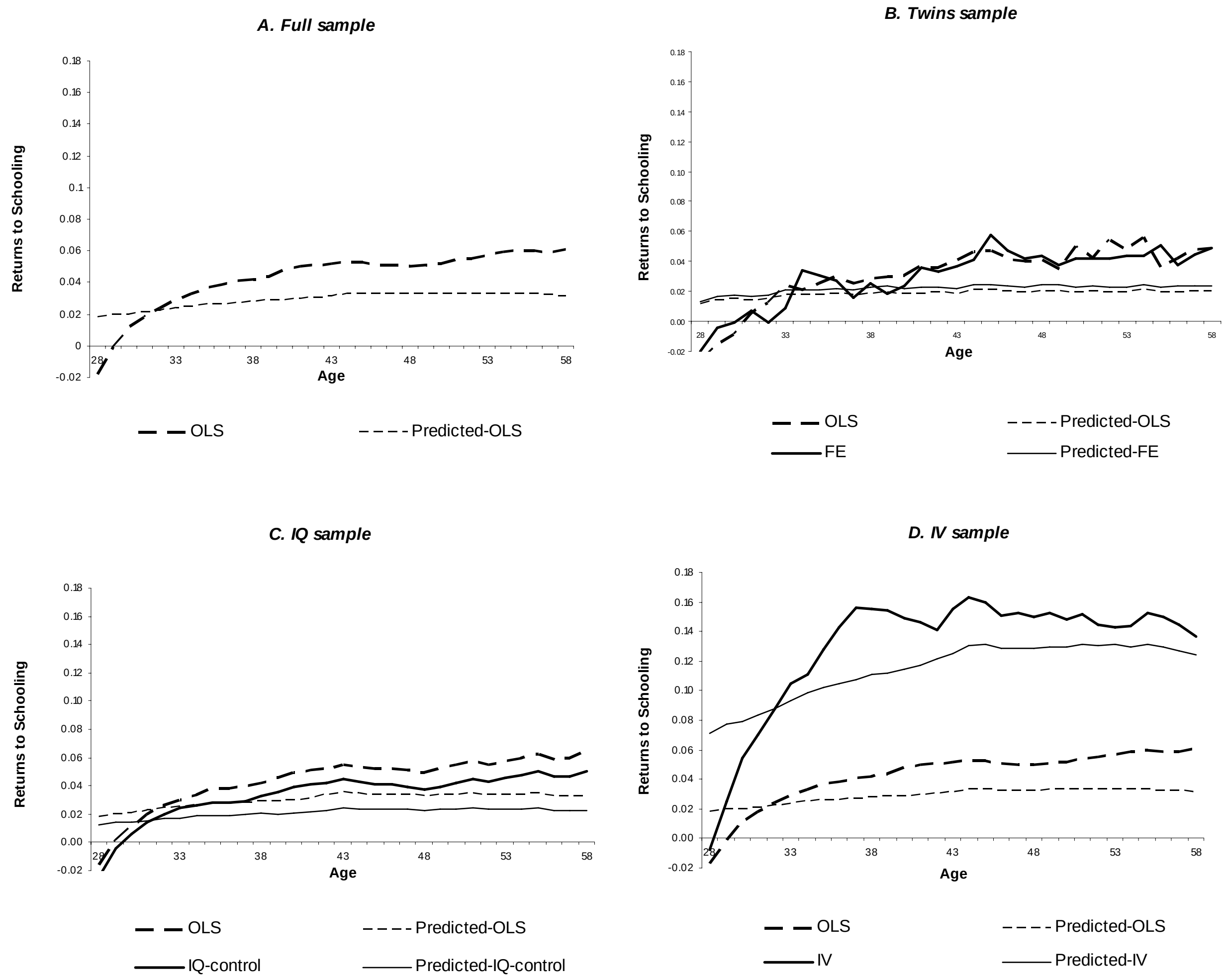
Note: This figure plots estimates of the slope coefficient λ_t from a regression of current earnings at age *t* on lifetime earnings, see equation (4). All regressions are performed separately for each sample. See notes in Table 1 for sample details.

Figure 6 Estimated and predicted lifetime return to schooling



Note: This figure uses estimates of λ_t and estimates of the returns to schooling in current earnings at age *t*, ρ_t, to plot the age-specific predicted returns to schooling in lifetime earnings, $\hat{\rho}(t) = \frac{\rho_t}{\lambda_t}$. See notes in Table 1 for sample details.

Figure 7 Estimated and predicted life-cycle variation in returns to schooling



Note: This figure uses estimates of λ_t and estimates of the returns to schooling in lifetime earnings, ρ, to plot the predicted returns to schooling in current earnings at age *t*, $\hat{\rho}_t = \rho \cdot \lambda_t$. See notes in Table 1 for sample details.

Tables

Table 1 Descriptive statistics

	Full sample		Twins sample		IQ sample		IV sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	Mean	Std.dev.	Mean	Std.dev.	Mean	Std.dev.	Mean	Std.dev.
Current earnings								
Age 28	235 063	(78 512)	232 509	(74 640)	236 010	(77 833)	235 378	(78 603)
Age 38	306 293	(123 271)	296 695	(102 809)	313 922	(128 760)	307 102	(123 656)
Age 48	336 366	(220 884)	322 534	(167 777)	342 822	(220 417)	337 457	(223 535)
Age 58	314 267	(201 531)	307 529	(154 548)	319 568	(201 888)	314 846	(202 647)
Lifetime earnings								
Age 20-58	255 859	(85 364)	249 850	(66 021)	259 289	(88 793)	256 445	(85 854)
Years of schooling	11.5	(3.0)	11.2	(3.0)	11.6	(3.0)	11.5	(3.0)
Father college	0.11	(0.32)	0.11	(0.31)	0.11	(0.32)	0.11	(0.32)
Mother college	0.05	(0.21)	0.05	(0.21)	0.05	(0.21)	0.05	(0.21)
Observations	56,832		702		14,938		53,915	

Notes: Full sample: Males born 1948-1950 with positive earnings from age 28-58. Twins sample: Male twins born 1948-1950, with positive earnings from age 28-58. IQ sample: Males born 1950, with positive earnings from age 28-58 and non-missing observations on IQ tests scores. IV sample: Male cohorts born 1948-1950, with positive earnings from age 28-58 and childhood municipality of residence for which we are able to identify the timing of the compulsory schooling reform. Schooling is measured at age 40. Father's and mother's education is represented by indicators for whether they have attained college/university degree by 1960.

Table 2 Returns to schooling in current and lifetime earnings

	Dependent variable: Log (earnings)						
	Full sample	Twins sample		IQ sample		IV sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	OLS	FE	OLS	IQ-control	OLS	IV
Panel A: Current earnings							
Age 28	-0.018***	-0.027**	-0.020	-0.016***	-0.025***	-0.017***	-0.008
	(0.002)	(0.013)	(0.015)	(0.002)	(0.003)	(0.002)	(0.009)
Age 33	0.029***	0.024***	0.009	0.030***	0.024***	0.029***	0.105***
	(0.001)	(0.007)	(0.010)	(0.001)	(0.002)	(0.001)	(0.007)
Age 38	0.042***	0.029***	0.025**	0.042***	0.033***	0.042***	0.155***
	(0.001)	(0.007)	(0.011)	(0.000)	(0.002)	(0.001)	(0.009)
Age 43	0.052***	0.041***	0.037***	0.055***	0.045***	0.052***	0.155***
	(0.001)	(0.006)	(0.011)	(0.001)	(0.002)	(0.001)	(0.010)
Age 48	0.050***	0.041***	0.044***	0.049***	0.037***	0.050***	0.150***
	(0.001)	(0.006)	(0.014)	(0.001)	(0.002)	(0.001)	(0.009)
Age 53	0.057***	0.048***	0.044***	0.058***	0.046***	0.057***	0.143***
	(0.001)	(0.011)	(0.008)	(0.000)	(0.002)	(0.001)	(0.008)
Age 58	0.061***	0.049***	0.049***	0.065***	0.050***	0.061***	0.137***
	(0.001)	(0.009)	(0.013)	(0.002)	(0.001)	(0.001)	(0.009)
Panel B: Lifetime earnings							
Age 20-58	0.025***	0.016***	0.019**	0.026***	0.018***	0.025***	0.099***
	(0.001)	(0.005)	(0.008)	(0.001)	(0.001)	(0.000)	(0.006)
Age 20-52	0.021***	0.012**	0.015*	0.021***	0.014***	0.021***	0.092***
	(0.001)	(0.005)	(0.008)	(0.001)	(0.001)	(0.001)	(0.006)
Observations	56,832	702		14,938		53,915	

Note: * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are robust to heteroskedasticity and clustered at the municipality level. See notes in Table 1 for sample details. Cohort dummies, municipality fixed effects and family background variables are included in all regressions. First stage IV estimation results are given in Table A.2.

Table 3 Returns to schooling in cross-sectional and lifetime earnings

	Dependent variable: Log (earnings)						
	Full sample	Twins sample		IQ sample		IV sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	OLS	FE	OLS	IQ-control	OLS	IV
Panel A: Cross-section earnings							
Cross-section 1985							
Return to schooling	0.030***	0.022***	0.011	0.029***	0.019***	0.030***	0.109***
	(0.001)	(0.004)	(0.009)	(0.001)	(0.001)	(0.001)	(0.005)
Mean age	32.9	32.9		31.9		32.9	
Cross-section 1995							
Return to schooling	0.050***	0.049***	0.043***	0.050***	0.039***	0.050***	0.139***
	(0.000)	(0.004)	(0.009)	(0.000)	(0.001)	(0.000)	(0.006)
Mean age	42.9	42.9		41.9		42.9	
Cross-section 2005							
Return to schooling	0.060***	0.052***	0.048***	0.061***	0.049***	0.060***	0.135***
	(0.001)	(0.006)	(0.009)	(0.001)	(0.001)	(0.001)	(0.005)
Mean age	52.9	52.9		51.9		52.9	
Panel B: Lifetime earnings							
Age 20-58 (imputed)	0.026***	0.025***	0.019***	0.027***	0.019***	0.026***	0.087***
	(0.001)	(0.003)	(0.005)	(0.000)	(0.000)	(0.000)	(0.006)
Age 20-52	0.021***	0.019***	0.015***	0.023***	0.015***	0.021***	0.082***
	(0.000)	(0.003)	(0.005)	(0.000)	(0.000)	(0.001)	(0.006)
Observations	180,730	2,288		113,247		171,703	

Note: * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are robust to heteroskedasticity and clustered at the municipality level. Cohort dummies, municipality fixed effects and family background variables are included in all regressions. First stage IV estimation results are given in Table A.2. Full sample: Males born 1948-1956 with positive earnings from age 28-58. Twins sample: Male twins born 1948-1956, with positive earnings from age 28-58. IQ sample: Males born 1950-1956, with positive earnings from age 28-58 and non-missing observations on IQ tests scores. IV sample: Male cohorts born 1948-1956, with positive earnings from age 28-58 and childhood municipality of residence for which we are able to identify the timing of the compulsory schooling reform. See Appendix for a detailed description of the method used to impute earnings.

Appendix

Table A.1 Description of variables

Variables	Description	Data source
Earnings		
Current earnings	The log of annual real earnings in a given year. Our earnings measure (*pensjonsgivende inntekt*) is the sum of pretax market income from wages, self-employment and work-related cash transfers, including unemployment benefits, sick leave benefits, and parental leave benefits. Annual earnings are adjusted for inflation and real wage growth using the standards of the Norwegian social security system.	Administrative Tax Records, 1967-2008
Lifetime earnings	The log of the annuitized value of annual real earnings from age 20 to 58, calculated using an annual real interest rate of 2.3 percent.	Administrative Tax Records, 1967-2008
Lifetime earnings (imputed)	Cohorts 1948-1950: The log of the annuitized value of annual real earnings from age 20 to 58 Cohorts 1950-1953: The log of the annuitized value of the sum of annual real earnings from age 20 to 55 and imputed earnings from age 56 to 58. Cohorts 1954-1956: The log of the annuitized value of the sum of annual real earnings from age 20 to 52 and imputed earnings from age 53 to 58.	Administrative Tax Records, 1967-2008
Education		
Years of schooling	The number of years of schooling corresponding to the highest completed level of education the individual has attained before turning 40.	National Education Database, 1970-2008
Family background		
Mother college	Indicator for whether the mother has attained a college degree according to the 1960 Census data.	National Population and Housing Census, 1960
Father college	Indicator for whether the father has attained a college degree according to the 1960 Census data.	National Population and Housing Census, 1960
Family income	Indicators for parent's position (quartile) in the distribution of family income (sum of mother's and father's taxable income) in 1970.	National Population and Housing Census, 1970
Other variables		
IQ test score	Full set of indicators for IQ test scores. The test scores are reported on a standard nine scale.	Norwegian Military Records, 1968-2008
Reform indicator	Indicator for whether the individual grew up in a municipality that implemented the education reform increasing the compulsory schooling from 7 to 9 years by the time the individual was expected to complete 7 years of pre-reform compulsory schooling (normally at age 14).	Lie (1973, 1974), Telhaug (1969), Aakvik et al (2010)
Vocational college	Indicator for whether the individual grew up in a municipality that had vocational college in its close proximity prior to the compulsory schooling reform.	Historical Education Records, 1963
Upper secondary	Indicator for whether the individual grew up in a municipality that had upper secondary school in its close proximity prior to the compulsory schooling reform.	Historical Education Records, 1963
Regional college	Indicator for whether the individual grew up in a municipality that had regional college in its close proximity prior to the compulsory schooling reform.	Historical Education Records, 1963
University	Indicator for whether the individual grew up in a municipality that had university in its close proximity prior to the compulsory schooling reform.	Historical Education Records, 1963

First stage results

Table A.2 presents results from separate estimations of equation (3) for cohorts 1948-1950 (columns 1 and 2), 1951-1953 (columns 3 and 4), 1954-1956 (columns 5 and 6), and the pooled sample of cohorts 1948-1956 (columns 7 and 8). We see that exposure to the compulsory education reform increased the number of years of schooling by nearly one-third of a year. There is also some evidence of heterogeneous responses to the reform. The change in compulsory schooling law had smaller impact on educational attainment of individuals with high educated mothers and rich parents. Moreover, we see that the reform effects were stronger among individuals who grew up in municipalities in close proximity to other school types, especially regional colleges and universities, and therefore had the possibility of pursuing higher education after completing the compulsory school. The first stage results are fairly similar across cohorts. We can also see that the first stages are strong with F-statistics on the excluded instruments exceeding 43, which means that we do not need to worry about problems due to weak instruments.

Table A.2 First stage results for IV estimations

	Dependent variable: Years of schooling							
	Cohorts 1948-1950		Cohorts 1951-1953		Cohorts 1954-1956		Cohorts 1948-1956	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Controls								
Father college	1.400***	(0.043)	1.305***	(0.048)	1.215***	(0.071)	1.346***	(0.028)
Mother college	1.170***	(0.061)	1.277***	(0.069)	1.106***	(0.102)	1.204***	(0.041)
Family income 2	0.393***	(0.035)	0.502***	(0.042)	0.325***	(0.061)	0.431***	(0.024)
Family income 3	0.848***	(0.037)	0.935***	(0.044)	0.665***	(0.062)	0.868***	(0.025)
Family income 4	1.692***	(0.039)	1.632***	(0.045)	1.361***	(0.068)	1.637***	(0.026)
Instruments								
Reform dummy	0.331**	(0.149)	0.356***	(0.085)	0.295***	(0.088)	0.426***	(0.043)
Reform x Mother college	-0.433**	(0.192)	-0.170	(0.105)	0.020	(0.114)	-0.113*	(0.059)
Reform x Father college	-0.077	(0.125)	0.086	(0.073)	0.081	(0.080)	-0.032	(0.041)
Reform x Family income 2	-0.101	(0.110)	-0.030	(0.070)	0.074	(0.072)	-0.020	(0.039)
Reform x Family income 3	-0.034	(0.111)	-0.189***	(0.070)	0.070	(0.073)	-0.129***	(0.039)
Reform x Family income 4	-0.103	(0.115)	-0.158**	(0.071)	-0.042	(0.078)	-0.267***	(0.040)
Reform x Vocational college	0.195***	(0.045)	0.106**	(0.051)	0.149***	(0.072)	0.161***	(0.029)
Reform x Upper secondary	0.093***	(0.045)	0.175***	(0.050)	0.030	(0.069)	0.110***	(0.028)
Reform x Regional college	0.422***	(0.046)	0.486***	(0.052)	0.587***	(0.068)	0.496***	(0.029)
Reform x University	0.330***	(0.051)	0.449***	(0.057)	0.318***	(0.073)	0.450***	(0.032)
F-value (instruments)	43.17		64.01		85.91		192.84	
Observations	53,915		57,332		60,456		171,703	

Note: * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are robust to heteroskedasticity and clustered at the municipality level. See notes in Table 1 and Table 3 for sample details. Cohort dummies and municipality fixed effects are included in all regressions.

Nearest neighbor matching for imputation of missing earnings for cohorts born 1951-1956

Consider an individual *i* born in 1953, for which we want to impute earnings for age 56-58. The nearest neighbor matching algorithm identifies the best individual match j^* among the 1948-1950 cohorts. The best match is defined as the individual observation that minimizing the Mahalanobis distance in annual real earnings from age 20 to age 55, between the individual and the potential matches, conditional on a set of covariates (see Rosenbaum and Rubin, 1985). The minimization problem as be expressed as finding an individual match j^* for individual *i* such that

$$(6) \qquad j^* = \min_{j \in D} \left\{ \sqrt{\sum_{t=20}^{55} \left(\frac{Y_{it} - Y_{jt}}{Z_t} \right)^2} \right\},$$

where Y_{it} is annual real earnings at age *t* for individual *i*, and Y_{jt} is annual real earnings at age *t* for a potential match $j \in D$, where D contains all individuals born in 1948-1950, who have the same value on the reform indicator, level of schooling, family background characteristics, and childhood county, as individual *i*. In order to construct Mahalanobis distance, we must weight the deviations $Y_{it} - Y_{jt}$ by the sample variance in annual earnings at age *t*, denoted by Z_t.

By following this procedure, we find matches cohorts born 1948-1950 for each individual born 1951-1953 and 1954-1956, by minimizing the Mahalanobis distance in annual real earnings from age 20 to age 55 and from age 20 to 52, respectively. Next, we impute the missing earnings observations after age 52 for cohorts 1954-1956, and after age 55 for cohorts 1951-1953, based on the earnings records of the individual matches.[16] The matching algorithm allows us to construct measures of lifetime earnings from age 20 to 58 and estimate returns to schooling in lifetime earnings for cohorts born 1951-1956. The results are given in panel B of Table A.3.

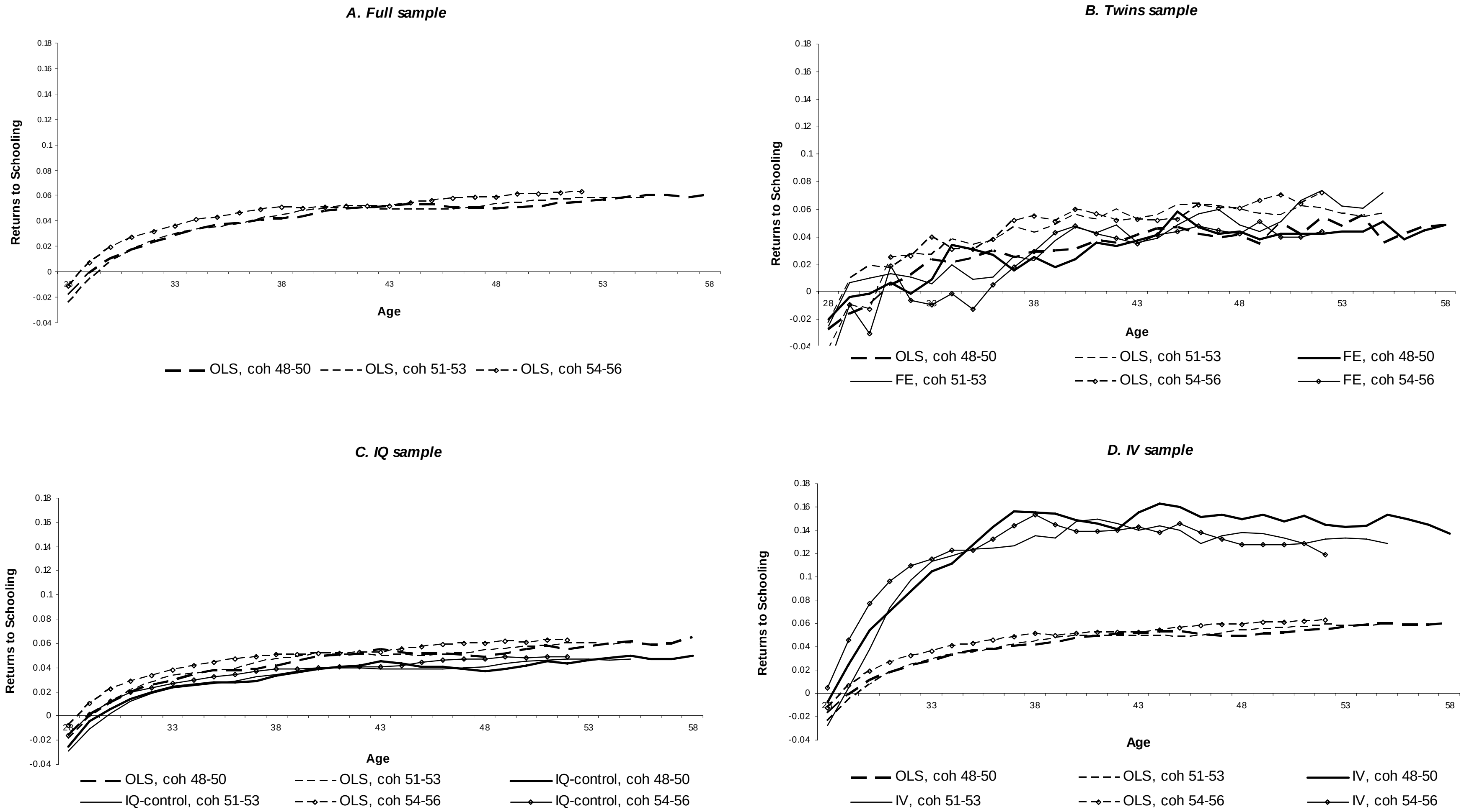
[16] To test the matching method, we have performed out-of-sample checks for ages where we have complete earnings records. These out-of-sample results suggest that the matching method performs very well.

Table A.3 Returns to schooling in current and lifetime earnings for birth cohorts 1951-1956

	Dependent variable: Log (earnings)													
	Birth cohorts 1951-1953							Birth cohorts 1954-1956						
	Full sample	Twins Sample		IQ sample		IV Sample		Full sample	Twins sample		IQ sample		IV sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	OLS	OLS	FE	OLS	IQ-control	OLS	IV	OLS	OLS	FE	OLS	IQ-control	OLS	IV
Panel A: Current earnings														
Age 28	-0.024***	-0.022**	-0.025	-0.019***	-0.029***	-0.024***	-0.028***	-0.011***	-0.042**	-0.054*	-0.008***	-0.016***	-0.012***	0.005
	(0.002)	(0.011)	(0.018)	(0.001)	(0.002)	(0.002)	(0.010)	(0.002)	(0.017)	(0.029)	(0.002)	(0.002)	(0.002)	(0.010)
Age 33	0.030***	0.027***	0.006	0.033***	0.023***	0.030***	0.113***	0.036***	0.040***	-0.009	0.038***	0.027***	0.036***	0.115***
	(0.001)	(0.010)	(0.012)	(0.001)	(0.001)	(0.001)	(0.009)	(0.001)	(0.012)	(0.018)	(0.001)	(0.001)	(0.001)	(0.008)
Age 38	0.045***	0.043***	0.023**	0.047***	0.034***	0.045***	0.135***	0.051***	0.055***	0.029**	0.051***	0.039***	0.051***	0.153***
	(0.001)	(0.007)	(0.010)	(0.001)	(0.001)	(0.001)	(0.010)	(0.001)	(0.015)	(0.013)	(0.001)	(0.001)	(0.001)	(0.008)
Age 43	0.049***	0.053***	0.036**	0.050***	0.039***	0.050***	0.140***	0.052***	0.053***	0.035**	0.053***	0.041***	0.052***	0.140***
	(0.001)	(0.009)	(0.016)	(0.001)	(0.002)	(0.001)	(0.010)	(0.001)	(0.012)	(0.011)	(0.001)	(0.002)	(0.001)	(0.007)
Age 48	0.053***	0.060***	0.049***	0.054***	0.041***	0.054***	0.138***	0.059***	0.061***	0.042***	0.060***	0.047***	0.059***	0.128***
	(0.001)	(0.009)	(0.017)	(0.001)	(0.001)	(0.001)	(0.009)	(0.001)	(0.012)	(0.013)	(0.001)	(0.001)	(0.001)	(0.007)
Age 52	0.058***	0.061***	0.074***	0.060***	0.047***	0.059***	0.132***	0.063***	0.072***	0.044**	0.063***	0.049***	0.063***	0.119***
	(0.001)	(0.009)	(0.019)	(0.001)	(0.001)	(0.001)	(0.010)	(0.001)	(0.016)	(0.017)	(0.001)	(0.001)	(0.001)	(0.007)
Age 55	0.058***	0.057***	0.072***	0.060***	0.047***	0.059***	0.129***	-	-	-	-	-	-	-
	(0.001)	(0.009)	(0.018)	(0.001)	(0.001)	(0.001)	(0.009)							
Panel B: Lifetime earnings														
Age 20-58	0.024***	0.029***	0.025***	0.026***	0.017***	0.024***	0.086***	0.029***	0.030***	0.015*	0.030***	0.021***	0.029***	0.086***
(imputed)	(0.001)	(0.005)	(0.008)	(0.001)	(0.001)	(0.001)	(0.007)	(0.000)	(0.009)	(0.008)	(0.001)	(0.001)	(0.001)	(0.007)
Age 20-52	0.019***	0.025***	0.021**	0.021***	0.013***	0.019***	0.080***	0.025***	0.025***	0.010	0.026***	0.017***	0.025***	0.083***
	(0.001)	(0.005)	(0.008)	(0.001)	(0.001)	(0.001)	(0.007)	(0.001)	(0.009)	(0.010)	(0.001)	(0.001)	(0.001)	(0.005)
Obs.	60,416	810		47,221		57,332		63,482	774		51,088		60,977	

Note: * p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors adjusted for clustering at the municipality level are given in parentheses. See notes in Table 1 and Table 3 for sample details. Cohort dummies, municipality fixed effects and family background variables are included in all regressions. First stage IV estimation results are given in Table A.2. See Appendix for a detailed description of the method used to impute earnings.

Figure A.1 Returns to schooling in current earnings for birth cohorts 1948-1956



Note: See notes in Table 1 for sample details. Cohort dummies, municipality fixed effects and family background variables are included in all regressions. First stage IV estimation results are given in Table A.2

Top-coded earnings data

Prior to 1986, our earnings data are top-coded, though at fairly high levels. In fact, less than 3 percent of the observations have right-censored earnings in any given year.[17] Yet to make sure that top-coding is not driving our results, we follow Atkinson (2005) in using a Pareto distribution to simulate earnings above the top-coding threshold. The Pareto distribution is known to be a desirable approximation of the uppermost part of earnings and distributions.

The Pareto distribution has following CDF

$$(7) \qquad G(y) = 1 - \left(\frac{\theta}{y}\right)^{\alpha}, \theta > 0, \alpha > 0, y \geq \theta,$$

and is thus fully characterized by parameters α and θ. $G^{-1}(q)$ denotes the q-quantile in the distribution G. Let $G^{-1}(q_2)$ be the top-censoring earnings threshold, where q_2 is the share of population with earnings below this threshold. Following Atkinson (2005), we estimate α by the following estimator

$$(8) \qquad \hat{\alpha} = \frac{\log\left(\dfrac{1-q_2}{1-q_1}\right)}{\log\left(\dfrac{G^{-1}(q_1)}{G^{-1}(q_2)}\right)},$$

where $G^{-1}(q_1)$ is some lower level of earnings with cumulative share given by q_1. For a given year, we choose the following three values of $G^{-1}(q_1)$: 90 %, 95 % and 99 % of the year's top-censoring threshold. From the estimator given in equation (8), we get three different estimates of α corresponding to the three choices of $G^{-1}(q_1)$. Using the average value of the three estimates of α, we estimate parameter θ as $\hat{\theta} = (1-q_2)^{1/\hat{\alpha}} G^{-1}(q_2)$, after inverting the CDF given in equation (7).

We estimate α and θ separately for each year between 1967 and 1985. In each year, we simulate as many observations from the estimated Pareto distribution as the number of top-censored observations. Next, the top-censored earnings are then replaced by the simulated earnings. Finally, we estimate returns to schooling in current and lifetime earnings using the

[17] The top-coding in the Norwegian earnings data is considerably less severe than in the earnings data provided by the US Social Security Administration, where between 22.5 and 62.2 percent of the sample is right-censored in the years 1960-1980 (Haider and Solon, 2006). Moreover, our calculations show that most individuals in our sample escaped top-coding during 1971-1975 and top-coding is not present in the earnings data for 1981.

simulated top-earnings data, for each of our samples. The results are given in Table A.4.[18] It is reassuring to find that the estimates of returns to schooling barely move.

Table A.4 Returns to schooling using simulated top-earnings

	Dependent variable: Log (earnings)						
	Full sample	Twins sample		IQ sample		IV sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	OLS	FE	OLS	IQ-control	OLS	IV
Panel A: Current earnings							
Age 28	-0.017***	-0.027**	-0.019	-0.016***	-0.025***	-0.017***	-0.008
	(0.002)	(0.013)	(0.015)	(0.002)	(0.003)	(0.002)	(0.009)
Age 33	0.029***	0.024***	0.009	0.030***	0.024***	0.029***	0.105***
	(0.001)	(0.007)	(0.010)	(0.001)	(0.002)	(0.001)	(0.007)
Age 38	0.042***	0.029***	0.025**	0.042***	0.033***	0.042***	0.155***
	(0.001)	(0.007)	(0.011)	(0.000)	(0.002)	(0.001)	(0.009)
Age 43	0.052***	0.041***	0.037***	0.055***	0.045***	0.052***	0.155***
	(0.001)	(0.006)	(0.011)	(0.001)	(0.002)	(0.001)	(0.010)
Age 48	0.050***	0.041***	0.044***	0.049***	0.037***	0.050***	0.150***
	(0.001)	(0.006)	(0.014)	(0.001)	(0.002)	(0.001)	(0.009)
Age 53	0.057***	0.048***	0.044***	0.058***	0.046***	0.057***	0.143***
	(0.001)	(0.011)	(0.012)	(0.000)	(0.002)	(0.001)	(0.008)
Age 58	0.061***	0.049***	0.049***	0.065***	0.050***	0.061***	0.137***
	(0.001)	(0.009)	(0.013)	(0.002)	(0.001)	(0.001)	(0.009)
Panel B: Lifetime earnings							
Age 20-58	0.025***	0.017***	0.019**	0.026***	0.018***	0.025***	0.099***
	(0.001)	(0.005)	(0.008)	(0.001)	(0.001)	(0.000)	(0.006)
Age 20-52	0.021***	0.012*	0.015*	0.021***	0.014***	0.021***	0.092***
	(0.001)	(0.005)	(0.008)	(0.001)	(0.001)	(0.001)	(0.006)
Observations	56,832	702		14,938		53,915	

Note: * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses are robust to heteroskedasticity and clustered at the municipality level. See notes in Table 1 for sample details. Top-censored earnings are simulated from a Pareto distribution. Cohort dummies, municipality fixed effects and family background variables are included in all regressions. First stage IV estimation results are given in Table A.2

[18] As an out-of-sample test of the simulation method, we perform the same exercise using earnings data for 1986 where there is no top-censoring. The simulated earnings using the Pareto method are very similar to the actual earnings. In fact, there is hardly any difference in the Gini coefficients (and other inequality measures) for the actual earnings distribution and the earnings distribution with simulated top-coded earnings.