Title
Unobserved Ability, Comparative Advantage, and the Rising Return to Education in the United States 1979-2000

Author
Deschenes, Olivier

Publication Date
2001-09-01

UNOBSERVED ABILITY, COMPARATIVE ADVANTAGE, AND THE RISING RETURN TO EDUCATION IN THE UNITED STATES 1979-2000.

Olivier Deschênes[1]

Department of Economics
University of California, Santa Barbara
Santa Barbara, CA 93106-9210

September 2001

Abstract: This paper quantifies the change in the causal effect of education on labor market earnings in the United States between 1979 and 2000. In absence of valid instrumental variables for schooling, a causal model for earnings and schooling that incorporates heterogeneity in absolute and comparative advantage across individuals is used to impose some structure on the observed relationship between schooling and earnings. A simple intuition arises from the model: if individuals with higher returns to schooling acquire more schooling, the relationship between log earnings and schooling will be convex in the population. Likewise, for a fixed cohort of individuals, the degree of convexity will rise over time if the causal return to education rises. Differences across cohorts in the mapping between schooling and ability will lead to permanent differences in the profiles of the earnings-schooling relationship. Changes in the observed relationship between schooling and earnings can therefore be decomposed into year-specific and cohort-specific factors corresponding to causal and confounding components. Using CPS data for cohorts of men born between 1930 and 1970, I find that the causal return to education increased by 30% between 1979 and 2000, after controlling for the confounding effects of time-varying ability and comparative advantage biases across cohorts.

Keywords: causal effect of education, unobserved ability, comparative advantage, self-selection, correlated random coefficients, repeated cross-sections.

---

[1] email: olivier@econ.ucsb.edu
I am indebted to David Card for many helpful suggestions and continual encouragement during the course of this research project. I also thank Orley Ashenfelter, Ken Chay, Bo Honoré, Dean Hyslop, Jeff Kling, Alan Krueger, Marco Manacorda, Cecilia Rouse, Till von Wachter, and seminar participants at Cornell, Princeton, University of California Santa Barbara, University College London, University of Illinois at Urbana-Champaign, Université de Montréal, University of Rochester and the University of Toronto for helpful suggestions.

1. Introduction.

In recent years, many researchers have attempted to measure the causal link between education and labor market earnings.[1] This interest is generated in part by public policy debates over investments in education, in part by the recent availability of better data sets, and in part by methodological advances in resolving the identification problems related to unobservable ability and endogenous schooling.[2]

At the same time, the wage structure of the U.S. labor market changed profoundly.[3] Since the early 1980s, the conventional measure of the economic return to education—the schooling coefficient in a human capital earnings regression[4]—almost doubled. Residual wage dispersion has also been increasing during that time period, a pattern that is often attributed to a rise in the return to unobservable skills, such as motivation or cognitive ability.[5]

These concurrent changes in the wage structure heighten the uncertainty regarding the causal link between schooling and earnings, especially whether it has changed during the last decades. One interpretation of the rising correlation between earnings and schooling is that there has been an increase in the causal effect of education on labor market earnings. This may be due to an increased demand for better-educated workers, or to the entry in the labor market of younger cohorts with higher marginal productivity of schooling (perhaps because of an improvement in school quality). An alternative interpretation of the increase in the conventionally measured return to schooling is based on unobservable variables that affect earnings and are correlated with schooling. A rise in the return to unobservable skills, or an increasing degree of ability-education sorting across cohorts might also have contributed to a rising correlation between schooling and earnings.[6]

What both of these interpretations have in common is the idea that changes in the educational wage structure can occur because of year effects (or "economy-wide" effects) and composition

[1] See for example Angrist and Krueger (1991), Ashenfelter and Rouse (1998), and Meghir and Palme (1999). Card (1999, 2000) presents a review of recent studies.

[2] Most advancements relate to the estimation of models with heterogeneous treatment effect using instrumental variables. See Angrist and Imbens (1995), Heckman and Vytlacil (1998), Card (1999, 2000), and Heckman, Tobias and Vytlacil (2000).

[3] See Katz and Autor (1999) for an extensive overview of recent studies of the wage structure in the U.S.

[4] Mincer (1974) showed that the schooling coefficient in a log earnings regression can be interpreted as the "return to education" if the only cost of schooling are foregone earnings and if the return to schooling is independent of schooling levels. In this paper, I will refer to this estimate as the "conventional measure of the return to schooling".

[5] See Juhn, Murphy and Pierce (1993).

[6] See Taber (1999), Blackburn and Neumark (1993), and Murnane, Willett and Levy (1995).

effects (or "cohort" effects). In the first interpretation these effects lead to a change in the average causal effect of education. In the second, they lead to time-varying "ability biases" in the observed relationship between earnings and schooling. While previous analysts have evaluated some of these explanations, the absence of a unifying framework addressing all of these issues simultaneously has greatly limited the scope of our understanding of the recent changes in the educational wage structure.

The purpose of this paper is to quantify the changes in the causal effect of education on labor market earnings during the last two decades in the United States. I begin by setting out a model describing the relationship between earnings and schooling in the context of repeated cross-sectional data, with unobserved absolute ability and heterogeneous returns to education. If the schooling decisions of individuals are influenced by unobserved ability, the average causal effect of education embedded in the conventional estimate of the return to education will be confounded. The two-factor model illustrates how the intertemporal changes in the average causal effect and in the confounding elements can be decomposed into year-specific and cohort-specific factors. Changes in the conventional measure of the return to education are potentially confounded through two distinct channels: (i) changes in the return to unobserved ability over time, and (ii) changes in the mapping between ability and completed education across cohorts.

A key implication of the model is that if people who have higher returns to education tend to acquire more schooling (i.e. if there are comparative advantage incentives in schooling decisions), the observed relationship between earnings and schooling will be convex (Mincer 1974, Rosen 1977). Moreover, for a fixed cohort of individuals the degree of convexity will increase over time if the year-specific component of the causal return to education rises. This simple prediction of the model separately identifies the year-specific causal return to education from the year-specific return to unobserved ability, using the coefficients from an augmented human capital earnings regression.

The empirical analysis is based on repeated cross-sectional data from the Current Population Survey for cohorts of men born between 1930 and 1970. I begin by documenting the marked increase in the convexity of the relationship between log earnings and schooling over the 1980s and 1990s that was noted by Mincer (1996). I then make use these changes in the U.S. wage structure to identify the parameters of the model. I find that the two-factor model of ability, schooling and earnings provides a relatively accurate description of the changes in the educational wage structure over the last twenty years. The estimates imply an increase in the causal return to education of about 30% between 1979 and 2000, as opposed to 50% for the conventional estimate of the return to schooling. The return to unobserved ability followed a different pattern, increasing by about 10% over the 1980s and decreasing by the same amount over the 1990s. The estimates indicate an important increasing inter-cohort trend in the correlation between educational attainment and individual-specific returns to education. This is consistent with a model where returns to schooling vary across individuals, and where comparative advantage incentives play a more significant role in the schooling decisions of younger cohorts. Taken as a whole, the evidence in this paper provides no indication that the rise in the conventional measure of the return to education is due to a time-varying unobserved ability bias.

The remainder of the paper is organized as follows. Section 2 provides a brief overview of changes in the educational wage structure between 1979 and 2000 in the United States. Section 3 discusses the identification of changes in the causal effect of education on earnings with repeated cross-sectional data. An empirical model of earnings and schooling with two factors of unobserved heterogeneity across the population is developed. Section 4 presents the estimation method used to decompose the estimated coefficients from a series of augmented earnings regressions into year-specific and cohort-specific components of the changing causal effect of education. Section 5 describes the empirical results and uses the parameters of the model to interpret the recent changes in educational-related wage differentials. Section 6 presents a sensitivity analysis, and Section 7 concludes.

2. The Changing Relationship Between Earnings and Schooling, 1979-2000.

A. *The data.*

This section presents a descriptive overview of the changes in the relationship between log earnings and years of education in the United States over the last decades. The data are taken from the monthly CPS Outgoing Rotation Group Earnings Files (ORG) covering the years 1979 to 2000. This choice is motivated by several practical reasons, most importantly, by the larger sample sizes provided by the ORG as opposed to the Annual Earnings Supplement in the March CPS. In order to focus on individuals who have completed their formal schooling and made a permanent transition to the labor market, the sample is restricted to men aged 26-60, who earned more than $250 (in

1999 constant dollars) in any survey week between 1979 and 2000.[7] To maintain independence of the samples from year to year, only individuals who are in their first rotation out the CPS samples are considered for the analysis.[8]

Table 1 presents summary statistics for the data used in the empirical analysis. Birth cohorts are defined by 5 year intervals, starting with men born in 1930-34 and ending with those born between 1965-69. The aggregation of single birth year cohorts into 5 year birth cohorts ensures large enough samples when the cohorts are followed on a year-to-year basis. Moreover, this definition is fine enough to group individuals who attended elementary and secondary school together, and were subject to similar influences from the educational and economic environments (for example school quality and expected gains to an additional year of education). An interesting feature of Table 1 relates to the differences across cohorts in educational attainment. Average education displays a rising inter-cohort trend for the cohorts born before 1950, followed by a decline for those born in the 1950s and early 1960s.[9] This is illustrated with greater detail in Figures 1 and 2 where for each birth cohort (or survey year), the average level of education and the fraction of college graduates are displayed. Figure 1 shows the remarkable decline in educational attainment for cohorts born after 1950, while Figure 2 illustrates the increase in educational attainment in the working population between 1979 and 2000. Cleary, this last pattern is not the consequence of a secular increase in educational attainment between successive birth cohorts.

B. *Changes in the earnings-schooling relationship.*

The standard human capital earnings function specifies log earnings as a linear function of years of education. The log-linear specification has been used in virtually all studies of the impact of education on earnings. Implicit in this specification is the assumption that returns to schooling do not vary across the population, or that any variation is unrelated to educational attainment. In general, however, any correlation between returns to schooling and educational attainment at the individual level will engender a nonlinear relationship between log earnings and schooling in the population as a whole (Mincer 1974, Rosen 1977, Card 1999).[10]

[7] This corresponds to the weekly earnings of an individual working 40 hours per week at the 1981 federal minimum wage. The sample restrictions are similar to those used in other analyses of the wage structure (see e.g. Katz and Autor (1999)). All earnings figures in this paper are reported in 1999 constant dollars.

[8] The rotation group structure of the CPS implies an overlap of about half of the sample from year-to-year.

[9] Card and Lemieux (2001) analyze the underlying sources of these trends.

[10] For example, a positive correlation between completed education and returns to schooling would arise if individuals

In order to obtain some simple evidence on the functional relationship linking earnings and schooling, I estimated an unrestricted regression of log earnings on a set of dummy variables for each schooling level available in the data.[11] Figure 3 displays the estimated dummies for 3 time periods: 1979-1980, 1989-1990 and 1999-2000.[12] The figure suggests that between 1979 and 2000 the shape of the earnings-schooling relationship changed. During the early 1980s, log earnings and education appeared to be linearly related, with the exception of a slight nonlinearity between 15 and 16 years of completed education, as noted previously by Card and Krueger (1992) and Heckman, Layne-Farrar and Todd (1996). Over the 1980s and 1990s, however, the profile shifted downwards for lower schooling levels (below 12 years of education) and upwards at the higher end of the schooling distribution. The shift in the profile of the earnings-schooling relationship not only reflects the uniform increase in educational wage differentials over the last twenty years, but also the "acceleration" of the differentials at the higher end of the schooling distribution.

Figure 3 suggests that the salient features of the changes in the educational wage structure can be reasonably approximated by a simple quadratic relationship between log earnings and years of completed education. Based on this conclusion, I fit a simple human capital wage regression with linear and quadratic schooling terms to the data from the March CPS from 1970-1999.[13] The coefficients on the linear and quadratic terms are shown in Figure 4, along with the coefficient on years of education from a conventional human capital regression that excludes the quadratic term. This figure clearly shows that the changes in the functional form relating earnings and schooling concurred with the rise in the conventional measure of the return to education. Starting in the early 1980s, the decrease in the linear term was offset by the remarkable increase in the quadratic term, implying a rise in the conventional measure of the return to education.[14] This figure also motivates the time period chosen for this study: Since all of the increase in the conventional measure of the return to education occurred in the 1980s and 1990s, this study will focus only on that time period.

---

based their schooling decisions on their expected gain from an additional year of education.

[11]The regressions also included controls for potential experience and race.

[12]In 1992 the coding of educational attainment in the CPS changed from a measure of highest grade completed to a degree-based measure. I use the approach suggested by Jaeger (1997) to linearize the degree-based measures into single years of completed education. Transforming the pre-1992 data with the new coding did not alter any of the results in the paper.

[13]The March CPS data were used here to provide a longer time-series of microdata on earnings and schooling.

[14]Remember that the two sets of regression coefficients are related in the following way. Consider the regressions $y = a_0 + b_0 x + e$ and $y = a_1 + b_1 x + b_2 x^2 + u$. Then using the "omitted-variable" formulas, it is easy to show that $b_0 \approx b_1 + 2b_2\overline{x}$, where $\overline{x}$ is the sample average of x.

For the reason listed above, the ORG will be used in the rest the empirical analysis.

The two panels of Figure 5 complete the descriptive analysis by presenting the linear and quadratic coefficients from the augmented log earnings regressions estimated separately for each birth cohort described in Table 1. Similar to the patterns illustrated in Figure 4, the ORG data also suggests that during the 1980s and 1990s, the contribution of the linear term to the measured return to schooling decreased while the contribution of the quadratic term increased substantially, in a proportional manner across cohorts.

Taken as a whole, the evidence in Figures 4 and 5 indicates important changes in the functional form relating years of education and log earnings during the last twenty years. The similarity of the trends in the linear and quadratic schooling coefficients among cohorts points out to a common underlying mechanism affecting all cohorts equally. Differences across cohorts in the levels of the two schooling components suggest the existence of permanent differences in the relationship between log earnings and schooling across cohorts.

3. Empirical Framework.

This section develops a statistical framework for interpreting the changes in the earnings-schooling relationship outlined in the previous section. The objective is to set out a simple and empirically tractable model that will enable us to assess whether the causal effect of education changed. I begin by specifying more precisely the concept of changing causal effect of education in the context of repeated cross-sectional data. Then, I show how the parameters of the causal model can be identified from a series of augmented human capital earnings regressions.

A. *Conceptual Framework.*

Most of the conceptual issues underlying causal inference in observational studies of the relationship between earnings and schooling can be encompassed in a simple model of endogenous schooling.[15] Empirically, the implications of endogenous schooling for the earnings function are conveniently described by a correlated random coefficient model.[16] To begin, consider a simple model in which

[15] See Becker (1975), Rosen (1977), Willis and Rosen (1979) and Willis (1986) for discussions of optimizing models of schooling determination.

[16] In the correlated random coefficient model, the treatment is correlated with the components of unobserved heterogeneity. See the applications in Garen (1984), Altonji and Dunn (1996), Meghir and Palme (1999), and Dustmann and Meghir (2001). Heckman and Vytlacil (1998), and Card (1999, 2000) discuss the correlated random coefficient model in the context of schooling models. In particular, Card (2000) shows that the cross-sectional version of (1) is consistent with a general class of optimizing models.

log earning are linearly related to years of education by an individual-specific slope and intercept.[17] More specifically, suppose that the log earnings of an individuals belonging to cohort c, observed in survey year t are determined by:

$$\log y_{ict} = \Psi_t a_{ic} + \delta_t b_{ic} S_{ic} + \varepsilon_{ict} \tag{1}$$

where a_{ic} is the unobserved ability of individual i, who belongs to cohort c; b_{ic} is the return to education specific to person i in cohort c; and S_{ic} denotes years of completed education. The factor loadings Ψ_t and δ_t represent the year-specific payoffs to ability and education common to all workers in year t, relative to a base period. In this model individual heterogeneity in earnings capacity affects both the intercept and the slope of the earnings function. At a given point in time, the variation in log earnings among individuals with the same level of schooling, and belonging to same cohort arises from two sources: individuals with higher values of a_{ic} will have higher earnings at all schooling levels (i.e. an "absolute advantage"), and individuals with higher values of b_{ic} will receive higher payoffs per year of education (i.e. a "comparative advantage").

According to this specification, the average causal effect of education in the population can vary over time for two reasons. First, if the average marginal productivity of schooling varies by cohort (i.e. different cohorts have different average values of b_{ic}), the average causal effect in the population will vary as younger cohorts enter and older cohorts exit the labor market. Second, changes in the year-specific payoff to schooling (as captured by δ_t) will shift the average causal effect in a proportional manner for all individuals. To reiterate, intertemporal changes in the causal effect of education may be due to cohort-specific factors (variation in b_c) and to year-specific factors (variation in δ_t): In what follows, I will refer to the cohort-specific and the year-specific components of the average causal effect of education.

If the unobserved determinants of log earnings are correlated with schooling, for example as the result of optimizing behavior in the schooling decision of individuals, the observed relationship between log earnings and schooling will be confounded. As a result, inference based on observed relationships, like the conventional estimate of the return to schooling, will be invalid. In a cross-section of data, the sources of confounding are the conventional "ability" bias (due to a correlation

[17] For simplicity, other determinants of earning capacity like labor market experience are omitted for the moment.

between a_{ic} and S_{ic}), and a "comparative advantage" bias (due to a correlation between b_{ic} and S_{ic}). With repeated cross-sectional data, the sources of confounding are more numerous: in addition to the cross-sectional ability and comparative advantage biases, a rise in the return to unobserved ability Ψ_t will confound a rise in the year-specific causal return to education δ_t. Moreover, if the correlation between educational attainment and the heterogeneity components varies across cohort, the magnitude of the absolute ability and comparative advantage biases will change as younger cohorts enter and older cohorts leave the labor market. To formalize these ideas, suppose that the conditional expectations of a_{ic} and b_{ic} are linear in years of completed education:

$$E[a_{ic}|S_{ic}] = a_c + \lambda_1^c(S_{ic} - \overline{S}_c) \tag{2.1}$$

$$E[b_{ic}|S_{ic}] = b_c + \lambda_2^c(S_{ic} - \overline{S}_c) \tag{2.2}$$

where a_c and b_c are cohort-specific averages of a_{ic} and b_{ic}. If individuals with higher return to schooling respond to the incentives of comparative advantage and acquire more schooling, the slope λ_2^c should be positive. The ability slope λ_1^c can be positive or negative depending on the model describing the optimizing behavior of individuals.[18]

In the model described by (1) and (2), the conventional measure of the return to education for cohort c observed in year t corresponds to:[19]

$$\delta_t b_c + \Psi_t \lambda_1^c + \delta_t \lambda_2^c \overline{S}_c \tag{3}$$

Thus, for a given cohort at a given point in time, the conventional estimate of the average return to education equals the sum of three components: $\delta_t b_c$, the average causal effect of education for cohort c in period t, $\Psi_t \lambda_1^c$ a time-varying ability bias specific to cohort c and $\delta_t \lambda_2^c \overline{S}_c$, a time-varying comparative advantage bias specific to cohort c. In the population as a whole, the magnitude of the absolute ability bias in period t will depend both on the amount of differential ability sorting across cohorts (captured by variation in λ_1^c) and on the return to unobservable ability in period t

[18] See Griliches (1977) for a discussion of this point. The evidence of non-hiearchical sorting in Willis and Rosen (1979) and Garen (1984) suggest that $\lambda_1^c < 0$ and $\lambda_2^c > 0$.

[19] This result was obtained under the assumption that the third central moment of S_{ic} is 0. Otherwise, the additional term $\lambda_2^c \frac{E[(S_{ic} - \overline{S}_c)^3]}{V(S_{ic})}$ must be included. See Deschenes (2001) for a derivation.

(Ψ_t).[20] Similarly, the magnitude of the comparative advantage bias in the population will depend on the amount of differential comparative advantage sorting across cohorts (captured by variation in $\lambda_2^c \overline{S}_c$), and on the return to education specific to period t (δ_t).

Equation (3) highlights the econometric difficulties associated with estimating the changing causal effect of education on earnings with observational data: changes in the conventional estimate of the average return to education may be driven by changes in λ_1^c, λ_2^c and Ψ_t as opposed to changes in b_c and δ_t.

B. *Identification of the changing causal effect of education.*

In absence of an exogenous source of variation in schooling levels, identification of the causal effect of education requires a statistical model to account for the biases created by the non-random allocation of schooling. In this paper, I propose a method using repeated cross-sectional data sets—without valid instrumental variables for schooling—to identify <u>changes</u> in the causal effect of education over time. The basic idea is to relate multiple observations on the estimated coefficients from an "augmented" earnings regression to the parameters of the causal model for log earnings and schooling. To proceed, note that equations (2.1) and (2.2) can be written as:

$$a_{ic} = a_c + \lambda_1^c (S_{ic} - \overline{S}_c) + u_{1ic} \tag{2.1'}$$

$$b_{ic} = b_c + \lambda_2^c (S_{ic} - \overline{S}_c) + u_{2ic} \tag{2.2'}$$

where u_{1i} and u_{2i} satisfy $E[u_{1ic}|S_{ic}] = E[u_{2ic}|S_{ic}] = 0$. Substitution of equations (2.1') and (2.2') into the earnings function (1) leads to:

$$\log y_{ict} = \Psi_t \{a_c + \lambda_1^c (S_{ic} - \overline{S}_c)\} + \delta_t \{b_c + \lambda_2^c (S_{ic} - \overline{S}_c)\} S_{ic} + v_{ict} \tag{4}$$

where $v_{ict} = \varepsilon_{ict} + \Psi_t u_{1ic} + \delta_t u_{2ic} S_{ic}$ is an heteroskedastic error term. Therefore under the maintained assumption that the conditional means of a_{ic} and b_{ic} are linear functions of years of education, and that the individual-specific relationship between log earnings and schooling is linear[21],

[20] Even with stationary returns to unobserved ability (Ψ_t=1 in all periods), an increase in the ability bias component $\Psi_t \lambda_1^c$ can arise if absolute ability sorting is more important for younger cohorts (i.e. if λ_1^c is rising across cohorts).

[21] The validity of this assumption will be evaluated in Section 6.

the regression function relating log earnings to years of education is quadratic:

$$E[\log y_{ict}|S_{ic}] = \Psi_t(a_c - \lambda_1^c \overline{S}_c) + \{\delta_t b_c + \Psi_t \lambda_1^c - \delta_t \lambda_2^c \overline{S}_c\} S_{ic} + \{\delta_t \lambda_2^c\} S_{ic}^2$$

or after collecting terms:

$$E[\log y_{ict}|S_{ic}] = \pi_0(c,t) + \pi_1(c,t) S_{ic} + \pi_2(c,t) S_{ic}^2 \tag{5}$$

$$\pi_0(c,t) \equiv \Psi_t(a_c - \lambda_1^c \overline{S}_c) \tag{5.1}$$

$$\pi_1(c,t) \equiv \delta_t b_c + \Psi_t \lambda_1^c - \delta_t \lambda_2^c \overline{S}_c \tag{5.2}$$

$$\pi_2(c,t) \equiv \delta_t \lambda_2^c \tag{5.3}$$

In a series of related papers, Garen (1984), Heckman and Vytlacil (1998), Card (1999,2000) and Woolridge (2000) have shown that with a valid instrumental variable, consistent estimates of all the causal parameters can be obtained by augmenting the standard human capital earnings function with a nonlinear function of schooling and the instrument. The additional regressors (i.e. the "control function") capture the correlation between schooling and the unobserved determinants of log earnings, thus purging the relationship between schooling and earnings of any ability and comparative advantage bias.[22]

In absence of a valid instrument for years of education this solution is not feasible. Nevertheless, equation (5) suggests an estimation method that is akin to the control function approach. Using repeated cross-sectional data sets, a series of quadratic earnings regressions like in (5) will be fitted to obtain estimates of $\pi_0(c,t)$, $\pi_1(c,t)$ and $\pi_2(c,t)$ for each cohort-year pair. Variation in $\pi_2(c,t)$ over time and across cohorts identifies the level λ_2^c for each cohort and δ_t, up to a normalization. Similarly, variation in $\pi_0(c,t)$ over time and across cohorts identifies Ψ_t (up to a normalization). Finally, differences in the levels of $\pi_0(c,t)$ and $\pi_1(c,t)$ across cohorts at a given point in time identify b_c and λ_1^c up to a single normalization. The normalizations are discussed with more detail in Section 4b.

The model above provides a simple, yet fairly general framework within which to interpret the data. Two important limitations warrant further discussion. First, the underlying earnings

[22]See Heckman and Robb (1985) for a general discussion of control function estimators in treatment effect models with non-random assignment.

function in (1) is linear in the endogenous schooling variable (as in Ashenfelter and Rouse (1998), and Heckman and Vytlacil (1998)). This is consistent with an optimizing model where the marginal return to schooling does not vary with schooling levels. In general, however, if there are diminishing returns to schooling in the production of human capital, the underlying earnings function should be concave. The assumption of linearity is maintened here because of practical reasons, moreover, recent evidence indicates that the marginal return to schooling is essentially invariant to schooling levels (Card 1999).[23] The second limitation relates to the cohort-specific selection parameters λ_1^c and λ_2^c. The assumption here is that these parameters are invariant once a cohort makes a permanent transition to the labor market. This amounts to assuming no ability-related attrition within-cohort, after all its members made a permanent transition to the labor market. This assumption would fail if within a cohort, less able workers permenantly exit the labor market earlier. Since this study focuses on a sample of prime-aged males, these effects are probably of second order.

4. Estimation.

In this section, I outline the two-step approach used to estimate the year-specific and cohort-specific parameters of the causal model. In the first step, estimates of the regression intercepts, linear and quadratic schooling coefficients in equation (5) are obtained by fitting a series of human capital earnings regressions, augmented with a quadratic schooling term. In the second step, the parameters of the causal model are estimated from the regression coefficients using a minimum-distance approach.[24]

A. *Estimation of the augmented earnings regressions.*

First, the following series of log earnings regressions is estimated from the repeated cross-sectional data generated by the ORG between 1979 and 2000:

$$\log y_{ict} = \pi_0(c,t) + \pi_1(c,t)S_{ict} + \pi_2(c,t)S_{ict}^2 + g(X_{ict}, \beta_t) + e_{ict} \tag{6}$$

[23]If the underlying earnings function is concave, a negative curvature parameter would load onto the quadratic schooling coefficient in (5.3). By assuming that this parameter is 0, we might overstate the role of comparative advantage selection.

[24]An alternative approach is to directly estimate the parameters with a nonlinear least square regression. The NLLS estimates are essentially identical to the OMD estimates reported in this paper. See Deschenes (2001).

In these regressions $\log y_{ict}$ is the log real weekly earnings (in 1999 constant dollars) of individual i, belonging to birth cohort c, and observed in survey year t; S_{ict} denotes years of completed education; $\pi_0(c,t)$, $\pi_1(c,t)$ and $\pi_2(c,t)$ are unrestricted regression coefficients for each cohort and year pair. Additional control variables are included in $g(X_{ict}, \beta_t)$, which contains a quartic in potential labor market experience, a race indicator, and interactions between race and education, with coefficients allowed to vary by year. Each regressor is deviated from its year-specific average. For the 127 distinct cohort-year pairs observed in the data, the regressions yield a total of 381 estimated coefficients, $\widehat{\pi}_0(c,t)$, $\widehat{\pi}_1(c,t)$ and $\widehat{\pi}_2(c,t)$ which will be used in the second step.[25]

B. *Estimation of the parameters via minimum-distance.*

For a given cohort-year pair, the relationship between the actual and the predicted coefficients of the causal model for earnings and schooling can be written as $\widehat{\pi}(c,t) = f(\theta_{ct}) + \eta_{ct}$, where:

$$\widehat{\pi}(c,t) = \begin{bmatrix} \widehat{\pi}_0(c,t) \\ \widehat{\pi}_1(c,t) \\ \widehat{\pi}_2(c,t) \end{bmatrix} \qquad \text{and} \qquad f(\theta_{ct}) = \begin{bmatrix} \Psi_t(a_c - \lambda_1^c \overline{S}_c) \\ \delta_t b_c + \Psi_t \lambda_1^c - \delta_t \lambda_2^c \overline{S}_c \\ \delta_t \lambda_2^c \end{bmatrix}$$

and where η_{ct} is a combination of sampling and specification errors. The optimal minimum distance (OMD) estimator of the parameters of interest $\theta_{ct} = [b_c, \lambda_1^c, \lambda_2^c, \delta_t, \Psi_t]'$ is obtained by minimizing the following quadratic form:

$$\frac{1}{T}\frac{1}{C}\sum_{c=1}^{C}\sum_{t=1}^{T}[\widehat{\pi}(c,t) - f(\theta_{ct})]' V_{ct}^{-1}[\widehat{\pi}(c,t) - f(\theta_{ct})]$$

where T is the number of survey years between 1979 and 2000, C is the number of birth cohorts, and V_{ct} is the asymptotic covariance matrix of $\widehat{\pi}_0(c,t)$, $\widehat{\pi}_1(c,t)$ and $\widehat{\pi}_2(c,t)$. Chamberlain (1984) shows that under certain regularity conditions, this minimum distance estimator is asymptotically efficient. In practice replacing V_{ct} by a consistent estimator does not change the asymptotic distribution of the OMD estimates. The heteroskedasticity-consistent estimator of White (1981) is used in the empirical analysis.

[25] Given the age restrictions imposed on the sample, not all birth cohorts are observed in all survey years. This explains the number of cohort-year pairs observed in the data. See Table 1 for more details.

C. *Additional restrictions for identification.*

The model for log earnings presented in Section 3 allows for two sources of heterogeneity among otherwise identical individuals (i.e. individuals belonging to the same birth cohort, having the same level of schooling, and observed during the same survey year). Changes over time in the contribution of these heterogeneity components to the log earnings of a given individual are captured by the factor loadings δ_t and Ψ_t. As a consequence, the factor loadings are only identified up to a normalization. In the empirical application, δ_t and Ψ_t are normalized to take a value of 1 in 1979. Given that, the causal return to education specific to year t, δ_t, will be identified from the ratio of $\pi_2(c,t)$ to $\pi_2(c,1979)$, while the return to unobserved ability specific to year t, Ψ_t, will be identified from the ratio of $\pi_0(c,t)$ to $\pi_0(c,1979)$.[26]

For a given cohort-year pair, the model has 3 implications for the level of the regression coefficients $\pi_0(c,t)$, $\pi_1(c,t)$ and $\pi_2(c,t)$, as illustrated in (5). Abstracting from the year-specific parameters δ_t and Ψ_t, the number of cohort-specific parameters is $4C$ (that is, a_c, b_c, λ_1^c and λ_2^c), where C is the number of birth cohorts. Clearly, not all of these parameters can be freely identified from the $3C$ regression coefficients. With the scale normalization on δ_t, the parameter corresponding to cohort-specific comparative advantage selection λ_2^c is directly identified off the quadratic schooling coefficient in (5.3). However, there is a tradeoff concerning which of the remaining parameters can be identified. Provided that an estimate of λ_1^c is available, the average marginal productivity of schooling for cohort c, b_c, will be identified from the linear schooling coefficient in (5.2). Therefore, some restrictions must be imposed on a_c and λ_1^c. As equation (5.1) makes clear, imposing less restrictions on the shape of λ_1^c across cohorts comes at the cost of imposing more restrictions on a_c. The approach I follow is guided by the objective of this study: assessing intertemporal changes in the causal effect of education by decomposing the observed relationship between log earnings and schooling into time-varying average causal effect, time-varying ability bias and time-varying comparative advantage bias. One normalization consistent with this objective is to let the ability slope parameter λ_1^c take a fixed value for <u>one</u> of the eight birth cohorts (henceforth the reference cohort), and assume no differences across cohorts in the average level of absolute ability (i.e. that $a_c{=}a$ for all cohorts).[27] This lack of identification is a drawback of using

[26]In practice the OMD estimator will "average" these ratios for all cohorts observed in year t. Cohorts for which the ratio is more precisely estimated will receive more weight in the calculations.

[27]An alternative possibility is to assume that the ability bias parameters are equal for 2 different cohorts. In that

observational data without valid instruments for years of education: with a source of exogenous variation in schooling, the levels of all the parameters of the model could be estimated. This study, however, focuses on intertemporal changes in the causal effect of education, and differences across cohorts in b_c, λ_1^c and λ_2^c, combined with the relative returns δ_t and Ψ_t are sufficient to make such an assessment credibly.

Information from other studies can be used to calibrate the range of potential values for λ_1 for the reference cohort. The evidence reviewed in Griliches (1977) and Card (1999) is suggestive of small absolute ability biases. Consequently, in the empirical implementation of the model described in equation (1)-(5), the correlation between ability and schooling for the reference cohort (the cohort of men born between 1945 and 1949) will be normalized to 0. The robustness of the findings to positive or negative normalizations on λ_1 for the reference cohort will be evaluated in Section 6.

5. Empirical Results.

A. *Results from the augmented earnings regressions.*

Tables 2, 3 and 4 presents the OLS estimates of the augmented earnings regressions following the specification of equation (6). The estimated regression intercept, linear and quadratic schooling coefficients are reported, as well as the corresponding standard errors.[28] In order to keep the size of the tables manageable, averages of the estimated coefficients for 3 year bands are reported. The full set of estimated regression coefficients have already been reported in Figure 5. The trends illustrated in Figure 5 are also apparent in these tables, moreover, they indicate that these regression coefficients are precisely estimated, with standard errors in the range of 0.25 to 0.75 for the linear term and in the range of 0.01 to 0.03 for the quadratic term. The next section will make use of this information to estimate the parameters of the causal model.

B. *Minimum-distance estimates of the two-factor model.*

The model presented in equation (1)-(5) encompasses a wide variety of models that have been considered in the literature to study changes in the wage structure and make causal inference on the effect of education on earnings. Since it encompasses all of the previous analyses, the two-factor

case, a model with a linear trend across cohorts in average ability (i.e. $a_c = a_0 + a_1 c$) could be estimated. This is not a major concern since a_c does not load into the schooling coefficients of the earnings function.

[28]For convenience, the estimated schooling coefficients and their standard errors are multiplied by 100.

model with unobserved ability, heterogeneity in the return to education and separate year-specific return to unobserved ability and education is considered first. Table 5 report the OMD estimates of the cohort-specific parameters b_c, λ_1^c and λ_2^c, as well as the goodness-of-fit statistic for the model. Table 6 reports the OMD estimates of the year-specific returns to education and unobserved ability (δ_t and Ψ_t) for selected years between 1979 and 2000. Figure 6 displays the complete set of estimates of δ_t and Ψ_t.

The first column contains the cohort-specific average marginal productivity of schooling, b_c. The parameters indicate that an additional year of education permanently increase log earnings in the range of 5.0% and 5.5% for the different cohorts, with the smallest values for the cohorts born after 1950. All the parameters are precisely estimated with standard errors of about 0.003. If anything, the entries in column 1 indicate a downward trend in b_c across cohorts, and therefore that any increase in the causal effect of education would have to be driven by a rise in the year-specific causal return to schooling δ_t.

The second column shows the level of ability bias specific to each cohort (λ_1^c), relative to the reference cohort. The estimated parameters are typically not significantly different from 0. More interestingly, the estimates indicate no inter-cohort trend in the correlation between unobserved ability and completed education. While the level of the estimated average education-slope and ability bias depends on the normalization on λ_1 for the reference cohort (men born between 1945 and 1949), the difference in the parameters across cohorts does not. The entries in column 2 provide no evidence suggesting that the rise in the conventional measure of the return to education is due to an increasingly higher correlation between absolute ability and schooling among younger cohorts.

Column 3 reports the estimates of λ_2^c, the slope in the relationship between person-specific returns to education and years of completed education specific to each cohort, multiplied by $\overline{S}_c$, the average schooling for cohort c. As shown in (3), this amounts to the contribution of comparative advantage bias in the conventional measure of the return to schooling. All the entries are positive and statistically significant, providing clear evidence of heterogeneity in the returns to schooling, and of the existence of a positive correlation between educational attainment and returns to education at the individual level. This is consistent with the results in Willis and Rosen (1979), Garen (1984), and Meghir and Palme (1999) who found that individuals pursued comparative advantage in their educational investments. As Table 5 suggests, the pursuit of comparative advantage played an important role in educational investments for the men born between 1930 and 1970 in the United States, moreover, the extent of educational self-selection based on individual returns to schooling appears to be more important for younger cohorts.[29]

Finally, column 4 displays the averages of the conventional measure of the return to education taken over their entire working lifetime of each cohort (i.e. for each cohort over all the years it is observed in the data).[30] By examining the entries across the rows of column 4, it appears that the "total effect" of education on earnings (i.e. including the bias terms) is similar across cohorts when averaged over their entire working histories (the maximal difference is 0.01). The contribution of the elements in columns 1-3 to these lifetime averages, however, differs substantially across cohorts. For the older cohorts, individual differences in the return to education correlated with schooling (as captured by $\lambda_2^c \overline{S}_c$) explain about 25% of the average total lifetime effect. For the younger cohorts born after 1950, these individual differences explain about 40% of the average lifetime total effect.

While the figures in Table 5 indicate no differences across cohorts in the extent of absolute ability sorting, it is possible that a rise in the year-specific return to unobserved ability could be the driving force behind the observed increase in the conventional measure of the return to education over the last twenty years. This possibility is investigated in Table 6 which contains the estimated relative returns to education and unobserved ability (δ_t and Ψ_t), for selected years. The estimated factor loading on years of completed education (δ_t) represents the causal return to education that is common to all labor market participants in year t (i.e. net of the cohort-specific parameters reported in Table 5), and thus is not confounded by the return to unobserved ability in year t (Ψ_t) or by the composition effects. As Table 6 indicates, the year-specific causal return to education and the return to unobserved ability followed different patterns over the 1980s and the 1990s. Whereas the return to education increased by about 30% more or less continuously over the whole period, the return to unobserved ability increased by about 10% over the 1980s,[31]

[29]The increasing trend in $\lambda_2^c \overline{S}_c$ across cohort is solely due to larger correlation between educational attainment and person-specific return to education for younger cohorts (larger λ_2^c) and not to an increase in average educational attainment across cohorts ($\overline{S}_c$). As Figure 2 showed, average educational attainment actually decreased for cohorts born after 1950.

[30]For the younger cohorts who are still in their prime working age, this might be an imperfect indicator of their lifetime average. Note columns 1 to 3 do not exactly add up to column 4 because the year-specific factor loadings δ_t and Ψ_t are not included in the entries of columns 1 to 3 but are implicitly included in column 4.

[31]This is consistent with the evidence in Chay and Lee (1999) who report that the return to unobserved ability increased by 10-20% over the 1980s.

but actually declined over the 1990s. These trends are illustrated in Figure 6, which displays the estimated δ_t and Ψ_t for each year between 1979 and 2000.

C. *Censoring in the CPS data.*

All measures of earnings are censored in the Current Population Survey, and in the ORG files usual weekly earnings are top-coded.[32] The summary statistics in Deschenes (2001) indicate that the censoring problem was most important between 1985 and 1987, and that the probability of having censored weekly earnings increases with educational attainment. Since log earnings and schooling are positively correlated, censoring of the earnings variable will likely result in underestimation of the effect of schooling on earnings. Indeed, the evidence in Chamberlain (1994) suggests that censoring of weekly earnings in ORG data leads to systematic underestimation of the returns to schooling. To account for the censoring problem, the regression coefficients in (6) were also estimated using a Tobit model. It is well known, however, that Tobit estimates of a censored regression model are not consistent when the distribution of the error terms is non-normal or heteroskedastic.[33] Since the causal model for log earnings in (1)-(5) implies heteroskedastic residuals, the regression coefficients were also estimated using Powell's (1984) censored least absolute deviation (CLAD) semiparametric estimator. This estimator is robust to deviations from normality and to heteroskedasticity.

Tables 7-10 presents the OMD estimates of the parameters of the causal model for earnings derived from the Tobit and CLAD estimates of equation (6). Examination of the entries in Tables 7 and 9 reveals that the estimated parameters are very similar to those contained in Table 5. On the basis of this simple comparison, there is no indication that the estimates of the cohort-specific parameters b_c, λ_1^c and λ_2^c reported in Table 5 are seriously biased by censoring. Tables 8 and 10 show the point estimates of the year-specific returns to education and unobserved ability (again for selected years) when OMD is based on Tobit and CLAD first-stage estimates. The full set of estimated year-specific parameters is displayed in Figures 7 and 8. As Figures 7 and 8 indicate, the censoring-corrected estimates of the year-specific causal return to education δ_t are slightly larger,

[32]From 1979 to 1988, any value of usual weekly earnings in excess of \$999 was recorded as \$999, similarly, from 1989 to 1997 weekly earnings were censored at \$1923, and from 1998 to 2000, weekly earnings were censored at \$2884.

[33]See for example Goldberger (1983). Chay and Honore (1998) estimate the bias due to nonnormality and heteroskedasticity in the context of a racial earnings gap when earnings are severely top-coded. They show that the racial wage gap can be biased downwards by 10 to 30%.

with the largest gap between 1985-1988, the period when the fraction of censored earnings in the ORG data was the highest. Interestingly, the top coding of weekly earnings has no impact of the estimated return to unobserved ability Ψ_t. If anything, the censoring-corrected estimates are slightly smaller.

D. *Minimum-distance estimates of alternative models.*

Table 11 displays the goodness-of-fit statistics associated with alternative models that have been considered in the literature studying the changes in the wage structure and attempting to make causal inference on the effect of education on earnings. In column 1 are the OMD goodness-of-fit statistics obtained when the regression coefficients in (6) are computed by OLS, while columns 2 and 3 correspond to those obtained by Tobit and CLAD. For convenience, the goodness-of-fit statistic associated with the two-factor model presented in Tables 5-10 is reported in row 1. The other rows of Table 11 test further simplifications against the two-factor model alternative.[34] The second row displays the statistical fit of a model where the average marginal productivity of schooling is constrained to be same for all the cohorts. The equality of the b_c across cohorts is rejected by the unconstrained model in row 1. The third row shows the fit of a model imposing no variation across cohorts in the correlation between educational attainment and return education at the individual level. Not surprisingly, this restriction is easily rejected by the unconstrained alternative. The fourth and fifth rows present the estimation results when stationarity restrictions are imposed. Row four shows that the hypothesis of a constant year-specific causal return to education between 1979 and 2000 (while allowing the return to ability Ψ_t to vary) is overwhelmingly rejected. Similarly, in row five, the stationarity of the return to unobserved ability is also rejected. In the sixth row, the goodness-of-fit statistic associated with a model estimated under the restriction of no comparative advantage selection (i.e. that λ_2^c=0) is reported. Variants of this model have been used in most cross-sectional analyses of the causal effect of education (see e.g. Angrist and Krueger (1991), Ashenfelter and Rouse, 1998). The large chi-square statistic suggests that this restriction is not supported by the data considered in this study. Finally, the last row displays the statistical fit of a "one-factor" model of ability in which the return to education and the return to unobserved ability

[34]The difference between the goodness-of-fit statistics associated with the additional restrictions and the "unrestricted" alternative (the two-factor model in our case) has an asymptotic χ^2 with degrees of freedom equal to the number of additional restriction being tested. See Chamberlain (1984).

are restricted to be equal in each time period (i.e. that $\delta_t = \Psi_t$).[35] This restriction is soundly rejected by the data, as indicated by the large chi-square statistic.

To summarize, OMD estimates based on OLS, Tobit and CLAD all clearly indicate that any further restriction imposed on the two-factor model is not consistent with the data. In particular, models assuming homogenous returns to schooling and stationary causal return to education appear to be greatly mis-specified.

6. Alternative Specifications and Sensitivity Analysis.

This section investigates the robustness of the findings in Tables 5-11 to alternatives specifications or assumptions. Particular attention is devoted to: (A) alternative normalizations on the magnitude of the ability bias for the reference cohort; (B) measurement error in reported schooling; (C) small-sample bias in OMD estimation; and (D) linearity of the underlying earnings function.

A. *Alternative restrictions for the reference cohort.*

To evaluate the robustness of the results reported in Table 5, the two-factor model was estimated under alternative normalizations on λ_1 for the reference cohort. Figure 9 displays the OMD estimates of b_c, λ_1^c, λ_2^c assuming negative (-0.02), zero (like in Table 5), and positive (0.02) ability bias for the reference cohort. As panels A and B of Figure 9 indicate, differences across cohorts in b_c and λ_1^c are preserved under the different normalizations. Panel C clearly shows that the cohort-specific levels of comparative advantage selection bias are invariant to the normalization on λ_1. Therefore, the evidence in Figure 9 suggests that the conclusions drawn from Table 5 are not driven by any particular restriction imposed on λ_1 for the cohort of men born between 1945 and 1949.

B. *Measurement error in reported schooling.*

An important issue in studies of the return to schooling is the possibility of measurement errors in schooling. In bivariate regressions, classical measurement error in the independent variable leads to a downward bias in the estimated coefficient. The magnitude of this attenuation bias depends on reliability of the observed variable, which for schooling has generally been found to be about 0.90.[36] In nonlinear models, however, the effects of measurement error are not as straightforward

[35] Card and Lemieux (1996) use a one-factor model to study changes in the wage structure over the 1980s.
[36] See for example Table 9 in Angrist and Krueger (1999).

to predict and often must be evaluated on a case-by-case basis.[37] Appendix A shows that in a regression of log earnings on schooling and schooling squared, the bias depends on the true values of the regression coefficients, and the on the characteristics of the joint distribution of reported and actual schooling. Since complete data on the joint distribution of true and reported schooling are not available, direct calculation of correction factors is not feasible. One solution is to use simulations to gain knowledge on the sign of the biases and evaluate the effect of measurement error bias on the OMD estimates. Schooling and log earnings data were simulated using the range of estimates of π_0, π_1 and π_2 reported in Tables 2 to 4 and assuming that the reliability of reported schooling is 0.90. The results are summarized in the top panel of Table A.1 and suggests that (i) the intercept is unaffected by measurement error in reported schooling; (ii) the linear schooling coefficient has a small (about 10%) upward or downward bias; and (iii) the quadratic schooling coefficient is always biased downwards by 25%.

Under the assumption that the extent of measurement error is the same across cohorts and constant over time, the simulations also provide approximate correction factors that can be used to adjust the regression coefficients before using them in the OMD estimation. This provides measurement error corrected OMD estimates of b_c, λ_1^c, λ_2^c, δ_t and Ψ_t. Table A.1 presents those corrected estimates for 4 different combinations of true values for π_0, π_1 and π_2, as well as the actual estimated parameters from Tables 5 and 6. As the entries indicate, only the year-specific return to education δ_t appears to be downward biased by measurement error in reported schooling. For the four cases considered, the corrected estimates of δ_t are always larger than the uncorrected ones, suggesting that the estimates of δ_t in Section 5 should be viewed as lower bounds. The rest of the evidence in Table A.1 contains no indication of serious measurement error bias in the estimates reported in Tables 5 to 11.

C. *Small sample bias in OMD estimation.*

Although minimum distance estimators are consistent and asymptotically normally distributed under relatively mild regularity conditions, in practice, the asymptotic distribution may provide a poor approximation to the finite sample distribution. In the context of OMD estimators, small sample bias may arise in the presence of a correlation between the sampling errors in the elements

[37]See Griliches and Ringstad (1970) and Hausman, Newey, Ichimura and Powell (1991) for discussions in different contexts.

of the weighting matrix and the sampling errors in the vector of sample moments. In particular the simulation evidence in Altonji and Segal (1996) and Clark (1996) indicates that in linear and nonlinear models of covariance structure, OMD estimates are biased downwards in absolute value. An alternative type of minimum-distance estimator can be used evaluate the importance of small sample bias in Tables 5 and 6. The Independently Weighted Optimal Minimum-Distance (IWOMD) was suggested by Altonji and Segal (1996) to circumvent the problem of a correlation between the sampling errors in π and $V(\pi)$. The basic idea of IWOMD is to randomly split the data into subgroups and use different subgroups to estimates the moments and the weighting matrices, thus eliminating the possibility of correlation between the sampling errors in π and $V(\pi)$.

The design of the CPS samples is particularly well suited for this type of sample splitting analysis. More specifically, the earnings data in the monthly ORG files are collected from individuals who are in their fourth or eight month in the CPS samples. Since individuals in the fourth and eight rotation subsamples are independently drawn from the same population, the two subsamples are independent of each other. Thus, IWOMD can be implemented using these two random partitions of the CPS samples, and the results are presented in Table A.2. For brevity only the results of the cohort-specific parameters from the two-factor model are reported here.[38] All the parameters in Table A.2 are very close in magnitude to those in Table 5 and none of the differences are statistically significant. The similarity between the entries in Tables A.4 and Table 5 suggest that those estimates are not subject to small sample bias.

D. *Linearity of the earnings function.*

Section 3 emphasized that the identification of the parameters of the two-factor model hinges greatly on linearity assumptions. The linearity of the conditional expectations of the unobserved determinants of log earnings, combined with the linearity of the underlying earnings function implies a quadratic regression relating schooling and log earnings in the population. The assumptions, however, are not readily testable since they involve unobservable components a_{ic} and b_{ic}. Therefore, it is important to assess the validity of the results presented in this paper when departures from linearity are considered. One possibility is to allow log earnings and schooling to be nonlinearly

[38] See Deschenes (2001) for the full set of IWOMD estimates.

related as follows:

$$\log y_{ict} = \Psi_t a_{ic} + \delta_t b_{ic} S_{ic}^{\theta} + \varepsilon_{ict} \tag{1"}$$

and where the conditional expectations of the unobserved ability factors are linear in a *transformation* of schooling:

$$E[a_{ic}|S_{ic}] = a_c + \lambda_1^c (S_{ic}^{\theta} - \overline{S_c^{\theta}}) \tag{2.1"}$$

$$E[b_{ic}|S_{ic}] = b_c + \lambda_2^c (S_{ic}^{\theta} - \overline{S_c^{\theta}}) \tag{2.2"}$$

The model presented in Section 3 is a special case of this generalized model when $\theta = 1$. Equations (1"), (2.1") and (2.2") implies that the population regression function relating log earnings and schooling is quadratic in the transformed schooling S_i^{θ}, and therefore the empirical approach described in Section 4 can be used, given a predetermined value for θ. Figure 10 presents the goodness-of-fit statistics associated with the OMD estimates of the parameters of the model (1")-(2.2") for values of θ ranging between 0.5 and 1.5. These statistics measure the adequacy of the two-factor model corresponding to different values of θ and provide a simple assessment of the robustness of the results presented in Tables 5-11 to departures from the underlying linearity assumptions. The statistical fit of the two-factor models reported in Figure 10 does not provide strong evidence against the linearity assumption. The smallest goodness-of-fit statistic reported is 339.4, corresponding to $\theta = 1.07$, while the goodness-of-fit statistic for the linear model is 391.0. Based on this exercise, it appears that the linearity of the underlying earnings function provides a reasonable approximation to more general function forms relating earnings and schooling.

7. Conclusion.

This study assesses the movements in the educational wage structure in the United States between 1979 and 2000. The main purpose is to quantify the change in the average causal effect of education using repeated cross-sectional data when instrumental variables for schooling are not available. A causal model for earnings, with absolute and comparative advantage in the schooling decision of individuals is used to impose some structure on the observed relationship between schooling and earnings. A simple intuition arises from the model: if individuals with higher returns to schooling acquire more schooling, the relationship between log earnings and schooling will be convex in the population. Likewise, for a cohort of individuals with a fixed distribution of ability after transition to the labor market, the degree of convexity of the relationship will rise if the year-specific causal return to education rises. This feature of the changing U.S. wage structure is well-documented in the first part of this paper.

Taken as a whole, the evidence in this paper suggests that the observed rise in the conventional measure of the return to education reflects a change in the causal effect of education on earnings. After parsing out the effects of differential ability and comparative advantage biases specific to each cohort, and allowing unobserved ability to have a differential rate of return, the estimates indicate that the causal return to education increased by about 30% between 1979 and 2000. By comparison, the conventional measure of the return to education increased by about 50%. The empirical background for this finding is the striking rise in the convexity of the earnings-schooling relationship that affected all cohorts uniformly over the 1980s and 1990s. The results also indicate that all of the increase in the average causal effect of education is due to year-specific factors as opposed to cohort-specific effects, for example the entry in the labor market of younger cohorts with higher average marginal productivity of schooling.

My estimates also imply that there are no differences across cohorts in the mapping between absolute ability and schooling, and virtually no changes in the return to unobserved ability between 1979 and 2000: The estimates show that the return to unobserved ability increased by 10% over the 1980s and decreased by 10% over the 1990s. Therefore, my results provide no indication that the rise in the conventional measure of the return to school is confounded by a time-varying ability bias.

Perhaps more importantly, my results indicate a significant increasing trend in comparative advantage sorting across cohorts. This is consistent with a model where returns to schooling vary across individuals, and where comparative advantage incentives play a more important role in the schooling decision of younger cohorts.

Finally, various specification tests suggest that the two-factor model developed here provides a reasonable characterization of the changes in the educational wage structure over the 1980s and 1990s. In particular, the specification tests have shown that models assuming homogenous returns to education in the population—as was used in previous analyses—are greatly mis-specified. A key element of future research will be to understand the sources of the growing association between educational attainment and returns at the individual level.

References.

Altonji, Joseph G. and Dunn (1996): "The effects of Family Background Characteristics on the Return to Schooling," *Review of Economics and Statistics* 78: 665-671.

Altonji, Joseph G. and Lewis M. Segal (1996): "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics* 14: 353-366.

Angrist, Joshua D. and Alan B. Krueger (1991): "Does Compulsory School Attendance Affect Schooling and Earnings," *Quarterly Journal of Economics* 106: 974-1014.

Angrist, Joshua D. and Guido W. Imbens (1995): "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of the American Statistical Association* 90: 431-442.

Angrist, Joshua D. and Alan B. Krueger (1999): "Empirical Strategies in Labor Economics," in Orley Ashenfelter and David Card, editors, *Handbook of Labor Economics*, volume 3A, North-Holland, Amsterdam and New York.

Ashenfelter, Orley and Cecilia E. Rouse. (1998): "Income, Schooling and Ability: Evidence from a New Sample of Twins," *Quarterly Journal of Economics*, 113: 253-284.

Becker, Gary S. (1975): *Human Capital* (2nd Edition), Chicago, University of Chicago Press.

Blackburn, McKinley and David Neumark (1993): "Omitted-Ability Bias and the Increase in the Returns to Schooling," *Journal of Labor Economics* 11: 521-544.

Card, David and Alan B. Krueger (1992): "Does School Quality Matter: Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy* 100: 1-40.

Card, David and Thomas Lemieux (1996): "Wage Dispersion, Returns to Skill and Black-White Wage Differentials," *Journal of Econometrics* 74: 319-361.

Card, David (1999): "The Causal Effect of Education on Earnings," in Orley Ashenfelter and David Card, editors, *Handbook of Labor Economics*, volume 3A, North-Holland, Amsterdam and New York.

Card, David (2000): "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," NBER Working Paper No. 7769.

Card, David and Thomas Lemieux (2001): "Dropout and Enrollment Trends in the Postwar Period: What Went Wrong in the 1970s?," in Jonathan Gruber (editor), *Risky Behavior among Youth: An Economic Analysis*, National Bureau of Economic Research.

Cawley, John, James Heckman, Lance Lochner and Edward Vytlacil (2000): "Understanding the Role of Cognitive Ability in Accounting for the Recent Rise in the Economic Return to Education," in Kenneth Arrow, Samuel Bowles and Steven Durlauf, editors, *Meritocracy and Economic Inequality*, Princeton University Press, Princeton NJ.

Cawley, John, James Heckman and Edward Vytlacil (1998): "Cognitive Ability and the Rising Returns to Education," NBER Working Paper No. 6388.

Chamberlain, Gary (1984): "Panel Data," in Zvi Griliches and Michael Intrilligator, editors, *The Handbook of Econometrics*, volume 2, North-Holland, Amsterdam and New York.

Chamberlain, Gary (1994): "Quantile Regression, Censoring and the Structure of Wages," in Christopher Sims editor, *Advances in Econometrics: Sixth World Congress*, Volume 2, Cambridge University Press.

Chay, Kenneth Y. and David S. Lee (1999): "Changes in relative Wages in the 1980s: Returns to Observed and Unobserved Skills and Black-White Wage Differentials," mimeo, University of California-Berkeley.

Chay, Kenneth, Y. and Bo E. Honore (1998): "Estimation of Semiparametric Censored Regression Model," *Journal of Human Ressources* 33: 4-38.

Clark (1996): "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics* 14: 353-366.

Deschenes, Olivier (2001): "Unobserved Ability, Comparative Advantage, and the Rising Return to Education in the United States: A Cohort-Based Approach," Industrial Relations Section Working Paper #456, Princeton University. Available online at: http://www.irs.princeton.edu.

Dustmann, Christian and Costas Meghir (2001): "Wages, Experience and Seniority," IFS Working

Paper W01/01.

Garen, John (1984): "The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable," *Econometrica* 52: 1199-1218.

Griliches, Zvi and Vidar Ringstad (1970): "Error-in-the-Variables in Nonlinear Contexts," *Econometrica* 38: 368-370.

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

Hausman, Jerry, A., Whitney K. Newey, Hidehiko Ichimura, and James L. Powell (1991): "Identification and Estimation of Polynomial Errors-in-Variables Models," *Journal of Econometrics* 68: 273-295.

Heckman, James J. and Richard Robb (1986): "Alternative Methods for Evaluating the Impact of Interventions," in James J. Heckman and Burton Singer, editors, *Longitudinal Analysis of Labor Market Data*, Cambridge University Press.

Heckman, James J., Anne Layne-Farrar and Petra Todd (1996): "Does Measured School Quality Really Matter? An Examination of the Earnings-Quality Relationship," in Gary Burtless editor, *Does Money Matter? The Effects of School Resources on Student Achievement and Adult Success,* Brookings Institution, Washington, DC.

Heckman, James J. and Edward Vytlacil (1998): "Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Rate of Return to Schooling When the Return is Correlated with Schooling," *Journal of Human Resources* 23: 974-987.

Heckman, James J. and Edward Vytlacil (2000): "Identifying the Role of Cognitive Ability in Explaining the Level of and Change in the Return to Schooling," NBER Working Paper No. 7820.

Heckman, James J., Justin L. Tobias and Edward Vytlacil (2000): "Simple Estimators for Treatment Parameters in a Latent Variable Framework with an Application to Estimating the Returns to Schooling," NBER Working Paper No. 7950.

Jaeger, David A. (1997): "Reconciling the Old and New Census Bureau Education Questions: Recommendations for Researchers," *Journal of Business and Economic Statistics* 15: 300-308.

Juhn, Chinhui, Kevin M. Murphy and Brooks Pierce (1993): "Wage Inequality and the Rise in the Returns to Skill," *Journal of Political Economy* 101: 410-42.

Katz, Lawrence F. and David Autor (1999): "Changes in the Wage Structure and Earnings Inequality," in Orley Ashenfelter and David Card, editors, *Handbook of Labor Economics*, volume 3A, North-Holland, Amsterdam and New York.

Meghir, Costas and Marten Palme (1999): "Assessing the Effect of Schooling on Earnings Using a Social Experiment," Working Paper, University College London.

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

Mincer, Jacob (1996): "Changes in Wage Inequality 1970-1990," NBER Working Paper No. 5823.

Murnane, Richard, J., John B. Willett and Frank Levy (1995): "The Growing Importance of Cognitive Skills in Wage Determination," *Review of Economics and Statistics*, 77: 251-266.

Powell, James L. (1984): "Least Absolute Deviations Estimation for the Censored Regression Models," *Journal of Econometrics* 25: 303-325.

Rosen, Sherwin (1977): "Human Capital: A Survey of Empirical Research," in Ronald Ehrenberg, editor, Research in Labor Economics, volume 1, Greenwich Connecticut: JAI Press.

Taber, Christopher (1999): "The Rising College Premium in the Eighties: Return to College or Return to Unobserved Ability?" mimeo, Northwestern University.

White, Halbert (1980): "A Heteroskedastic-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica* 48: 817-838.

Willis, Robert (1986): "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Function," in Orley Ashenfelter and Richard Layard, editors, *Handbook of Labor Economics*, North-Holland, Amsterdam and New York.

Willis, Robert and Sherwin Rosen (1979): "Education and Self-Selection," *Journal of Political Economy* 79: S7-S36.

Woolridge, Jeffrey M. (2000): "Instrumental Variables Estimation of the Average Treatment Effect in the Correlated Random Coefficient Model," mimeo, Department of Economics, Michigan State University.

Appendix A. Measurement Error in Quadratic Regressions.

Suppose that schooling is measured with classical measurement error:

$$S_i^o = S_i + v_i$$

This implies that schooling squared is also measured with error:

$$S_i^{o2} = S_i^2 + 2S_i v_i + v_i^2$$

Note that in general there will be a correlation between the errors of measurement in the two regressors. Denote the true regression relating y to S_i and S_i^2 and the regression of y on the observed regressors respectively as:

$$\begin{aligned} y &= \alpha + \beta S_i + \gamma S_i^2 + \varepsilon_i \\ y &= \alpha + \beta S_i^o + \gamma S_i^{o2} + \varepsilon_i - v_i \beta - \gamma v_i^2 - 2\gamma S_i v_i \end{aligned}$$

The bias in the OLS estimates of β and γ is easily obtained by considering the linear projections of v_i, v_i^2 and $v_i S_i$ onto S_i^o and S_i^{o2}:

$$\begin{aligned} v_i &= \lambda_{11} S_i^o + \lambda_{12} S_i^{o2} + \xi_{1i} \\ v_i^2 &= \lambda_{21} S_i^o + \lambda_{22} S_i^{o2} + \xi_{2i} \\ v_i S_i &= \lambda_{31} S_i^o + \lambda_{32} S_i^{o2} + \xi_{3i} \end{aligned}$$

The bias in the OLS estimates is:

$$\begin{aligned} \widehat{\beta} - \beta &= -\beta \lambda_{11} - \gamma(\lambda_{21} - 2\lambda_{31}) \\ \widehat{\gamma} - \gamma &= -\beta \lambda_{12} - \gamma(\lambda_{22} - 2\lambda_{32}) \end{aligned}$$

Therefore the bias in the OLS estimates depends on the true values of β and γ and some characteristics of the joint distribution of S_i and S_i^o as captured by (λ_{11}, λ_{12}, λ_{21}, λ_{22}, λ_{13} and λ_{32}).

Figure 1: Average Educational Attainment among Full-time Male Workers, by Year of Birth.

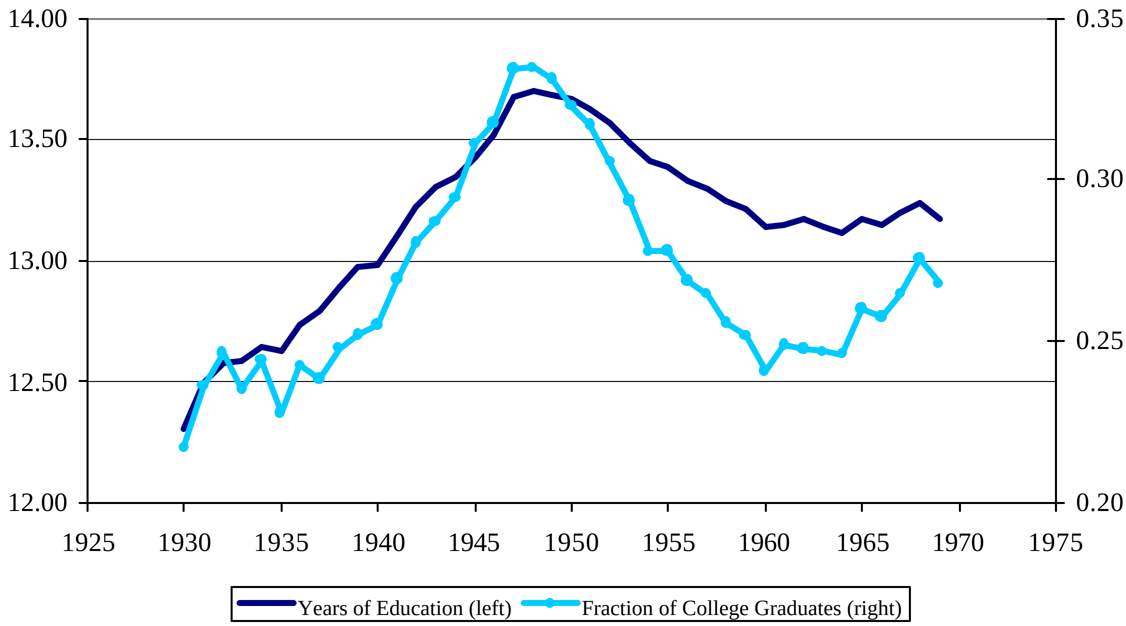


Figure 2: Average Educational Attainment among Full-time Male Workers, by Calendar Year 1979-2000.

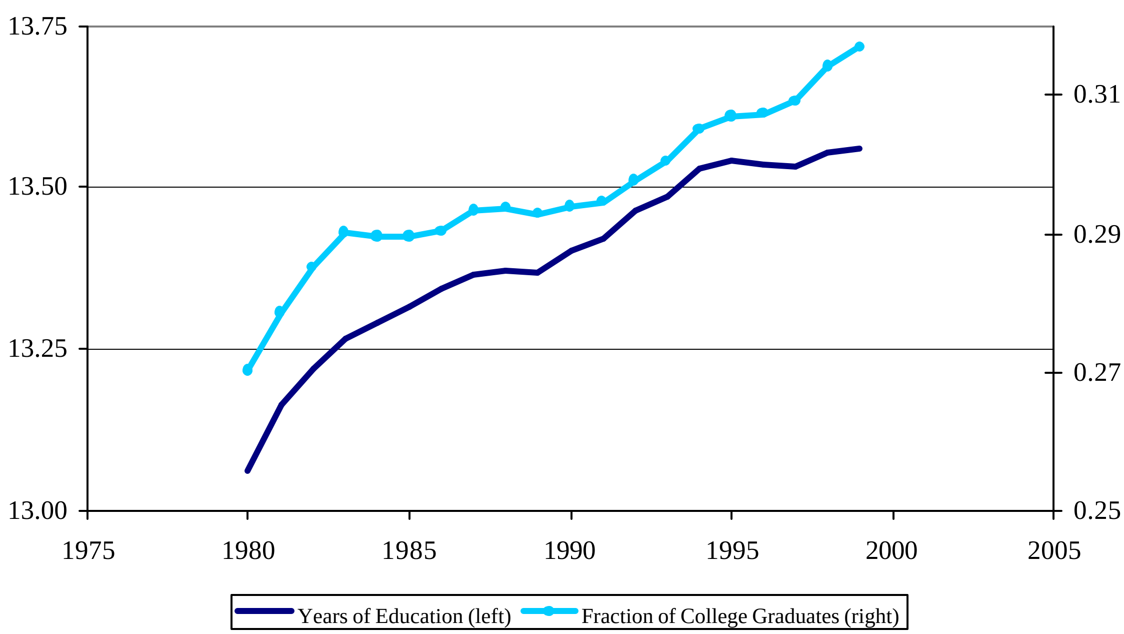


Figure 3: Adjusted Mean Log Earnings by Years of Education, 1979-2000.
[Relative to the mean log earnings of high school graduates]

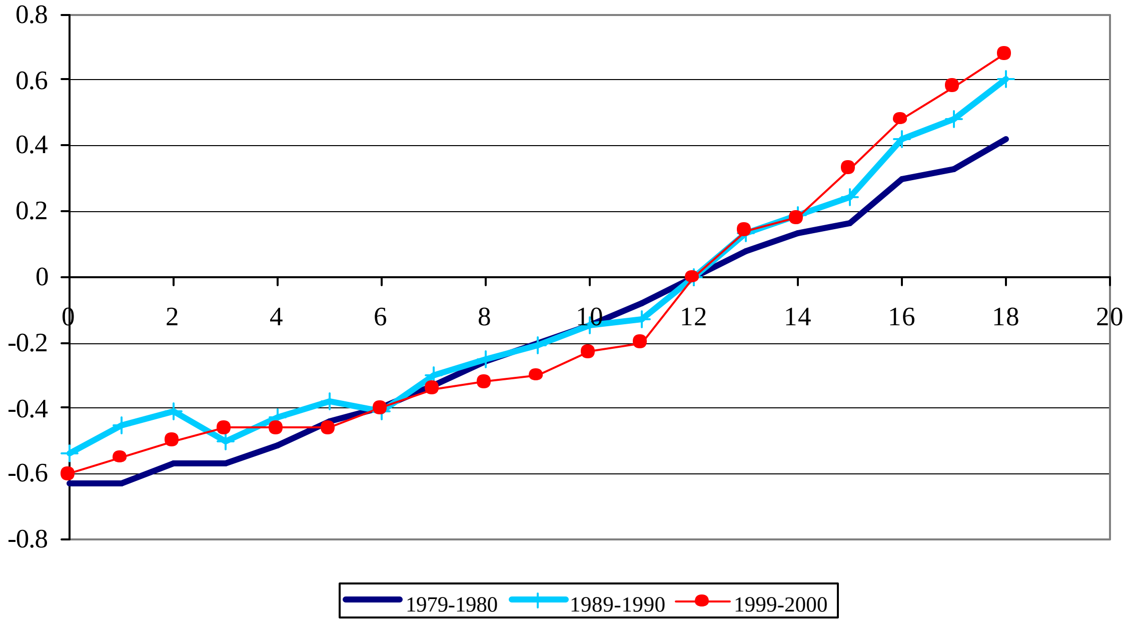


Figure 4: Changes in the Earnings-Schooling Relationship and the Rising Return to Education, 1970-1999.

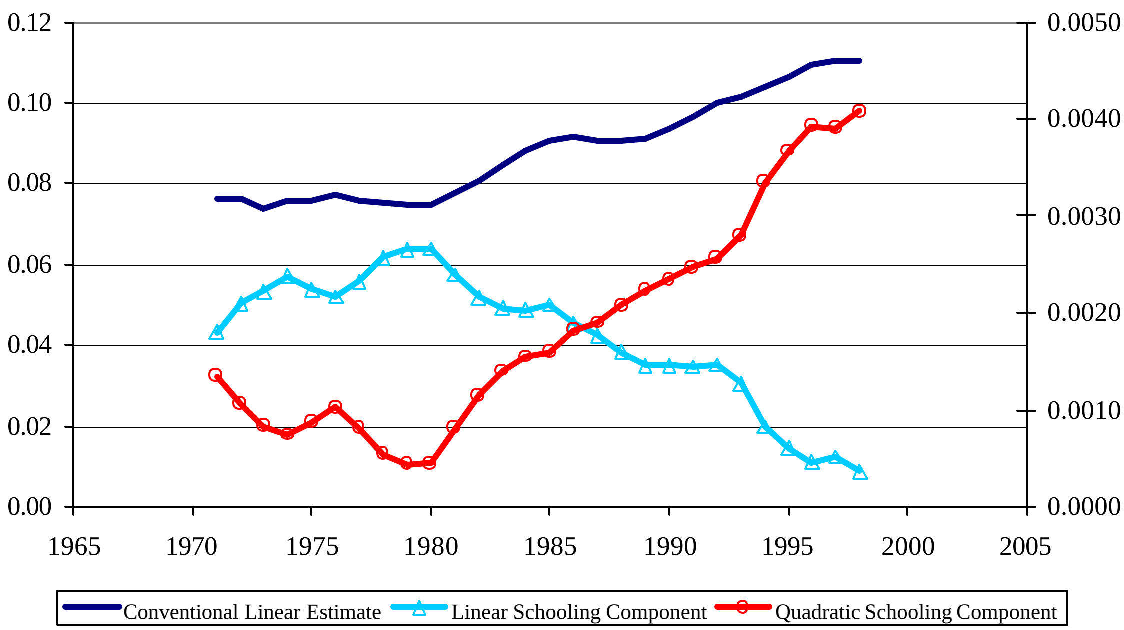


Figure 5: (a) Profile of the Earnings-Schooling Relationship: Linear Schooling Component, 1979-2000.

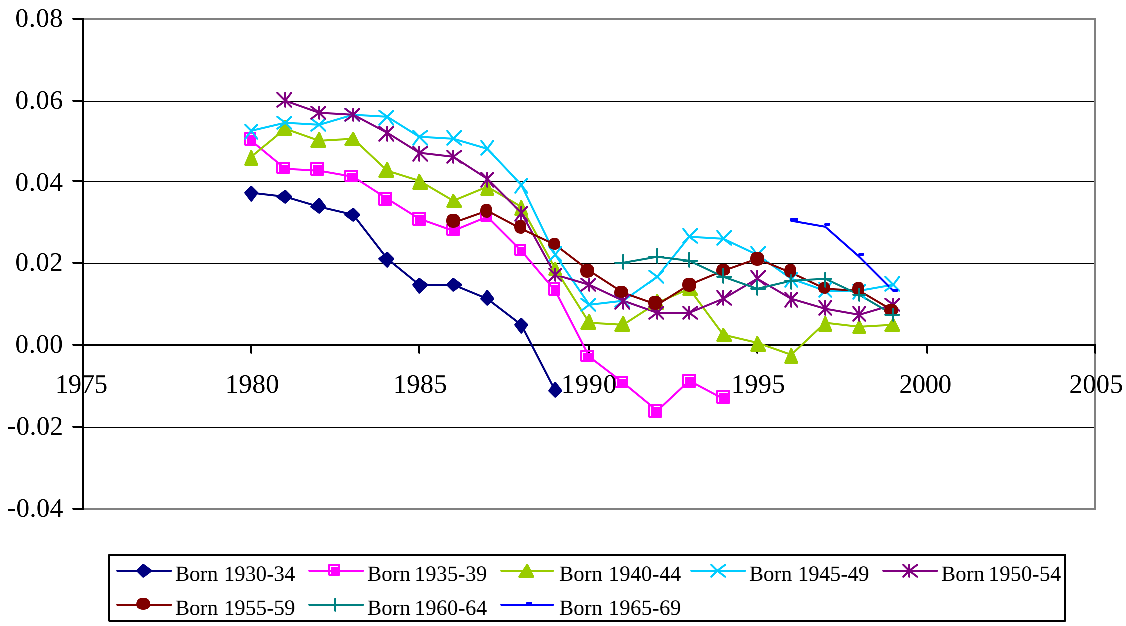


Figure 5: (b) Profile of the Earnings-Schooling Relationship: Quadratic Schooling Component, 1979-2000.

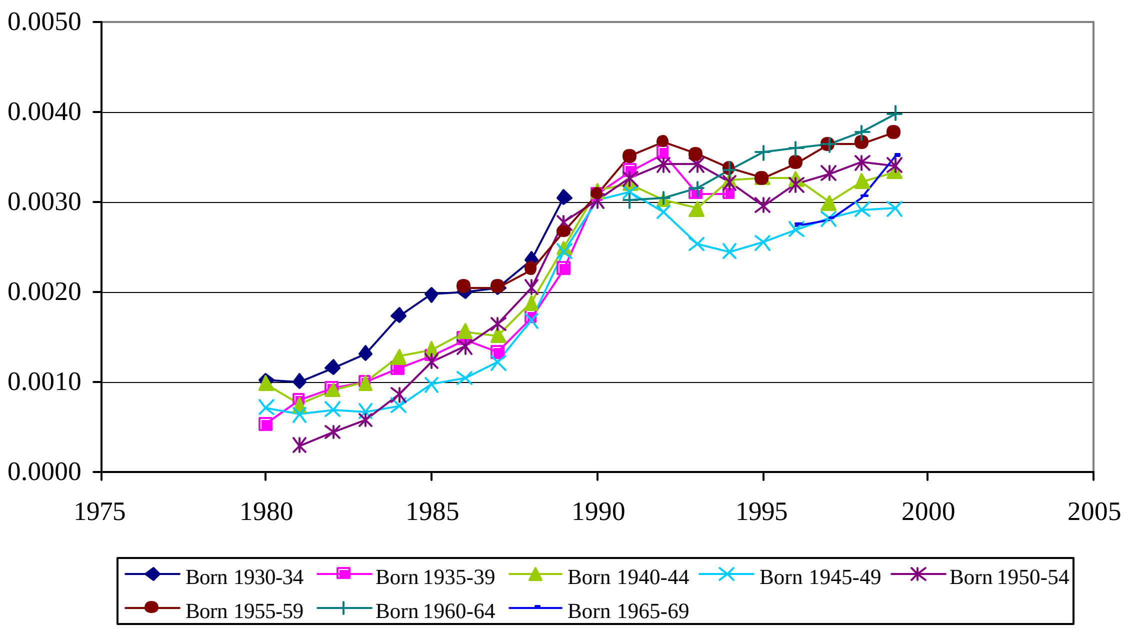


Figure 6: Estimates of the Year-Specific Returns to Education and Unobserved Ability (1979=1).
[First step regressions estimated by OLS]

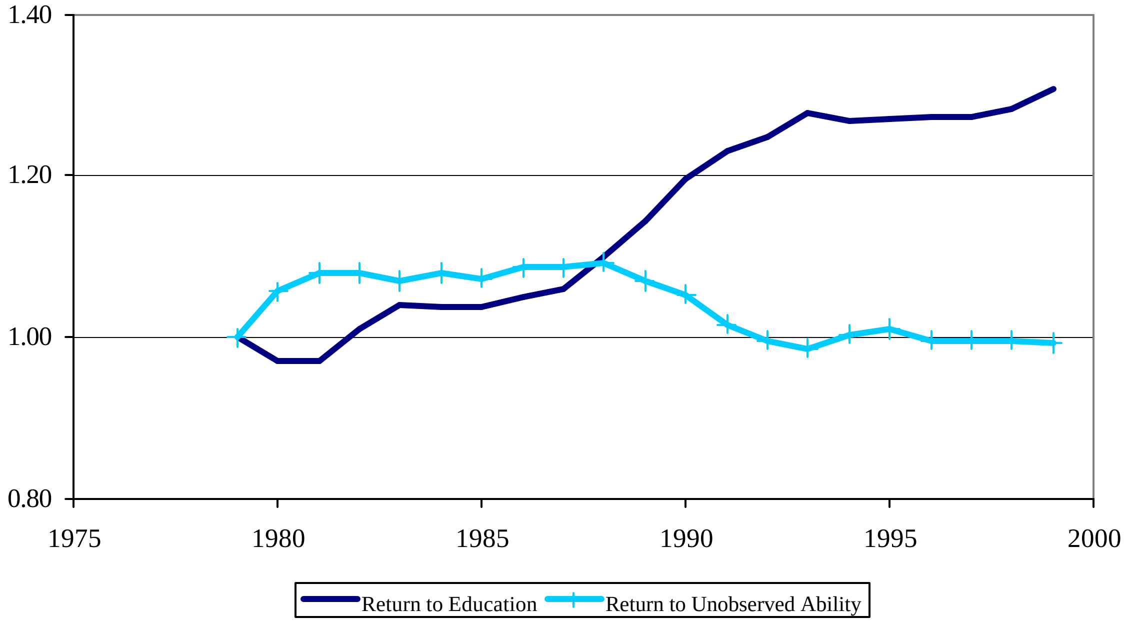


Figure 7: Estimates of the Year-Specific Returns to Education and Unobserved Ability (1979=1).
[First step regressions estimated by OLS and Tobit]

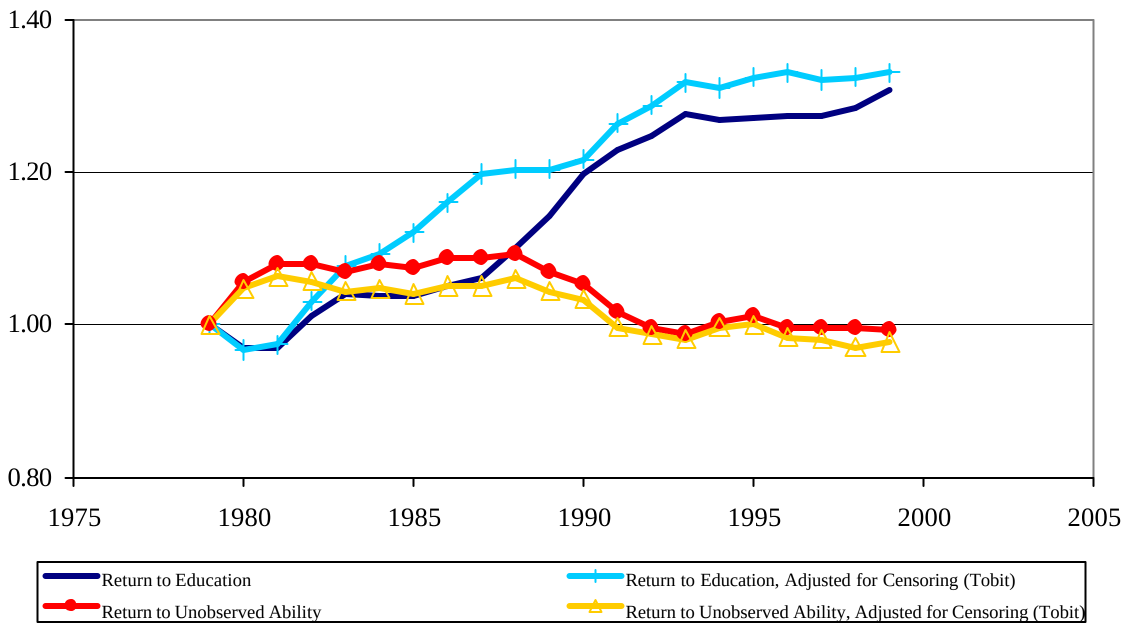


Figure 8: Estimates of the Year-Specific Return to Education (1979=1).
[First step regressions estimated by OLS, Tobit and CLAD]

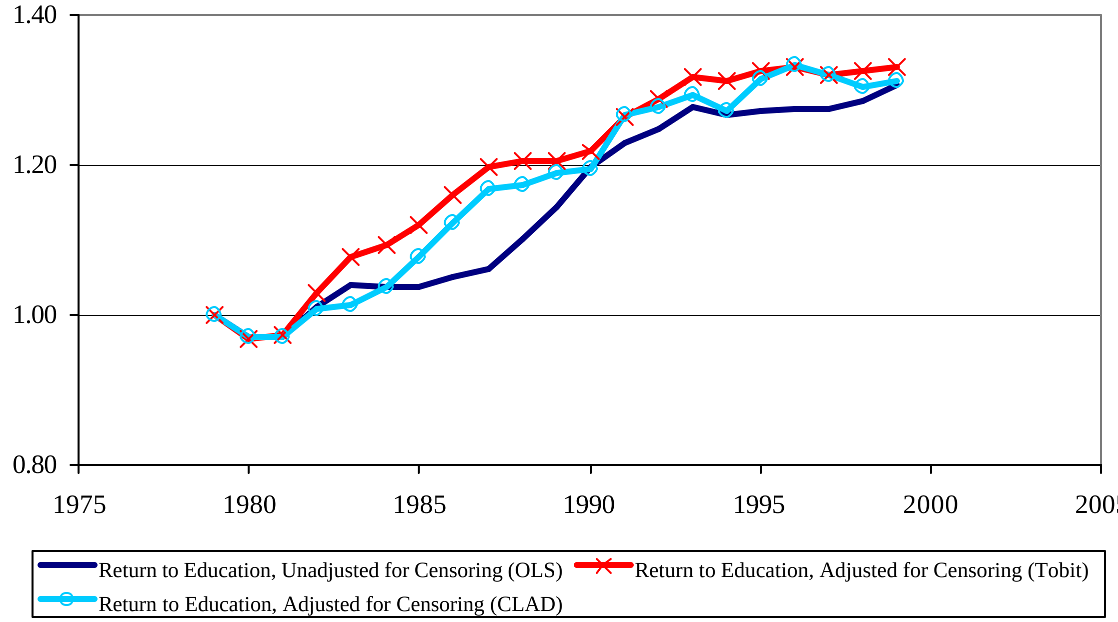
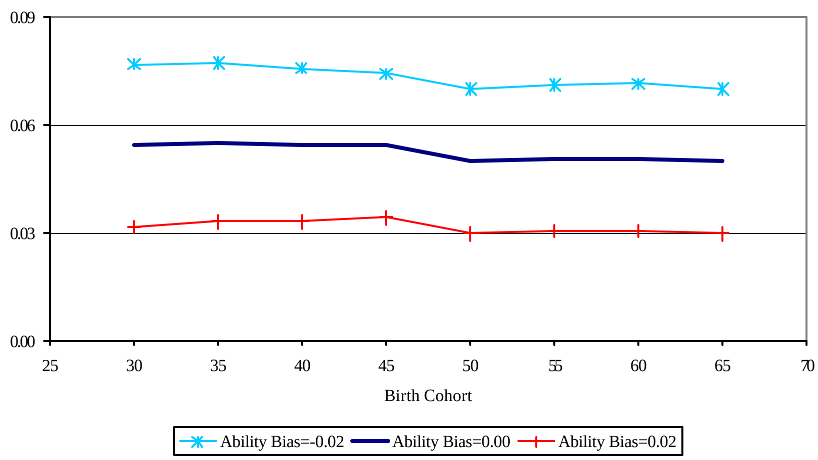
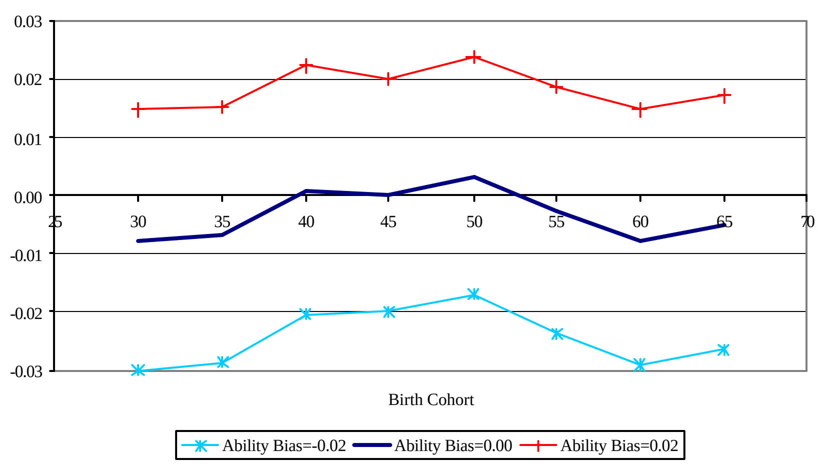


Figure 9: Estimates of the Cohort-Specific Parameters
under Alternative Normalizations for the Level of Ability Bias of the Reference Cohort.

[A] Average Marginal Productivity of Schooling (b_c)



[B] Relative Ability Bias (λ_1^c)



[C] Comparative Advantage Bias ($\lambda_2^c \overline{S}_c$)

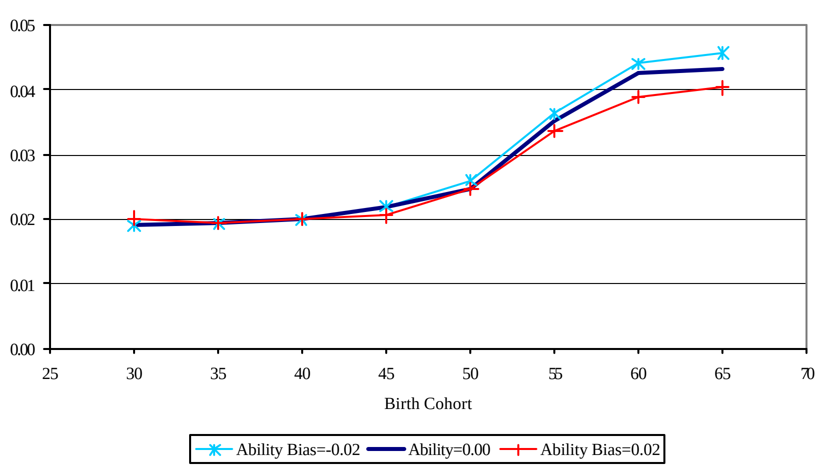


Figure 10: Goodness-of-Fit Statistics Associated with Different Transformations of Years of Education.
[First step regressions estimated by OLS]

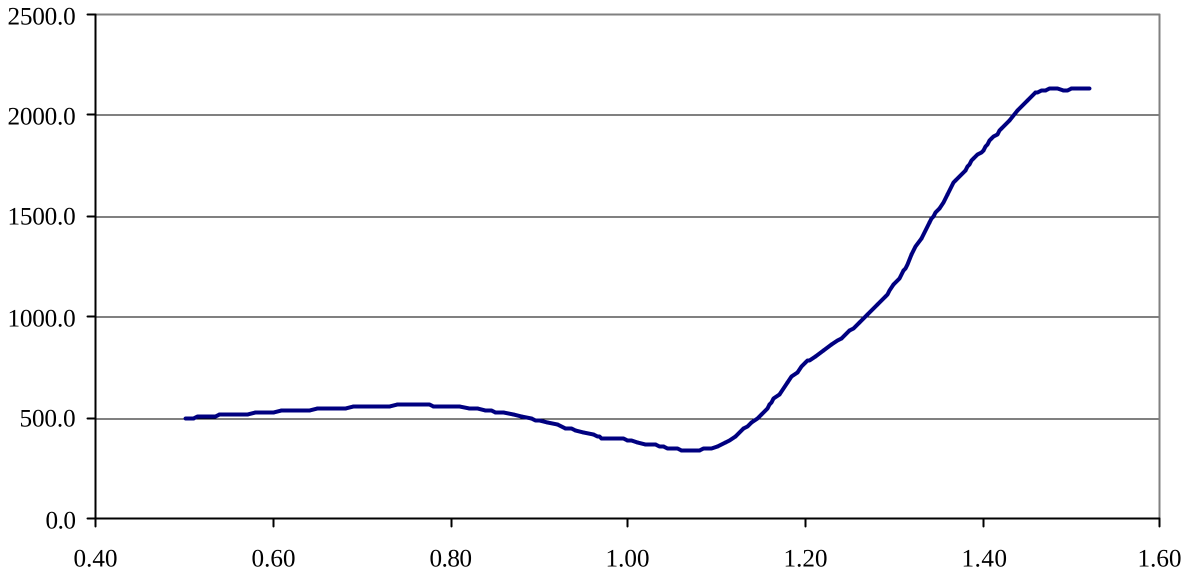


Table 1: Summary Statistics and Description of the Cohorts.

	Born 1930-34	**Born 1935-39**	**Born 1940-44**	**Born 1945-49**	**Born 1950-54**	**Born 1955-59**	**Born 1960-64**	**Born 1965-69**
Log Real Weekly Earnings	6.61 (0.46)	6.61 (0.47)	6.63 (0.48)	6.61 (0.47)	6.54 (0.47)	6.51 (0.48)	6.46 (0.49)	6.41 (0.48)
Years of Education	12.52 (3.26)	12.84 (3.06)	13.27 (2.92)	13.68 (2.79)	13.62 (2.60)	13.41 (2.49)	13.33 (2.45)	13.41 (2.48)
Entry Year into the sample	1979	1979	1979	1979	1980	1985	1990	1995
Exit Year from the sample	1990	1995	2000	2000	2000	2000	2000	2000
Age at Entry	47.03 (1.43)	41.94 (1.43)	36.88 (1.40)	31.85 (1.39)	28.01 (1.43)	28.01 (1.41)	28.08 (1.40)	28.08 (1.42)
Age at Exit (or last year in sample)	57.93 (1.40)	57.81 (1.41)	57.74 (1.39)	52.81 (1.37)	47.94 (1.41)	42.98 (1.42)	38.04 (1.42)	33.04 (1.42)
Experience at Entry	28.70 (3.63)	23.32 (3.49)	17.83 (3.34)	12.41 (3.26)	8.60 (2.87)	8.70 (2.78)	8.86 (2.86)	8.67 (2.67)
Experience at Exit (or last year in sample)	39.24 (3.62)	38.88 (3.25)	38.38 (3.29)	33.04 (3.19)	28.28 (2.97)	23.54 (2.91)	18.72 (2.95)	13.62 (3.00)
Observations	36,072	50,756	75,635	100,247	108,626	85,191	55,173	25,535

Notes: Standard deviations in parentheses.
The sample includes men aged 26-60 working for pay on a full-time basis,
and earning at least $250 per week in 1999 constant dollars.

Table 2: OLS Estimates of the Earnings-Schooling Relationship Intercepts, 1979-2000.

	Born 1930-34	**Born 1935-39**	**Born 1940-44**	**Born 1945-49**	**Born 1950-54**	**Born 1955-59**	**Born 1960-64**	**Born 1965-69**
1979-1981	6.55 [0.02]	6.54 [0.01]	6.56 [0.01]	6.58 [0.01]	6.53 [0.01]	---	---	---
1982-1984	6.55 [0.02]	6.56 [0.01]	6.57 [0.01]	6.56 [0.01]	6.55 [0.01]	---	---	---
1985-1987	6.55 [0.02]	6.56 [0.01]	6.57 [0.01]	6.56 [0.01]	6.56 [0.01]	6.56 [0.01]	---	---
1988-1990	6.55 [0.03]	6.56 [0.02]	6.56 [0.01]	6.55 [0.01]	6.55 [0.01]	6.57 [0.01]	---	---
1991-1993	---	6.52 [0.02]	6.53 [0.01]	6.53 [0.01]	6.53 [0.01]	6.57 [0.01]	6.57 [0.02]	---
1994-1996	---	6.58 [0.03]	6.55 [0.02]	6.52 [0.01]	6.51 [0.01]	6.56 [0.01]	6.54 [0.01]	6.54 [0.02]
1997-2000	---	---	6.57 [0.02]	6.57 [0.01]	6.57 [0.01]	6.60 [0.01]	6.62 [0.01]	6.59 [0.02]

Notes: Robust standard errors in brackets.
See the notes in Table 1 for a description of the sample used.
The entries are the estimates of $\pi_0(c,t)$ for 3-year bands.

Table 3: Estimated Profile of the Earnings-Schooling Relationship 1979-2000: Linear Component.

	Born 1930-34	**Born 1935-39**	**Born 1940-44**	**Born 1945-49**	**Born 1950-54**	**Born 1955-59**	**Born 1960-64**	**Born 1965-69**
1979-1981	3.71	5.02	5.87	6.17	6.35	---	---	---
	[0.41]	[0.45]	[0.47]	[0.42]	[0.56]			
1982-1984	3.21	4.10	5.15	5.71	5.67	---	---	---
	[0.48]	[0.52]	[0.50]	[0.49]	[0.50]			
1985-1987	1.42	2.77	3.65	5.16	4.60	2.98	---	---
	[0.50]	[0.51]	[0.53]	[0.54]	[0.51]	[0.51]		
1988-1990	-1.09	1.38	3.55	3.14	1.84	1.81	---	---
	[0.55]	[0.55]	[0.54]	[0.50]	[0.50]	[0.49]		
1991-1993	---	-1.58	0.59	1.12	0.63	0.87	1.99	---
		[0.65]	[0.63]	[0.55]	[0.56]	[0.54]	[0.56]	
1994-1996	---	---	-0.43	1.86	1.42	1.96	1.31	2.84
			[0.64]	[0.60]	[0.58]	[0.57]	[0.56]	[0.74]
1997-2000	---	---	0.51	1.52	0.85	1.11	1.00	1.77
			[0.61]	[0.52]	[0.46]	[0.45]	[0.45]	[0.46]

Notes: All entries are multiplied by 100.
Robust standard errors in brackets.
See the notes in Table 1 for a description of the sample used.
The entries are the estimates of $\pi_1(c,t)$ for 3-year bands.

Table 4: Estimated Profile of the Earnings-Schooling Relationship 1979-2000: Quadratic Component.

	Born 1930-34	**Born 1935-39**	**Born 1940-44**	**Born 1945-49**	**Born 1950-54**	**Born 1955-59**	**Born 1960-64**	**Born 1965-69**
1979-1981	0.10 [0.02]	0.05 [0.02]	0.09 [0.02]	0.06 [0.02]	0.03 [0.02]	---	---	---
1982-1984	0.13 [0.02]	0.10 [0.02]	0.10 [0.02]	0.07 [0.02]	0.06 [0.02]	---	---	---
1985-1987	0.20 [0.02]	0.15 [0.02]	0.16 [0.02]	0.10 [0.02]	0.14 [0.02]	0.21 [0.02]	---	---
1988-1990	0.30 [0.02]	0.22 [0.02]	0.13 [0.02]	0.13 [0.02]	0.17 [0.02]	0.17 [0.02]	---	---
1991-1993	---	0.35 [0.02]	0.30 [0.02]	0.30 [0.02]	0.34 [0.02]	0.36 [0.02]	0.30 [0.02]	---
1994-1996	---	---	0.33 [0.02]	0.26 [0.02]	0.30 [0.02]	0.33 [0.02]	0.36 [0.02]	0.27 [0.03]
1997-2000	---	---	0.32 [0.02]	0.28 [0.02]	0.34 [0.02]	0.37 [0.02]	0.39 [0.02]	0.33 [0.02]

Notes: All entries are multiplied by 100.
Robust standard errors in brackets.
See the notes in Table 1 for a description of the sample used.
The entries are the estimates of $\pi_2(c,t)$ for 3-year bands.

Table 5: Optimal Minimum Distance Estimates of the Cohort-Specific Parameters.
[First step regressions estimated by OLS]

Birth Cohort	**(1) Average Marginal Productivity of Schooling** (b_c)	**(2) Relative Ability Bias** (λ_1^c)	**(3) Comparative Advantage Bias** ($\lambda_2^c \bar{S}_c$)	**(4) Average Conventional Return to Schooling**
1930-34	0.0542	-0.0076	0.0188	0.0721
	[0.004]	[0.004]	[0.005]	[0.001]
1935-39	0.0550	-0.0069	0.0193	0.0799
	[0.003]	[0.003]	[0.005]	[0.001]
1940-44	0.0543	0.0008	0.0199	0.0882
	[0.003]	[0.003]	[0.005]	[0.001]
1945-49	0.0541	0.0000	0.0205	0.0862
	[0.002]	---	[0.004]	[0.001]
1950-54	0.0497	0.0033	0.0245	0.0866
	[0.003]	[0.003]	[0.005]	[0.001]
1955-59	0.0506	-0.0028	0.0349	0.0918
	[0.003]	[0.004]	[0.007]	[0.001]
1960-64	0.0507	-0.0076	0.0427	0.0892
	[0.003]	[0.004]	[0.008]	[0.001]
1965-69	0.0498	-0.0051	0.0429	0.0789
	[0.004]	[0.005]	[0.008]	[0.001]
Observations	381	381	381	537,150
N×Objective	391.0	391.0	391.0	
[χ^2]	[357.4]	[357.4]	[357.4]	---
(d.f.)	(315)	(315)	(315)	

Notes: Robust standard errors in brackets.
The 381 coefficients of the augmented human capital earnings regression (6) are estimated by OLS. The ability bias parameter for the cohort of men born 1945-49 is normalized to be 0.00. The standard errors in column 3 are computed using a Taylor series approximation. The entries in column 4 are the within-cohort lifetime averages of the OLS estimates of the conventional measure of the return to education (3).

Table 6: OMD Estimates of the Year-Specific Returns to Education and Unobserved Ability.
[Selected Years, First step regressions estimated by OLS]

Year	**Return to Education** (δ_t)	**Return to Unobserved Ability** (Ψ_t)
1979	1.00 ---	1.00 ---
1982	1.00 [0.03]	1.07 [0.02]
1985	0.99 [0.03]	1.10 [0.02]
1988	1.02 [0.03]	1.10 [0.02]
1991	1.18 [0.04]	1.05 [0.02]
1994	1.27 [0.04]	1.02 [0.02]
1997	1.28 [0.04]	0.98 [0.02]
2000	1.35 [0.04]	0.97 [0.02]
N×Objective [χ^2] (d.f.)	391.0 [357.4] (315)	391.0 [357.4] (315)

Notes: Robust standard errors in brackets.

Table 7: Optimal Minimum Distance Estimates of the Cohort-Specific Parameters.
[First step regressions estimated by Tobit]

Birth Cohort	**(1) Average Marginal Productivity of Schooling** (b_c)	**(2) Relative Ability Bias** (λ_1^c)	**(3) Comparative Advantage Bias** ($\lambda_2^c \bar{S}_c$)	**(4) Average Conventional Return to Schooling**
1930-34	0.0501	-0.0074	0.0225	0.0777
	[0.003]	[0.004]	[0.006]	[0.002]
1935-39	0.0506	-0.0071	0.0231	0.0862
	[0.003]	[0.003]	[0.006]	[0.002]
1940-44	0.0497	0.0011	0.0226	0.0933
	[0.003]	[0.003]	[0.005]	[0.002]
1945-49	0.0500	0.0000	0.0233	0.0900
	[0.002]	---	[0.005]	[0.001]
1950-54	0.0466	0.0023	0.0272	0.0884
	[0.002]	[0.003]	[0.005]	[0.001]
1955-59	0.0472	-0.0035	0.0362	0.0936
	[0.003]	[0.003]	[0.007]	[0.002]
1960-64	0.0477	-0.0080	0.0413	0.0886
	[0.003]	[0.004]	[0.008]	[0.002]
1965-69	0.0463	-0.0032	0.0416	0.0838
	[0.003]	[0.005]	[0.008]	[0.003]
Observations	381	381	381	537,150
N×Objective	475.0	475.0	475.0	---
[χ^2]	[357.4]	[357.4]	[357.4]	
(d.f)	(315)	(315)	(315)	

Notes: Robust standard errors in brackets.
The 381 coefficients of the augmented human capital earnings regression (6) are estimated by Tobit. The ability bias parameter for the cohort of men born 1945-49 is normalized to be 0.00. The standard errors in column 3 are computed using a Taylor series approximation. The entries in column 4 are the within-cohort lifetime averages of the Tobit estimates of the conventional measure of the return to education (3).

Table 8: OMD Estimates of the Year-Specific Returns to Education and Unobserved Ability.
[Selected Years, First step regressions estimated by Tobit]

Year	**Return to Education (δ_t)**	**Return to Unobserved Ability (Ψ_t)**
1979	1.00	1.00
	---	---
1982	1.02	1.05
	[0.03]	[0.02]
1985	1.07	1.06
	[0.03]	[0.02]
1988	1.18	1.06
	[0.03]	[0.02]
1991	1.22	1.03
	[0.04]	[0.02]
1994	1.31	1.01
	[0.04]	[0.02]
1997	1.33	0.96
	[0.04]	[0.02]
2000	1.35	0.98
	[0.04]	[0.02]
N×Objective	475.0	475.0
[χ^2]	[357.4]	[357.4]
(d.f)	(315)	(315)

Notes: Robust standard errors in brackets.

Table 9: Optimal Minimum Distance Estimates of the Cohort-Specific Parameters.
[First step regressions estimated by CLAD]

Birth Cohort	**(1) Average Marginal Productivity of Schooling** (b_c)	**(2) Relative Ability Bias** (λ_1^c)	**(3) Comparative Advantage Bias** ($\lambda_2^c \bar{S}_c$)	**(4) Average Conventional Return to Schooling**
1930-34	0.0541 [0.003]	-0.0192 [0.003]	0.0351 [0.009]	0.0743 [0.001]
1935-39	0.0566 [0.003]	-0.0066 [0.003]	0.0180 [0.004]	0.0856 [0.001]
1940-44	0.0578 [0.003]	0.0019 [0.003]	0.0212 [0.005]	0.0945 [0.001]
1945-49	0.0588 [0.002]	0.0000 ---	0.0192 [0.004]	0.0901 [0.001]
1950-54	0.0583 [0.003]	-0.0016 [0.003]	0.0232 [0.005]	0.0910 [0.001]
1955-59	0.0602 [0.003]	-0.0042 [0.003]	0.0295 [0.006]	0.0975 [0.001]
1960-64	0.0572 [0.003]	-0.0076 [0.004]	0.0387 [0.008]	0.0961 [0.001]
1965-69	0.0567 [0.004]	-0.0068 [0.006]	0.0322 [0.007]	0.0833 [0.001]
Observations	381	381	381	537,150
N×Objective [χ^2] (d.f.)	677.6 [357.4] (315)	677.6 [357.4] (315)	677.6 [357.4] (315)	---

The 381 coefficients of the augmented human capital earnings regression (6) are estimated by CLAD. The ability bias parameter for the cohort of men born 1945-49 is normalized to be 0.00. The standard errors in column 3 are computed using a Taylor series approximation. The entries in column 4 are the within-cohort lifetime averages of the CLAD estimates of the conventional measure of the return to education (3).

Table 10: OMD Estimates of the Year-Specific Returns to Education and Unobserved Ability.
[Selected Years, First step regressions estimated by CLAD]

Year	**Return to Education** (δ_t)	**Return to Unobserved Ability** (Ψ_t)
1979	1.00 ---	1.00 ---
1982	1.00 [0.04]	1.00 [0.01]
1985	1.07 [0.04]	0.99 [0.01]
1988	1.20 [0.04]	0.96 [0.01]
1991	1.21 [0.05]	0.97 [0.01]
1994	1.26 [0.05]	0.98 [0.02]
1997	1.32 [0.05]	0.95 [0.02]
2000	1.34 [0.05]	0.99 [0.02]
N×Objective [χ^2] (d.f)	677.6 [357.4] (315)	677.6 [357.4] (315)

Notes: Robust standard errors in brackets.

Table 11: Goodness-of-Fit Statistics for Alternative Models.

	(1) First Step: OLS	(2) First Step: Tobit	(3) First Step: CLAD
1. Two-Factor Model	391.0	475.0	677.6
[d.f]	[315]	[315]	[315]
(p-value)	(0.00)	(0.00)	(0.00)
2. Two-Factor Model, Constant b_c across cohorts	11.1	8.2	35.4
$\{\chi^2\}$	{14.1}	{14.1}	{14.1}
[d.f]	[7]	[7]	[7]
(p-value)	(0.13)	(0.32)	(0.00)
3. Two-Factor Model, Constant λ_2^c across cohorts	219.2	181.8	113.9
$\{\chi^2\}$	{14.1}	{14.1}	{14.1}
[d.f]	[7]	[7]	[7]
(p-value)	(0.00)	(0.00)	(0.00)
4. Two-Factor Model, Stationary δ_t	527.4	637.0	818.7
$\{\chi^2\}$	{31.4}	{31.4}	{31.4}
[d.f]	[20]	[20]	[20]
(p-value)	(0.00)	(0.00)	(0.00)
5. Two-Factor Model, Stationary Ψ_t	303.9	302.6	423.3
$\{\chi^2\}$	{31.4}	{31.4}	{31.4}
[d.f]	[20]	[20]	[20]
(p-value)	(0.00)	(0.00)	(0.00)
6. Two-Factor Model, No Comparative Advantage Bias ($\lambda_2^c=0$)	2453.0	3379.5	1560.4
$\{\chi^2\}$	{15.5}	{15.5}	{15.5}
[d.f]	[8]	[8]	[8]
(p-value)	(0.00)	(0.00)	(0.00)
7. One-Factor Model ($\delta_t=\Psi_t$)	653.5	783.1	1041.5
$\{\chi^2\}$	{31.4}	{31.4}	{31.4}
[d.f]	[20]	[20]	[20]
(p-value)	(0.00)	(0.00)	(0.00)

The entries in row 1 are the goodness-of-fit statistics associated with the two-factor model presented in Tables 5 to 10. The entries in rows 2-7 are the differences in the goodness-of-fit statistics between each row and the two-factor model of row 1.

Table A.1: Simulations and Measurement Error Corrected Optimal Minimum Distance Estimates.
[selected years and cohorts]

	Actual (from Table 5)	**Design 1** (p_1=, p_2=) **(0.08,0.001)**	**Design 2** (p_1=, p_2=) **(0.06,0.002)**	**Design 3** (p_1=, p_2=) **(0.04,0.003)**	**Design 4** (p_1=, p_2=) **(0.02,0.004)**
[A] Simulations Results[1]					
Intercept	---	0.98	0.98	0.99	0.99
Linear Schooling Term	---	1.08	1.03	0.94	0.75
Quadratic Schooling Term	---	1.25	1.25	1.25	1.25
[B] Measurement Error Corrected OMD Estimates[2]					
Return to Education					
1979	1.00	1.00	1.00	1.00	1.00
1989	1.20	1.30	1.34	1.40	1.58
1999	1.30	1.43	1.47	1.55	1.76
Return to Unobserved Ability					
1979	1.00	1.00	1.00	1.00	1.00
1989	1.08	1.06	1.06	1.07	1.07
1999	0.99	0.98	0.98	0.99	1.00
Cohort Born 1940-44					
Average Slope (b_c)	0.0543	0.0591	0.0567	0.0525	0.0438
Relative Ability Bias (λ_1^c)	0.0008	0.0012	0.0011	0.0009	0.0003
Comparative Advantage Bias ($\lambda_2^c \overline{S}_c$)	0.0245	0.0239	0.0239	0.0226	0.0212
Cohort Born 1960-64					
Average Slope (b_c)	0.0507	0.0567	0.0548	0.0515	0.0444
Relative Ability Bias (λ_1^c)	-0.0076	-0.0075	-0.0069	-0.0059	-0.0037
Comparative Advantage Bias ($\lambda_2^c \overline{S}_c$)	0.0427	0.0493	0.0480	0.0453	0.0413

(1) The entries in the top panel correspond to the ratios of the true/estimated OLS coefficients from the simulations. The simulations were based on 10,000 replications from the following model with 2,000 observations, and assuming that the reliability ratio of reported schooling is 0.90:

$\log y = 5.0 + \pi_1 S + \pi_2 S^2 + e$
$e \sim N(0 , 0.2)$
$S \sim N(12 , 9)$
$S^o = S + v, \quad v \sim N(0 , \sigma_v^2)$

(2) These entries are the measurement-error corrected OMD estimates of the two-factor model.
The ability bias parameter for the cohort of men born 1945-49 is normalized to be 0.00.
The correction factors were obtained by the simulations and are reported in the top panel of the table.

Table A.2: IWOMD Estimates of the Cohort-Specific Parameters.
[First step regressions estimated by OLS]

Birth Cohort	(1) **Average Marginal Productivity of Schooling** (b_c)	(2) **Relative Ability Bias** (λ_1^c)	(3) **Comparative Advantage Bias** $(\lambda_2^c \overline{S}_c)$	(4) **Average Conventional Return to Schooling**
1930-34	0.0535	-0.0055	0.0175	0.0724
	[0.004]	[0.004]	[0.005]	[0.001]
1935-39	0.0541	-0.0060	0.0193	0.0805
	[0.003]	[0.003]	[0.005]	[0.001]
1940-44	0.0532	0.0011	0.0199	0.0882
	[0.003]	[0.003]	[0.005]	[0.001]
1945-49	0.0535	0.0000	0.0219	0.0863
	[0.002]	---	[0.005]	[0.000]
1950-54	0.0484	0.0036	0.0259	0.0866
	[0.003]	[0.003]	[0.005]	[0.000]
1955-59	0.0497	-0.0014	0.0349	0.0922
	[0.003]	[0.003]	[0.007]	[0.001]
1960-64	0.0490	-0.0061	0.0427	0.0870
	[0.003]	[0.004]	[0.008]	[0.001]
1965-69	0.0481	-0.0046	0.0443	0.0796
	[0.004]	[0.005]	[0.009]	[0.001]
Observations	762	762	762	1,074,468
N×Objective	987.9	987.9	987.9	---
$[\chi^2]$	[758.5]	[758.5]	[758.5]	
(d.f)	(696)	(696)	(696)	

Notes: Robust standard errors in brackets.
The 762 coefficients of the augmented human capital earnings regression (6) are estimated by OLS. For IWOMD, the sample is partitioned in two independent groups (the CPS rotation groups), and each one is used to compute OMD estimates. See the text for more details. The ability bias for the cohort of men born 1945-1949 is normalized to be 0.00. The standard errors in column 3 are computed using a Taylor series approximation. The entries in column 4 are the within-cohort lifetime averages of the conventional measure of the return to education (3).