UNOBSERVED ABILITY, COMPARATIVE ADVANTAGE, AND THE RISING RETURN TO EDUCATION IN THE UNITED STATES 1979-2002.

Olivier Deschênes[1]

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

June 2001:
This version: December 2003

Abstract: This paper quantifies the change in the causal effect of education on labor market earnings in the United States between 1979 and 2002. In absence of valid instrumental variables for schooling, I develop a causal model for earnings and schooling that incorporates heterogeneity in absolute and comparative advantage across individuals. The model 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. 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 effects 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 20-40% between 1979 and 2002, after controlling for the confounding effects of time-varying ability and comparative advantage biases across cohorts. This increases explains most of the observed change in the educational wage structure in the U.S. over that time period.

JEL classification: J24, I21, C21.
Keywords: wage structure, causal effect of education, ability bias, comparative advantage.

[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, Peter Kuhn, Thomas Lemieux, 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 an earnings regression—almost doubled.[4] 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 (Juhn, Murphy and Pierce 1993).[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 is based on the unobservable determinants of earnings that are potentially correlated with schooling. A rise in the return to unobservable skills, or an increasing degree of ability-education sorting across cohorts may also have contributed to a rising observed association between schooling and earnings (see e.g. Taber 2001).[6]

[1] See for example Angrist and Krueger (1991), Ashenfelter and Rouse (1998), and Meghir and Palme (1999). Card (2001) presents an overview of recent studies.

[2] Most advancements relate to the estimation of models with heterogeneous treatment effect using instrumental variables. See Angrist and Imbens (1995), Card (2001), Heckman and Vytlacil (1998), Heckman, Tobias and Vytlacil (2000), and Wooldridge (1997, 2003).

[3] See Katz and Autor (1999) for an extensive overview of recent changes in 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] Lemieux (2003) shows that the choice of data set and the wages series analyzed greatly influences the patterns in residual wage inequality over the period 1976-2003. Some of his estimates suggest that residual wage inequality only accounts for 25% of the overall rise in wage inequality over that time period.

[6] The analysis by Taber (2001) only considers the first possibility. See also Blackburn and Neumark (1993), and

What both of these interpretations have in common is the idea that changes in the educational wage structure can arise along two dimensions: First, economy-wide changes in the demand for educated workers, or for unobserved skills will generate year effects in the observed return to education. Second, the observed relationship between schooling and wages may change over time because of composition or cohort effects.[7] Combined, both dimensions of change in the educational wage structure entail a time-varying causal effect of education and time-varying ability biases. While previous analysts have evaluated some of these explanations separately, the absence of a unifying framework addressing all of these possibilities 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 assess 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 a repeated cross-section setting. Following Willis and Rosen (1979), but unlike most recent studies, two factors of unobserved ability are considered: absolute ability and heterogeneous returns to education. If the schooling decisions of individuals are influenced by the unobserved ability factors, the conventional estimate of the return to education will differ from the average causal effect of education. Consequently, 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 individuals 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. I show below that 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 estimated coefficients from an augmented human capital

Murnane, Willett and Levy (1995).

[7] Gossling, Machin and Meghir (2000) document the importance of cohort effects in explaining changes in the male wage distribution in the U.K.

earnings regression. This identification result is robust to nonlinearity in the structural earnings function and in the conditional mean function of the unobserved ability factors.

The empirical analysis is based on repeated cross-sectional data from the Current Population Survey for the years 1979-2002, and includes cohorts of men born between 1930 and 1969. I begin by documenting the marked increase in the convexity of the relationship between log earnings and schooling over the 1980s and 1990s, as was first noted by Mincer (1996). In addition, I show that this increase was experienced by all cohorts of workers in the labor market. I then make use of 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. A series of specification tests indicates that the two-factor model greatly dominates the more conventional one-factor model of ability that has been the cornerstone of the literature.

The estimates show an increase in the year-specific causal return to education of about 21-36% between 1979 and 2002, as opposed to 62% for the conventional estimate of the return to schooling. The return to unobserved ability increased by 8-13% during the same time period. The estimates also 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 undermines the contention that the rising return to unobserved ability is the single driving force behind the important increase in the cross-sectional association between schooling and earnings over the 1980s and 1990s.

The remainder of the paper is organized as follows. Section 2 describes the data and provides a brief overview of the changes in the educational wage structure between 1979 and 2002 in the United States. Section 3 presents a framework for analyzing changes in the educational wage structure. Section 4 discusses the identification and estimation of the year-specific parameters. Section 5 describes the identification and estimation of the year-specific and cohort-specific parameters. Section 6 presents a sensitivity analysis, and Section 7 concludes.

2. The Changing Relationship Between Earnings and Schooling, 1979-2002.

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 2002. This choice is motivated by several practical reasons, most importantly, by the larger sample sizes and the wide span of cohorts provided by the ORG. Other data sets, such as the National Longitudinal Survey of Youth (NLSY), have also been used by previous analysts. However the NLSY only tracks a small range of cohorts over time. As shown by Heckman and Vytlacil (2001), this feature of the design of the NLSY greatly limits the capacity of researchers to separate out age, cohort, and time effects. Since one objective of this study is to assess the contribution of changes in the mapping between ability and education across cohorts to the changing educational wage structure, the ORG files from the CPS are a better-suited data source.

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. Following most of the literature, I use the hourly wage rate as the dependent variable the analysis. This choice is motivated by the fact that most theories of wage determination pertain to the hourly wage rate. Hourly wage rates are constructed following Lemieux (2003): For workers paid by the hour, which represent about 50% of the workforce during that time period, I use the hourly wage rate reported in the ORG files. Weekly earnings (and hours worked) are also reported for workers not paid on an hourly basis. For these workers, I construct the hourly wage rate by dividing weekly earnings by usual hours of work. Following DiNardo, Fortin and Lemieux (1996), all the models in this paper are weighted by weekly hours of work. This weighted scheme ensures that workers who are more strongly attached to the labor market receive more weight in the estimation. An alternative would be to restrict the sample to full-time workers. These considerations do not affect the main conclusion of this paper.

Nominal hourly wage rates are converted to 2002 constant dollars using the GDP deflator for personal consumption expenditures. To limit the potential influence of outliers, I deleted all observations with an hourly wages below 5.00 or above 100.00 in 2002 constant dollars. The results in this paper are not sensitive to this restriction. Finally, to maintain the 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. Throughout the paper, "cohort" is intended to mean one of these 8 groups. 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). Based on this specification, there are 137 cohort/year pairs in the sample. Each entry in Table 1 represents the cohort-specific average of the variable listed, for all the years in which a cohort is observed in the sample. The first row shows that younger cohorts have lower real hourly wage rates, reflecting a combination of age differences, and of the overall decline in average real earnings in the United States over the 1980s. 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 2002. Clearly, this last pattern is not the consequence of a secular increase in educational attainment between successive

[8] The rotation group structure of the CPS implies an overlap of about half of the sample from year to year. Without independent samples from year to year, the estimated regression coefficients would be serially correlated.

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

birth cohorts. Finally, the fraction of Hispanic workers has increased over time, due in most part to immigration. This implies that Hispanic workers are more strongly represented in recent cohorts. In order to deal with this issue, controls for Hispanic ethnicity are included in all models. Estimation results from samples that are restricted to white non-hispanic men are available from the author. The results of this paper are robust to the inclusion or exclusion of Hispanic workers.

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

The standard human capital earnings function specifies log wages as a linear function of years of education, with a constant coefficient. The log-linear specification has been used in countless 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]

In order to obtain some simple evidence on the functional form relationship linking earnings and schooling, I estimated an unrestricted regression of log hourly wage 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: 1979-1981, 1989-1991 and 1999-2002.[12] The figure suggests that between 1979 and 2002 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

[10] For example, a positive correlation between completed education and returns to schooling would arise if individuals based their schooling decisions on their expected gain from an additional year of education.

[11] The regressions also included a quartic in experience, indicators for race, ethnicity, marital status, and dummies for metropolitan area and census division.

[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. Additional results are available from the author upon request.

only reflects the uniform increase in educational wage differentials over the last twenty years, but also the spreading 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 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 1970-2002.[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. For the reason listed above, the ORG data, and not the March CPS will be used in the rest the empirical analysis.

The three panels of Figure 5 complete the descriptive analysis by presenting the intercept, linear, and quadratic coefficients from the augmented log wage regressions estimated separately for each of the birth cohort described in Table 1.[15] Panel (A) reports the intercepts from each regression. As shown below, this component of the earnings-schooling relationship provides an important source of identification in models of the changing educational wage structure. All the covariates included in the regression are standardized to have mean 0 in every year, thereby ensuring that the intercept represents average log hourly wage rate. The patterns indicate a rise in average real hourly wages (in 2002 constant dollars) in the early 1980s, followed by a decline in the late 1980s, and a rise in the late 1990s. The patterns in panels (B) and (C) of Figure 5 are similar to the patterns

[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 $y = a_1 + b_1 x + b_2 x^2 + u$. It then follows that $b_0 \approx b_1 + 2b_2\overline{x}$, where $\overline{x}$ is the sample average of x.

[15] The estimates displayed in Figure 5 are smoothed using a 3 year moving-average for each cohort. The full set of estimates is available in an appendix from the author.

illustrated in Figure 4, suggesting 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. More striking is the similarity of the trends in each components across cohorts.[16]

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 intercept, linear, and quadratic schooling coefficients among cohorts points out to a common underlying mechanism affecting all cohorts equally. The differences across cohorts in the levels of the components suggest the existence of permanent differences in the relationship between log earnings and schooling across cohorts. This paper will exploit these empirical facts as its source of identification.

3. Conceptual 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 characterizing 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.

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.[17] Empirically, the implications of endogenous schooling for the earnings function are conveniently described by a correlated random coefficient model.[18] More specifically, suppose that the log real hourly wage of an individual belonging to cohort c, observed in survey year t is

[16] Card and Lemieux (2001) interpret similar trends in the college wage premium in the context of model with imperfect substitution of workers with the same education, but different age.

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

[18] 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, 2001) discuss the correlated random coefficient model in the context of schooling models. In particular, Card (2001) shows that the cross-sectional version of (1) is consistent with a general class of optimizing models.

determined by:

$$\log y_{ict} = \Psi_t a_{ic} + \delta_t b_{ic} S_{ic} - \delta_t k_c S_{ic}^2 + \varepsilon_{ict} \tag{1}$$

where a_{ic} is the unobserved ability (i.e. the part of earnings capacity that does not interact with schooling) of individual i, who belongs to cohort c; b_{ic} is the part of the education slope specific to person i in cohort c; and S_{ic} denotes years of completed education.[19] In this model the marginal benefit of an additional year of schooling is $b_{ic} - k_c S_{ic}$, so that b_{ic} represents the intercept of the marginal benefit equation. The parameter $k_c \geqslant 0$ allows for curvature in the human capital production function, capturing diminishing returns to schooling in producing earnings capacity (Becker 1975). Note that the degree of curvature may vary across cohorts. 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.

This model allows individual heterogeneity in earnings capacity to affect both the intercept and the slope of the earnings function. At a given point in time, the variation in log wages 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 wages at all schooling levels (an "absolute advantage"), and individuals with higher values of b_{ic} will receive higher payoffs per year of education (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 return to schooling varies across cohort (i.e. different cohorts have different average values of $b_{ic} - k_c S_{ic}$), the average causal effect of education 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 (denoted by δ_t) will shift the average causal effect of education in a proportional manner for all individuals, irrespective of their cohorts. To reiterate, inter-temporal changes in the causal effect of education may be due to cohort-specific factors (variation in b_c and k_c) and to year-specific factors (variation in δ_t): In what follows, I will refer to these as the cohort-specific and the year-specific components of the average

[19] Other determinants of earning (e.g. labor market experience) are omitted for the moment but will be included in the empirical analysis.

causal effect of education.

If the unobserved determinants of log hourly wage are correlated with schooling, for example because of optimizing behavior in the schooling decisions, the observed relationship between log hourly wage 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, there are two sources of confounding: (i) the "ability bias" (due to a correlation between a_{ic} and S_{ic}); and (ii) the "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 in the conventional measure of the return to education. Moreover, if the mapping 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.[20] To formalize these ideas, suppose that the conditional expectations of a_{ic} and b_{ic} are linear in years of completed education:

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

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

where a_c and b_c are cohort-specific averages of a_{ic} and b_{ic}. The linearity assumption implies that u_{1ic} and u_{2ic} are mean-independent of S_{ic} (i.e. $E[u_{1ic}|S_{ic}] = E[u_{2ic}|S_{ic}] = 0$). 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.[21] Throughout I assume that a_c, b_c, λ_1^c, and λ_2^c are time-invariant once the cohort makes a permanent transition to the labor market. This amounts to assuming no ability-related attrition within cohort, after all its

[20] Differential ability-education mappings across cohort could arise, for example, if school enrollment decisions become more meritocratic (Herrnstein and Murray 1994).

[21] 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 $\lambda_2^c > 0$.

members made a permanent transition to the labor market. Since this study focuses on a sample of prime-aged males, there is a limited scope for this kind of attrition. In the model described by (1) and (2), the conventional measure of the return to education for cohort c observed in year t ($\beta_{ols}(c,t)$) corresponds to:[22]

$$\beta_{ols}(c,t) = \delta_t(b_c - k_c\overline{S}_c) + \Psi_t\lambda_1^c + \delta_t\lambda_2^c\overline{S}_c \tag{3}$$

Thus, for a given cohort at a given point in time, the conventional estimate of the return to education equals the sum of three components: $\delta_t(b_c - k_c\overline{S}_c)$, 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 (Ψ_t).[23] 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 causal 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.

As discussed above, most previous analyzes of the changing educational wage structure considered a restricted form of this model where heterogeneity in the education slope and cohort-specific mappings between ability and education are ruled out (i.e. it is assumed that $b_{ic} = b$, (or $\lambda_2^c = 0$) and $\lambda_1^c = \lambda_1$). As such, our current state of knowledge may be limited due to the more restrictive models employed in the previous literature. This study will contribute to this debate by using different strategies to estimate the parameters of the more general model, and by performing the appropriate specification tests of the more restrictive models.

[22] 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 Appendix B for a derivation.

[23] 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).

4. Identification and Estimation of the Year-Specific Parameters.

The model above illustrates that both levels and changes in the cross-sectional relationship between log wages and schooling are confounded. A standard econometric solution to this problem of endogeneity is the method of instrumental variables. In principle, instrumental variables techniques can be used to estimate the average causal effect of education (and its change) using relatively weak functional form assumptions. Unfortunately, instrumental variables for schooling are not readily available in large repeated cross-sectional data sets like the CPS. Moreover in some cases the instruments may only be weakly correlated with educational attainment[24]. An alternative approach to instrumental variables is to estimate a structural model of earnings determination and schooling choice using maximum likelihood methods (see e.g. Cameron and Taber (2000), Taber (2001), and Belzil and Hansen (2002)). Under distributional assumptions for the components of unobserved heterogeneity, this approach allows the specification and estimation of more flexible models. However, results from such studies are sometimes difficult to interpret, and may be non-robust to departures from the distributional assumptions. A final possibility would be the use of fixed-effect estimators that eliminate the confounding effect of time-invariant ability. However, in absence of instruments for education, fixed-effects estimates of the return to education are not identified (Hausman and Taylor 1980).

A. *Identification.*

In this paper, I propose a method based inter-cohort comparisons to identify changes in the causal effect of education over time when instrumental variables for schooling are not available. The basic idea is to relate multiple observations on the estimated coefficients from an "augmented" log wage regression to the parameters of the causal model for log wages and schooling. I begin by showing the that year-specific causal effect of education and the return to unobserved ability (δ_t and Ψ_t) are identified up to a normalization under very weak assumptions. The following assumption is identifies δ_t and Ψ_t:

A1: $E[a_{ic}|S_{ic}] = a_c + \lambda_1^c(S_{ic} - \overline{S}_c)$ and $E[b_{ic}|S_{ic}] = b_c + \lambda_2^c(S_{ic} - \overline{S}_c)$

[24] Several studies analyze the consequences of weak IV in similar contexts. See Bound et al. (1995), Angrist and Krueger (1995), and Stock and Staiger (1997).

This linearity assumption would follow from the joint normality assumptions typically used in the literature (see e.g. Taber 2001), but only linearity of the conditional mean is required here. In Section 6 I show that this assumption can be relaxed. As long as the conditional means of a_{ic} and b_{ic} are known functions of S_{ic} (linear or not), δ_t and Ψ_t are identified. To proceed, substitute equations (2.1) and (2.2) into the log wage function (1):

$$\log y_{ict} = \Psi_t[a_c + \lambda_1^c(S_{ic} - \overline{S}_c)] + \delta_t[b_c + \lambda_2^c(S_{ic} - \overline{S}_c)]S_{ic} - \delta_t k_c S_{ic}^2 + v_{ict} \tag{4}$$

where $v_{ict} = \varepsilon_{ict} + \Psi_t u_{1ic} + \delta_t u_{2ic} S_{ic}$ is an heteroskedastic error term.[25] Therefore under the assumption that the conditional means of a_{ic} and b_{ic} are linear functions of years of education, the regression function relating log wages to years of education is quadratic:

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

After re-arranging:

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

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

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

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

In a series of related papers, Garen (1984), Heckman and Vytlacil (1998), Card (2000) and Wooldridge (1997, 2003) have shown that with a valid instrumental variable, consistent estimates of all the parameters of this model 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 wages, thus purging the relationship between schooling and wages of any ability and

[25] The equation for v_{ict} shows that residual wage dispersion and returns to unobserved ability are not mechanically linked in this two-factor model.

comparative advantage bias.

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 log wage 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. A simple intuition regarding the source of identifying variation for Ψ_t and δ_t arise from equations (5.1) and (5.3) above. Variation over time in the quadratic term $\pi_2(c,t)$ identifies the year-specific causal return to education δ_t up to a normalization (δ_t=1 in 1979). Similarly, variation over time in the intercept term $\pi_0(c,t)$ identifies the return to unobserved ability Ψ_t up to a normalization (Ψ_t=1 in 1979). Figure 5 shows the time-series variation in these components. Since the purpose of the paper is to identify the change in the causal effect of education over time, the normalizations on δ_t and Ψ_t are innocuous.

While the time-series variation in $\pi_1(c,t)$ is not necessary to identify δ_t and Ψ_t, I discuss below a model which exploits this variation as well. As a by-product of this estimation, estimates of $(a_c - \lambda_1^c \overline{S}_c)$ and $(\lambda_2^c - k_c)$ are also obtained. It is worth emphasizing that the identification of δ_t and Ψ_t is not driven by the assumption of linearity of $E[a_{ic}|S_{ic}]$ and $E[b_{ic}|S_{ic}]$. For example, a quadratic model for the conditional expectations of a_{ic} and b_{ic} would imply a cubic regression function for log wages and schooling. In that case, the causal return to education, δ_t, is identified off the time-series variation in the cubic schooling coefficient in the log wages regression. I explore this possibility in Section 6.

B. *Estimation of the augmented wage regressions.*

First, the following series of log wage regressions is estimated from the repeated cross-sectional data in the ORG between 1979 and 2002:

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

In these regressions $\log y_{ict}$ is the log real hourly wage rate (in 2002 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. In all models the set of covariates in $g(X_{ict}, \beta_t)$ includes a quartic in potential labor market experience, a race indicator, an Hispanic ethnicity indicator, a marital status indicator, dummies for the census divisions and a dummy for metropolitan areas.[26] I refer to this set of covariates as the "Basic" set of covariates. Estimates of $\pi_0(c,t)$, $\pi_1(c,t)$ and $\pi_2(c,t)$ from these regressions on the basic set of covariates have already been displayed in Figure 5. In some specifications, dummies for single-digit industry and white-collar occupations are also included in the set of covariates.[27] I refer to this set of covariates as the "Full" set of covariates. Finally, in some specifications, the experience profile is allowed to vary by year and by cohort. Otherwise, the effect of each covariates is allowed to vary by year only. For the 137 distinct cohort-year pairs observed in the data, the regressions yield a total of 274 estimated coefficients, $\widehat{\pi}_0(c,t)$ and $\widehat{\pi}_2(c,t)$, which can be used to estimate Ψ_t and δ_t.

C. *Minimum-distance estimation of the parameters.*

Consider equations (5.1) and (5.3). For a given cohort-year pair, the relationship between the actual and the predicted coefficients of the causal model for log wages and schooling can be written as $\widehat{\pi}(c,t) = f(\theta_{ct}) + \eta_{ct}$, where:

$$\widehat{\pi}(c,t) = \begin{bmatrix} \widehat{\pi}_0(c,t) \\ \widehat{\pi}_2(c,t) \end{bmatrix} \quad \text{and} \quad f(\theta_{ct}) = \begin{bmatrix} \Psi_t(a_c - \lambda_1^c \overline{S}_c) \\ \delta_t(\lambda_2^c - k_c) \end{bmatrix} \tag{7}$$

and where η_{ct} is a combination of sampling and specification errors. The parameters of interest are δ_t and Ψ_t. At this stage, the cohort-specific parameters that can be estimated from (7), $(a_c - \lambda_1^c \overline{S}_c)$ and $(\lambda_2^c - k_c)$, are viewed as nuisance parameters. Below I consider a model that interprets both the year-specific and cohort-specific parameters. The optimal minimum distance (OMD) estimator of the parameters of interest is obtained by minimizing the following quadratic form:

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

[26] All regressors are deviated from their year-specific averages. Since the dependent variable is the log hourly wage, each regression is weighted by hours of work (see DiNardo, Fortin and Lemieux 1996).

[27] See Cawley, Heckman and Vytlacil (1999) for a recent analysis of wage determination of occupation choice.

where T is the number of survey years between 1979 and 2002, C is the number of birth cohorts, and V_{ct} is the asymptotic covariance matrix of $\widehat{\pi}_0(c,t)$ 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 estimator. The heteroskedasticity-consistent estimator of White (1981) is used in the empirical analysis. The asymptotic variance of the OMD estimator of θ is $F'V^{-1}F$, where F is the Jacobian of $f(\theta)$. All standard errors reported in Appendix Table 1 are computed using this formula.

D. *Results.*

Figure 6 displays the OMD estimates of the year-specific causal return to education and return to unobserved ability (δ_t and Ψ_t) for every year between 1979 and 2002. The full set of estimates and the corresponding standard errors are also reported in Table 1 in the appendix. Since these estimates only use information on $\widehat{\pi}_0(c,t)$ and $\widehat{\pi}_2(c,t)$, I label them "partial-information" estimates. Estimates corresponding to three different specifications of the log wage regressions are showed. The solid line corresponds to the regressions including the basic set of covariates, the starred line corresponds to the regressions including the full set of covariates, and the squared line corresponds to regressions that include cohort-specific quadratic experience profiles in addition to the full set of covariates.

The estimates show that the year-specific causal return to education (top panel) increased by 21-36% between 1979 and 2002. All the point estimates are statistically significant, with standard errors ranging from 0.03-0.05. Figure 6 also reveals important information about the timing of the change in the year-specific causal return to education: Most of the increase is concentrated in the early part of the 1980s. This result is consistent with the overall trends in educational-based wage differentials in the United States (Katz and Autor 1999).

Panel (B) of Figure 6 shows that the return to unobserved ability increased over the 1980s and 1990s, by 8-13%. The rate of increase is not uniform over the whole time period. As the top line indicates, the return to unobserved ability increased over the 1980s, was relatively constant over the early 1990s, and then increased again from the late 1990s onwards. This pattern is consistent with the casual observations about the changes in the U.S. wage structure. For example, most of the increase in residual wage dispersion between 1970 and 2002 can be accounted for by increases in the early 1980s (Card and DiNardo 2002, Lemieux 2003).

However, this figure makes clear that the observed rise in educational wage differentials cannot be solely attributed to increases in the return to unobserved ability: The rise in the return of unobserved ability explains at most 20-25% of the rise in the conventional measure of the return to education.[28] This is in sharp contrast with the results of Taber (2001) who finds no evidence of an increasing causal effect of schooling over the 1980s using a dynamic programming selection model. It should noted that the partial-information estimates only pertain to economy-wide changes in the educational wage structure. As the model above illustrates, changes in the educational wage structure can also be caused by changes in ability-education mapping across cohorts. The next section investigates this possibility.

5. Identification and Estimation of the Cohort-Specific Parameters.

A. *Identification.*

Unlike the year-specific parameters, the cohort-specific parameters are not all freely identified without additional assumptions. This shortcoming is a drawback of using 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 are identified. Other analysts have relied on distributional and functional form assumptions to resolve similar identification problems (see e.g. Taber 2001, Belzil and Hansen 2002). Moreover, the models previously considered are typically special cases of the model in Section 4 (i.e. assuming homogeneous returns to schooling or no differential ability-education mappings across cohorts), which in turns simplifies the identification problem.

To illustrate the identification issues in the present context, consider the model in equations (5.1)-(5.3). As shown earlier, Ψ_t and δ_t can be identified from the time-series variation in $\pi_0(c,t)$

[28] The convential estimate of the return to education (i.e. from a Mincerian regression) increased from 62% between 1979 and 2002.

and $\pi_2(c,t)$. Similarly, the time-series variation in all three components $\pi_0(c,t)$, $\pi_1(c,t)$ and $\pi_2(c,t)$ can be used to estimate Ψ_t and δ_t in addition to the cohort-specific parameters.

The identifying variation for the cohort-specific parameters is generated by variation across cohorts in the levels of regression coefficients. As illustrated by equations (5.1)-(5.3), 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 $\pi_2(c,t)$. However, the number of cohort-specific parameters is $5 \times C$ (that is, a_c, b_c, λ_1^c, λ_2^c and k_c), where C is the number of birth cohorts. Clearly, not all of these parameters can be freely identified: Only $3 \times C$ cohort-specific parameters can be identified.

The approach I follow is guided by the objective of this study: assessing inter-temporal changes in the causal effect of education by decomposing the observed relationship between log wages and schooling into time-varying average causal effect, time-varying ability bias and time-varying comparative advantage bias. First, I follow most of the literature by considering the case where the structural earnings function is linear, i.e. that $k_c = 0$ (see e.g. Ashenfelter and Rouse (1998), and Heckman and Vytlacil (1998)). Under this assumption, the parameters capturing comparative advantage selection (λ_2^c) are readily estimable from the levels of the quadratic schooling coefficients across cohorts. Since economic theory predicts that $k_c \geqslant 0$, because of diminishing returns to schooling in the production of earnings capacity, the estimates derived under the assumption $k_c = 0$ will provide a lower bound on the degree of comparative advantage bias across cohorts.[29] This reduces the number of parameters by C, but an additional C restrictions must be imposed. Provided that an estimate of λ_1^c is available, the average marginal productivity of schooling for cohort c, b_c, is 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. Since this study focuses on inter-temporal changes in the causal effect of education, only differences in b_c, λ_1^c and λ_2^c across cohorts, combined with the year-specific parameters δ_t and Ψ_t are required. One normalization consistent with this objective is to impose that $\sum_c \lambda_1^c = 0$, and assume no differences across

[29] To see this, note that from equation (5.3), $\lambda_2^c = \delta_t^{-1}\pi_2(c,t) + k_c$. Thus λ_2^c will increase linearly with the curvature parameter k_c. Consequently, under the assumption that the average schooling level for a cohort remains constant as k_c changes, the degree of comparative advantage bias will rise as k_c rises. This is illustrated in Figure 8.

cohorts in the average level of absolute ability (i.e. that a_c=a for all cohorts).[30] For this reason, I refer to the estimates of λ_1^c as "relative ability bias". To summarize the following assumptions are sufficient to identify inter-temporal <u>changes</u> in the educational wage structure:

A1: $E[a_{ic}|S_{ic}] = a_c + \lambda_1^c(S_{ic} - \overline{S}_c)$ and $E[b_{ic}|S_{ic}] = b_c + \lambda_2^c(S_{ic} - \overline{S}_c)$

A.2: $k_c = 0$

A.3: $\sum_c \lambda_1^c = 0$

A.4: $a_c = a$

Three final points are worth emphasizing. First, identification does not hinge on the linearity of the conditional means of the unobservables a_{ic} and b_{ic}. Departures from linearity will be considered in Section 6. None of the results presented in the paper are affected by this linearity assumption. Second, several specifications that have been considered in the literature arise as special case of the model developed here (e.g. constant ability-education mapping across cohorts). Consequently, goodness-of-fit tests for these alternative specifications can be readily implemented. Finally, identification of the year-specific causal return to education and ability, δ_t and Ψ_t, does not require A.2-A.4, as shown in the previous section.

B. *Minimum-distance estimation of the parameters.*

I now discuss how the cohort-specific and year-specific parameters can be estimated from information on $\pi_0(c,t)$, $\pi_1(c,t)$ and $\pi_2(c,t)$. The regression coefficients used for minimum-distance estimation are obtained from the regression models described in Section 4.B. Consider equations (5.1)-(5.3). For a given cohort-year pair, the relationship between the actual and the predicted coefficients of the causal model for log wages and schooling can be written as $\widehat{\pi}(c,t) = f(\theta_{ct}) + \eta_{ct}$, where:

$$\widehat{\pi}(c,t) = \begin{bmatrix} \widehat{\pi}_0(c,t) \\ \widehat{\pi}_1(c,t) \\ \widehat{\pi}_2(c,t) \end{bmatrix} \quad \text{and} \quad f(\theta_{ct}) = \begin{bmatrix} \Psi_t(a - \lambda_1^c \overline{S}_c) \\ \delta_t b_c + \Psi_t \lambda_1^c - \delta_t \lambda_2^c \overline{S}_c \\ \delta_t \lambda_2^c \end{bmatrix} \tag{8}$$

and where η_{ct} is a combination of sampling and specification errors. By imposing that $\sum_c \lambda_1^c = 0$,

[30] This last assumption is not a major concern since a_c does not load into the schooling coefficients of the earnings function.

the optimal minimum distance (OMD) estimator of the parameters of interest is readily obtained by minimizing the following quadratic form:

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

where T is the number of survey years between 1979 and 2002, C is the number of birth cohorts, and V_{ct} is the heteroskedasticity-consistent estimator of the covariance matrix of $\pi(c,t)$ (White 1981). Again, the asymptotic variance of the OMD estimator of θ is $F'V^{-1}F$, where F is the Jacobian of $f(\theta)$. All standard errors are computed using this formula.

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 (see e.g. Altonji and Segal 1996). In the context of OMD estimators, small sample bias may arise in the presence of a correlation between the sampling errors in the elements of the weighting matrix and the sampling errors in the vector of sample moments. 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 $\widehat{\pi}$ and $V(\widehat{\pi})$. In the present application, IWOMD and OMD estimates were remarkably similar.[31]

C. *Basic results.*

Figure 7 displays the OMD estimates of the year-specific causal return to education and return to unobserved ability (δ_t and Ψ_t) for every year between 1979 and 2002. The full set of estimates and the corresponding standard errors are also reported in Table 2 in the appendix. Since these estimates of δ_t and Ψ_t use information on $\widehat{\pi}_0(c,t)$, $\widehat{\pi}_1(c,t)$ and $\widehat{\pi}_2(c,t)$, I label them "full-information" estimates. Again, the estimates corresponding to three different specifications of the log wage regressions are showed. The solid line corresponds to the regressions including the basic set of covariates, the starred line corresponds to the regressions including the full set of covariates, and the squared line corresponds to the regressions with the full set of covariates, with cohort-specific

[31] These results are available from the author upon request.

quadratic experience profiles.

In general, the full-information estimates of δ_t and Ψ_t are very similar to the partial-information estimates displayed in Figure 6. Again, these show that the year-specific causal return to education (top panel) increased by 22-39%, while the return to unobserved ability (bottom panel) increased by 7-13% between 1979 and 2002. Any difference in the patterns in Figure 6 and 7 can be attributed to the fact that the estimates in Figure 7 are derived using the time-series variation in the linear schooling component ($\pi_1(c,t)$) as well.

The estimates of the cohort-specific parameters are reported in Table 2. In addition, the goodness-of-fit statistics associated with each model are reported. Each panel contains a separate series of estimates, corresponding to the three different specifications of the log wage regressions. In each panel, the first column shows the average marginal productivity of schooling (b_c), the second column displays the relative ability bias (λ_1^c), while the third column presents the comparative advantage bias ($\lambda_2^c\overline{S}_c$). Across the different specifications, the estimates of the marginal productivity of schooling indicate that an additional year of education permanently increase log hourly wages by 2.3% to 4.4%. All these parameters are precisely estimated, with standard errors ranging from 0.002-0.004. At the same time, the entries show a limited range of variation in b_c across cohorts. Depending on which specification is chosen, the maximal difference in b_c for two cohorts ranges from 0.01-0.02. A direct consequence of this is that any increase in the causal effect of education over time will have to be driven by a rise in the year-specific causal return to schooling (δ_t).

The estimates of the relative ability bias (λ_1^c) indicate no inter-cohort trend in the correlation between unobserved ability and completed education. Again, because of the normalization the levels of the estimates is not meaningful, only the differences across cohorts are. The entries of Table 2 show that across the different specifications there is no systematic pattern for the differences across cohorts. The differences are small, and typically do not exceed their standard errors. Therefore there is no evidence in Table 2 suggesting that the rise in the conventional measure of the return to education is due to a higher correlation between absolute ability and schooling among younger cohorts. This evidence contradicts the interpretation of Herrnstein and Murray (1994) who attribute part of the increase in the conventional measure of the return to education to the increased correlation between ability and education that would result from an increasing application of meritocratic principles in school admissions.

Estimates of cohort-specific comparative advantage bias are constructed by multiplying λ_2^c, the slope in the relationship between person-specific returns to education and years of completed education specific to each cohort, and $\overline{S}_c$, the average schooling for cohort c. As shown by equation (3), this amounts to the contribution of comparative advantage bias in the conventional measure of the return to schooling. Across all specifications 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 cross-sectional 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 2 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.[32] In fact, for the cohorts born in the mid-1950s, comparative advantage bias explains approximately 40% of the conventional measure of the return to schooling.[33] Finally, I discuss the implications of relaxing the assumption of linearity for the structural earnings function. As argued earlier, the estimates of comparative advantage bias derived under this assumption provide a lower bound on the level of comparative advantage bias. Figure 8 shows the estimates corresponding to various degrees of curvature in the earnings function. Since economic theory predicts that $k_c \geqslant 0$–or in other words that the structural earnings function is concave in years of education–I consider values ranging from 0 to 0.005. The results in Figure 8 show that the level of comparative advantage bias increases as k_c increases.[34] The ranking

[32] 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 1 showed, average educational attainment actually decreased for cohorts born after 1950.

[33] This figure is obtained by comparing the comparative advantage bias for a cohort, $\lambda_2^c\overline{S}_c$, to the conventional measure of the return to schooling, $\beta_{ols}(c,t)$ (see equation (3)) , averaged over all years in which a cohort is observed in the sample.

[34] This interpretation is correct under the assumption that the average schooling level of each cohort is not affected

across cohorts is preserved here since I consider the same curvature parameter for all cohorts. In principle this parameter could vary across cohorts, but must remains constant over time for a given cohort. Figure 8 clearly shows that the estimates reported in Table 2 provide lower bounds on the contribution of comparative advantage bias in the conventional measures of the return to education.

D. *Censoring in the CPS data.*

All measures of earnings are censored in the Current Population Surveys. In the ORG files, the wage rate of workers paid by the hour is topcoded at $99.99 throughout the sample period, and less than 0.5% of the hourly wage rates are topcoded. As explained in Section 2, weekly earnings are used to construct a measure of hourly wage rate for workers not paid on an hourly basis. The topcodes for weekly earnings changed over the sample period, and the summary statistics in Deschênes (2001) indicate that the fraction of weekly earnings censored is the highest between 1985 and 1987.[35] Since log wages and schooling are positively correlated, censoring of the log wage variable will likely result in underestimation of the effect of schooling on wages. However, since 50% of the sample is paid on an hourly basis, which entails little or no incidence of censoring, this is unlikely to cause major problems in this application. As shown in the first column of Table 4 in the appendix, on average only 3% of the observations have a topcoded hourly wage rate. The single year with the highest fraction of topcoded hourly wage is 1988, where 9% of the hourly wage rates are topcoded.

To account for the censoring problem, the regression coefficients in (6) were also estimated Powell's (1984) censored least absolute deviation (CLAD) semiparametric estimator. This estimator was chosen since it is robust to departures from the normality and heteroskedastcity assumptions that would be required for maximum-likelihood estimation.[36] Tables 3 and 4 report the OMD estimates of the parameters of the causal model for earnings obtained when the log wage regres-

by changes in k_c. Schooling choice models suggest that holding everything else constant, individual schooling levels will decline as k_c increase (as the marginal benefit to schooling declines faster). However, without information on the distribution of marginal benefits and marginal costs across individuals and cohorts, it is impossible to assess how the average education level of each cohort will be altered by changes in k_c.

[35] 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.

[36] For example, it is well-known that the Tobit estimator is not robust to departures from normality and to heteroskedasticity (see e.g. Goldberger 1983).

sions are estimated via CLAD. For brevity, only the "full-information" estimates are reported. Table 3 shows the estimates of δ_t and Ψ_t, as well as the proportion of censored observations in each year. An examination of the entries in Table 3 reveals that the estimates of the year-specific causal return to education and to return to unobserved ability are very similar to those obtained when the regression are estimated by OLS. Given that only 3% of the observations have topcoded hourly wages, this result is not surprising. The estimates of the cohort-specific parameters, reported in Table 4, are also very similar to those reported in Table 2. Therefore, the evidence from the CLAD estimates shows that none of the results reported in the previous section are seriously affected by the topcode of hourly wages in the CPS data.

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

Table 5 displays the goodness-of-fit statistics associated with alternative models that have been considered in the literature studying the changes in the wage structure. In column 1 are the OMD goodness-of-fit statistics obtained when the regression coefficients in (6) are estimated by OLS, while column 2 corresponds to the CLAD regression estimates. In both cases, the basic set of covariates are included in the regressions. None of the results are changed when other specifications of the log wage regressions are considered. For convenience, the goodness-of-fit statistics associated with the two-factor model presented in Tables 2 and 4 are reported in row 1. The other rows of Table 5 test further simplifications against the two-factor model alternative. In each row, the difference between the goodness-of-fit statistics of a restricted model and the goodness-of-fit statistic of the "unrestricted" model in row 1 is reported, as well as the corresponding degrees of freedom and p-values.[37]

Row 2 tests the restriction that the average marginal productivity of schooling is the same for all cohorts. The equality of the b_c across cohorts is rejected at the conventional level, as indicated by the low p-values. This implies that any change in the causal effect of education over the 1980s and 1990s must be generated by a change in the year-specific causal return to education, δ_t. The third row tests a model imposing no variation across cohorts in the mapping between educational

[37] 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).

attainment and person-specific returns education. Under this restriction, the degree of comparative advantage bias does not vary across cohorts. Not surprisingly, this restriction is easily rejected by the data. Rows 4 and 5 present specification tests of models with no ability bias ($\lambda_1^c = 0$) and no comparative advantage bias ($\lambda_2^c = 0$). Clearly, the evidence in Table 5 rejects both models. This is particularly important since the assumption of homogeneous returns to schooling has been used in most cross-sectional analyses of the causal effect of education (see e.g. Ashenfelter and Rouse, 1998). The large chi-square statistics suggests that this restriction is not supported by the data considered in this study. The sixth and seventh rows test models with stationarity restrictions imposed on the year-specific parameters. Row 6 shows that the hypothesis of a time-invariant year-specific causal return to education (while allowing the return to unobserved ability to vary) is rejected. Similarly, in row 7, the stationarity of the return to unobserved ability is also rejected. Finally, the last row displays the statistical fit of a model with a single return to skill, in which the year-specific return to education and the return to unobserved ability are restricted to be equal in each time period (i.e. that $\delta_t = \Psi_t$). Variants of this model have been used in several studies of the wage structure (see e.g. Juhn, Murphy and Pierce 1993, Card and Lemieux 1996). In this case, the restriction of a single price of skill is rejected.

To summarize, the evidence in Table 5 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 stationarity of the causal return to education and return to unobserved ability appear to be greatly misspecified.

F. *Interpretation.*

Taken as a whole, the evidence in Figures 6-7 and Tables 2-5 points to important conclusions regarding the changing educational wage structure in the U.S. over the 1980s and 1990s. First, most of the increase in educational-based wage differentials is attributable to an increase in year-specific causal return to education (δ_t), and not to an increase in the return to unobserved ability (Ψ_t). This is shown graphically in the top panel of Figure 9, which displays the conventional estimate of the return to education, from linear log wage regressions for 1979-2002. In addition estimates of δ_t and Ψ_t from regressions using the basic set of covariates are also reported. For comparisons purposes, all values were normalized to the 1979 values.

Based on this figure, it clear that the rise in the return to unobserved ability contributed to the the widening of the education wage differentials, but the scope of its contribution is limited. As such, the rise in the causal return to education explains 60% of the observed rise in the Mincerian estimate of the return to schooling. These results undermine the contention that the rising return to unobserved ability is the major driving force behind the observed changes in the educational wage structure over the 1980s and 1990s.

Second, the connection between the estimated return to unobserved ability and the patterns of residual wage dispersion is showed in the bottom panel. This figure shows the standard deviation of residuals from log hourly wage regressions estimated separately for each year, and normalized to its 1979 value, along with estimates of Ψ_t from the model. The similarly of the trends in the two series is remarkable. The evidence confirms the role of a rise in the return to unobserved ability for the in explaining the increase in residual wage dispersion, as argued by Juhn, Murphy and Pierce (1993) and others. Moreover, the similarity of the trends in Figure 9(B) reinforces the credibility of the estimates of return to unobserved ability reported in this paper.

6. Sensitivity Analysis.

This section investigates the robustness of the findings in Section 5 to alternatives assumptions. Particular attention is devoted to: (A) measurement error in reported schooling; and (B) linearity of the conditional expectation of the unobservables.

A. *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.[38] In nonlinear models, however, the effects of measurement error are not as straightforward to predict and often must be evaluated on a case-by-case basis.[39] 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 3 in the appendix 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 3 in the appendix 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 Table 2. As the entries indicate, the year-specific causal return to education δ_t appears to be downward biased by measurement error in reported schooling, but none of the other parameters are seriously affected. 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 this paper should be viewed as lower bounds. The rest of the evidence in Appendix Table 3 shows no indication of serious measurement error bias in the estimates reported in Section 5.

B. *Linearity of the conditional expectation of the unobservables.*

Sections 4 and 5 show that under the assumption of linearity for the conditional means of the

[38] See for example Table 9 in Angrist and Krueger (1999).

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

unobservables, the year-specific parameters δ_t and Ψ_t can be identified from repeated observations on coefficients from cross-sectional regressions with quadratic terms in years of education. This result holds true even after allowing non-linearity in the structural earnings function (see Section 4). However, it hinges on a linear relationship between the unobservables a_{ic} and b_{ic} and years of education. As noted above, this assumption has appeared in the previous literature via joint normality or other distributional assumptions.

Under the assumptions maintained in Section 5, estimates of the year-specific parameters are readily obtained when the conditional means of the unobservables are known non-linear functions of the endogenous variable schooling. Clearly, identification of a model with two unobserved ability factors that are correlated with the endogenous variable is not possible without knowledge of the functional form of the relationship between the unobservables and the endogenous variable. To proceed, suppose that the conditional expectations of the unobserved ability factors are quadratic in schooling:

$$E[a_{ic}|S_{ic}] = a_c + \lambda_1^c(S_{ic} - \overline{S_c}) + \gamma_1^c(S_{ic}^2 - \overline{S_c^2}) \tag{2.1”}$$

$$E[b_{ic}|S_{ic}] = b_c + \lambda_2^c(S_{ic} - \overline{S_c}) + \gamma_2^c(S_{ic}^2 - \overline{S_c^2}) \tag{2.2”}$$

Allowing for a non-linear relationship between the unobservables and schooling is potentially important since selection bias may be nonlinear around the mass points of the education distribution. Substituting these equations in (1), we obtain the following cubic log wage regression function:

$$E[\log y_{ict}|S_{ic}] = \pi_0(c,t) + \pi_1(c,t)S_{ic} + \pi_2(c,t)S_{ic}^2 + \pi_3(c,t)S_{ic}^3$$

$$\pi_0(c,t) = \Psi_t(a_c - \lambda_1^c\overline{S}_c - \gamma_1^c\overline{S}_c)$$

$$\pi_1(c,t) = \delta_t b_c + \Psi_t\lambda_1^c - \Psi_t\gamma_2^c\overline{S_c^2} - \delta_t\lambda_2^c\overline{S}_c$$

$$\pi_2(c,t) = \Psi_t\gamma_1^c + \delta_t(\lambda_2^c - k_c)$$

$$\pi_3(c,t) = \delta_t\gamma_2^c$$

With a normalization on λ_1^c and γ_1^c (as described in Section 5), the parameters of this model can be estimated from the coefficients $\pi_0(c,t)$, $\pi_1(c,t)$, $\pi_2(c,t)$, and $\pi_3(c,t)$ using OMD. Consequently, parameter estimates that are robust to nonlinear relationship between the unobservables and schooling are readily obtained using the same two-step procedure outlined in Section 5.

Figure 10 displays the estimated year-specific parameters from the nonlinear model, when the log wage regressions include the basic set of covariates.[40] For comparison, the corresponding estimates from Figure 7, derived under the assumption of a linear relationship between the unobservables a_{ic} and b_{ic} and schooling are also reported. The results are remarkably similar. It is apparent from the evidence that our conclusions regarding the year-specific causal return to education and the return to unobserved ability remain unchanged when more flexible models for the unobservables are considered.

7. Conclusions.

This paper analyzes the educational wage structure in the United States between 1979 and 2002. The main purpose is to assess the change in the 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 and interpret 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, 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 mostly 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 a differential rate of return for unobserved ability, the estimates indicate that the causal return to education increased by about 20-40% between 1979 and 2002. By comparison, the conventional measure of the return to education increased by about 60%. The

[40] The estimated cohort-specific parameters are available from the author upon request.

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 a year-specific factor common to all cohorts as opposed to cohort-specific factors.

As others have argued before, an increase in the return to unobserved ability also contributed to the widening of the educational wage differentials. However, my estimates indicate that the changes in the return to unobserved ability do not provide a single explanation for the changes in the educational wage structure over the 1980s and 1990s. Using a wide range of specifications, I find that the return to unobserved ability increased by 7-13% between 1979 and 2002, thereby explaining at most 20% of the observed increase in the cross-sectional estimate of the return to schooling. My estimates also imply that there are no differences across cohorts in the mapping between absolute ability and schooling. Therefore, my results provide no evidence for the claim that the observed changes in the educational wage structure are attributable to confounding due to time-varying absolute ability biases.

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—that have dominated the empirical literature—are greatly misspecified. 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 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.

Belzil, Christian and Jorgen Hansen (2002): "Unobserved Ability and the Return to Schooling," Econometrica, 70: 575-591.

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

Cameron, Steve and Christopher Taber (2000): "Borrowing Constrainsts and the Returns to Schooling," NBER Working Paper No. 7761.

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

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.

Card, David and Thomas Lemieux (2001): "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *Quarterly Journal of Economics*, 116: 705-746.

Card, David and John DiNardo (2002): "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles," *Journal of Labor Economics* 20: 733-783.

Cawley, John, James Heckman and Edward Vytlacil (1999): "Meritocracy in America: Wages Within and Across Occupations," Industrial Relations, 38: 250-296

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

Chay, Kenneth, Y. and David S. Lee (2000): "Changes in Relative Wages in the 1980s: Returns to Observed and Unobserved Skills and Black-White Wage Differentials," *Journal of Econometrics* 99: 1-38.

Deschênes, 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.

DiNardo, John, Nicole Fortin and Thomas Lemieux (1996): "Labor Market Institutions and the Distribution of Wages 1973-1992: A Semi-Parametric Approach," *Econometrica* 64: 1001-1044.

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.

Gossling, Amanda, Stephen Machin, and Costas Meghir (2000): "The Changing Distribution of Males Wages in the U.K.," *Review of Economic Studies* 67: 635-666.

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.

Hausman, Jerry, A., W. Taylor (1981): "Panel Data and Unobservable Individual Effects," *Econometrica* 49: 1377-1398.

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 (2001): "Identifying the Role of Cognitive Ability in Explaining the Level of and Change in the Return to Schooling," *Review of Economics and Statistics*, 83: 1-12.

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.

Herrnstein, Richard J., and Charles Murray (1994): *The Bell Curve*, New York: Free Press.

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.

Lemieux, Thomas (2003): "Composition Effects, Wage Measurement, and the Growth in Within-Group Wage Inequality," Working Paper, University British Columbia.

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.

Staiger, Douglas and James H. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica* 65: 557-586.

Taber, Christopher (2001): "The Rising College Premium in the Eighties: Return to College or Return to Unobserved Ability," *Review of Economic Studies* 68: 665-691.

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. (1997): "On Two Stage Least Squares Estimation of the Average Treatment Effect in a Random Coefficient Model," *Economic Letters* 56: 129-133.

Woolridge, Jeffrey M. (2003): "Further Results on Instrumental Variables Estimation of Average Treatment Effects in the Correlated Random Coefficient Model," *Economic Letters* 79: 185-191.

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}).

Appendix B. Derivation of Equation (3).

First, note that the slope coefficient in the linear projection of S_i^2 on S_i is $2\mu + E[(S_i - \mu)^3]/Var[S_i]$. This follows from the fact that $Cov(S_i^2, S_i) = E[(S_i - \mu)^3] + 2\mu V(S_i)$. Therefore, the slope coefficient is:

$$\frac{Cov(S_i^2, S_i)}{Var(S_i)} = 2\mu + \frac{E[(S_i - \mu)^3]}{Var[S_i]}$$

Now consider the model in (1) and (2) for a single cohort and a single cross-section of data:. The model for log earnings is:

$$\begin{aligned} &\log y_i = a_i + b_i S_i + \varepsilon_i \\ &a_i = a + \lambda_1(S_i - \overline{S}) + u_{1i} \\ &b_i = b + \lambda_2(S_i - \overline{S}) + u_{2i} \end{aligned}$$

where ε_i is uncorrelated with S_i. With the assumption of linearity of conditional expectations of a_i and b_i, u_{1i} and u_{2i} will satisfy $E[u_{1i}|S_i] = 0$ and $E[u_{2i}|S_i] = 0$. Substituting in the earnings equation:

$$\log y_i = \{a + \lambda_1(S_i - \overline{S}) + u_{1i}\} + \{b + \lambda_2(S_i - \overline{S}) + u_{2i}\}S_i + \varepsilon_i$$

The OLS estimate of the education slope in a regression of $\log y_i$ on S_i (the conventional measure of the return to education) is given by:

$$b_{ols} = \frac{Cov(\log y_i, S_i)}{Var(S_i)} = \frac{Cov(a_i, S_i)}{Var(S_i)} + \frac{Cov(b_i S_i, S_i)}{Var(S_i)}$$

The two components of b_{ols} are given by:

$$\frac{Cov(a_i, S_i)}{Var(S_i)} = \lambda_1$$

$$\frac{Cov(b_i S_i, S_i)}{Var(S_i)} = b - \lambda_2 \overline{S} + \lambda_2 \frac{Cov(S_i^2, S_i)}{Var(S_i)} + \frac{Cov(u_{2i} S_i, S_i)}{Var(S_i)}$$

Note that $E[u_{2i} S_i^2] = 0$ because u_{2i} is mean-independent of S_i. Therefore the OLS estimate of the education coefficient is given by:

$$b_{ols} = b + \lambda_1 + \lambda_2 \overline{S} + \lambda_2 \frac{E[(S_i - \overline{S})^3]}{V(S_i)}$$

Moreover, under the assumption that the third central moment of S_i is 0 (i.e. that $E[(S_i - S)^3] = 0$), the projection of S_i^2 on S_i has a slope equal to $2\overline{S}$. It then follows that:

$$b_{ols} = b + \lambda_1 + \lambda_2 \overline{S}$$

FIGURE 1: AVERAGE EDUCATIONAL ATTAINMENT, BY BIRTH COHORT

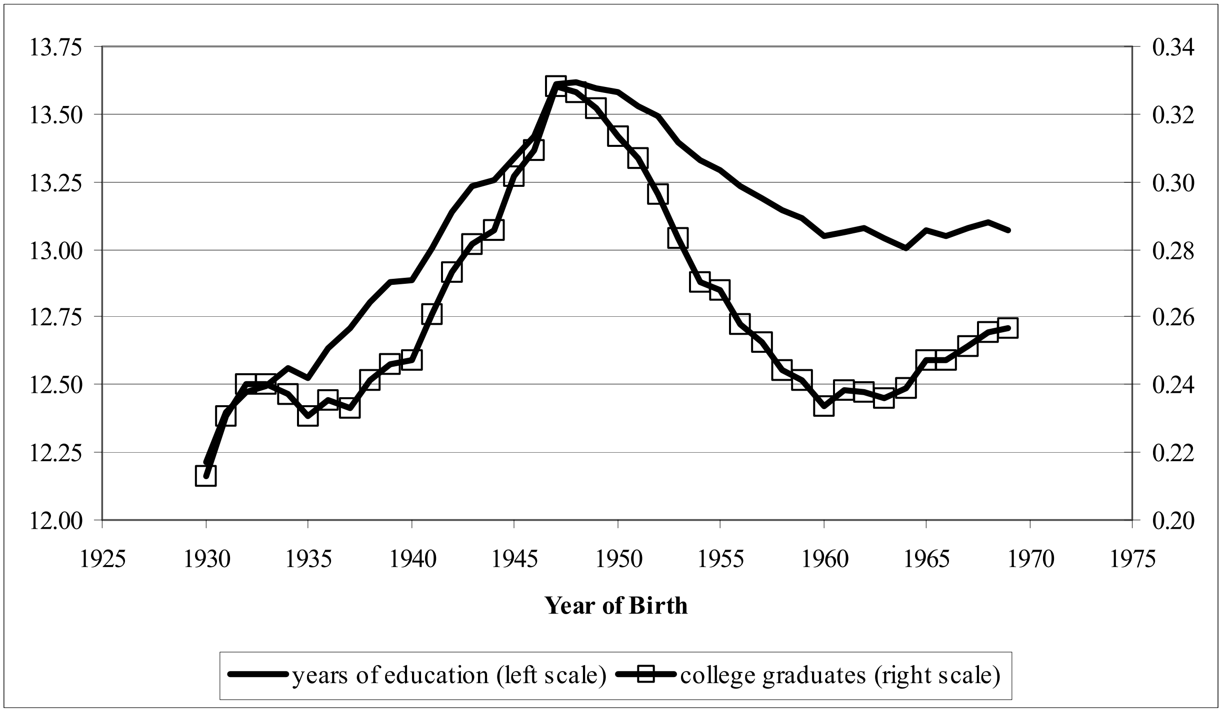


FIGURE 2: AVERAGE EDUCATIONAL ATTAINMENT, BY YEAR 1979-2002

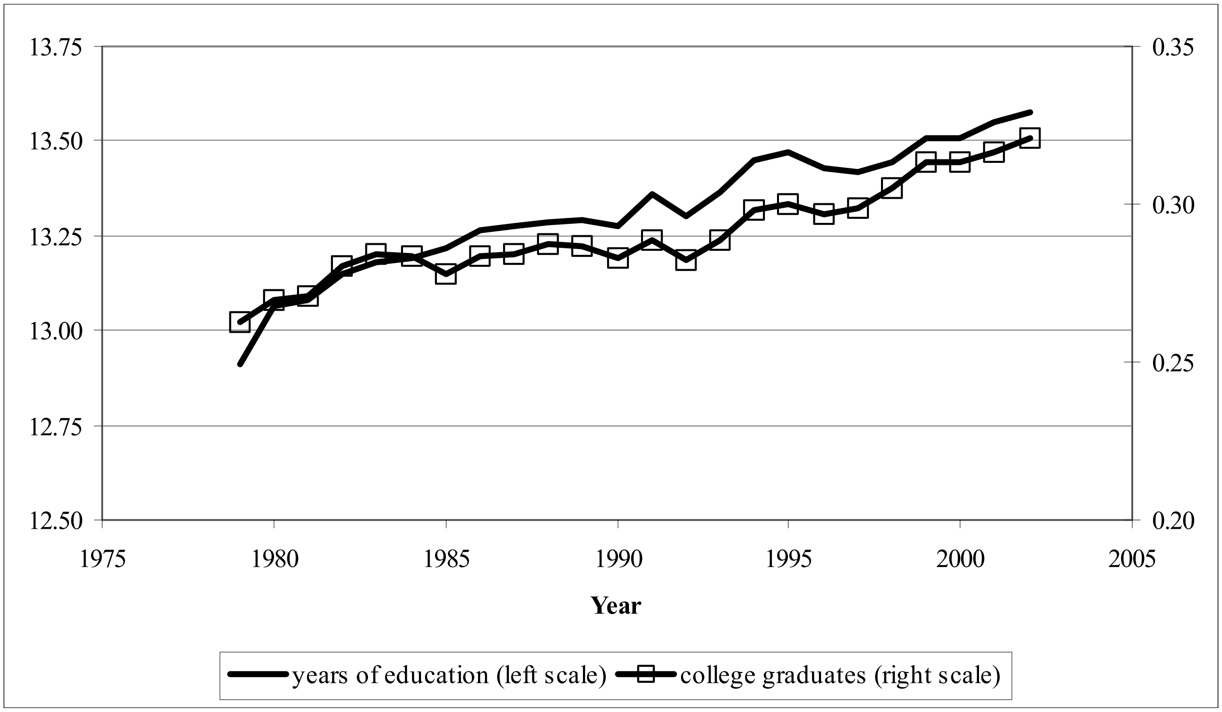


FIGURE 3: ADJUSTED MEAN LOG HOURLY WAGE BY YEAR OF EDUCATION, 1979-2002
[Relative to mean log hourly wage of high school graduates]

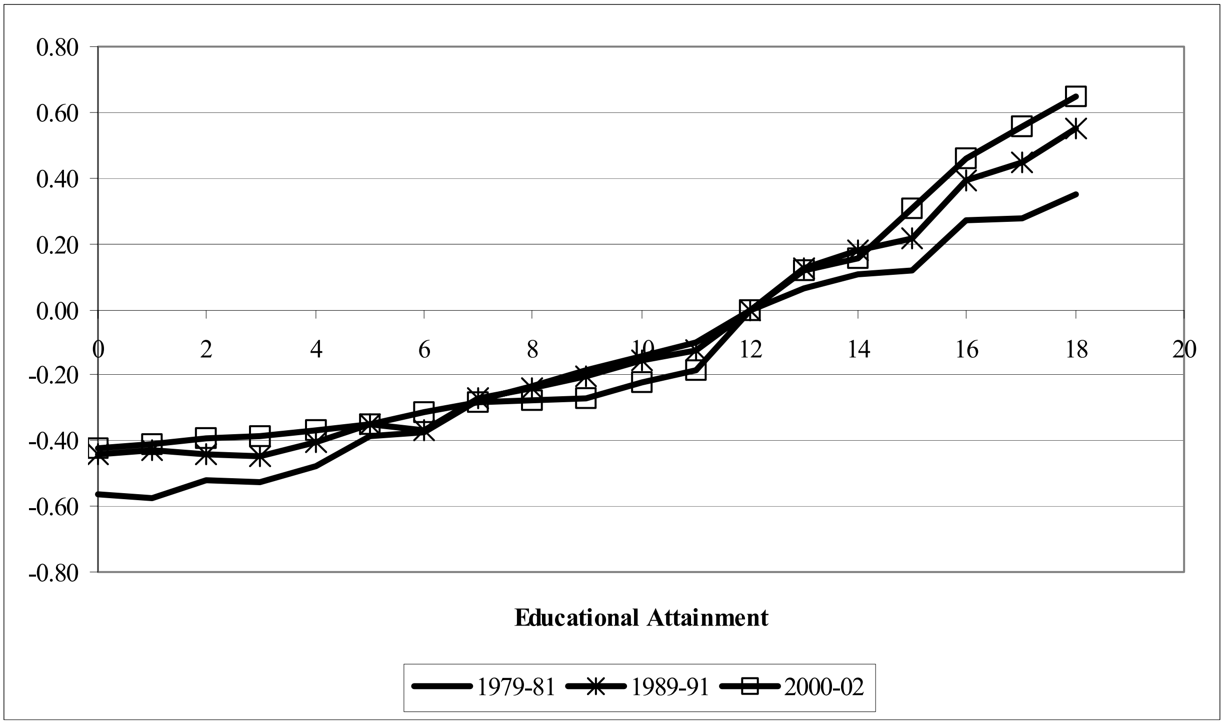
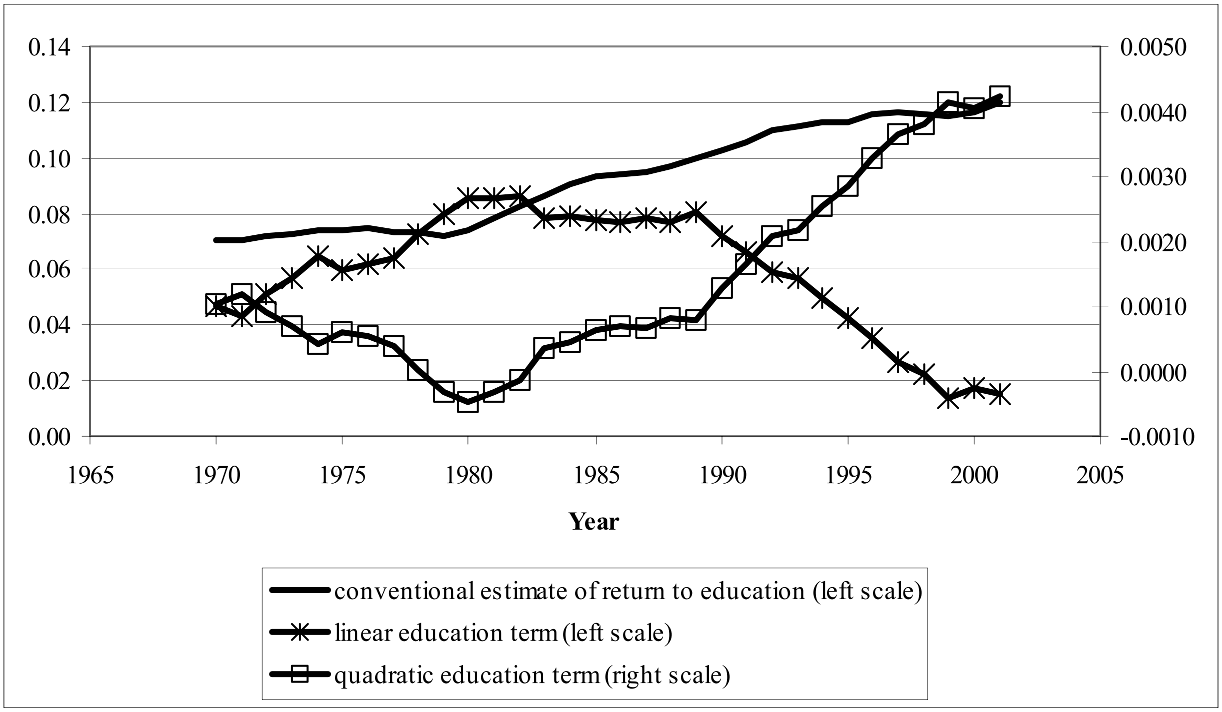


FIGURE 4: CHANGES IN THE EARNINGS-SCHOOLING RELATIONSHIP, 1970-2002



Data are from the 1971-2003 March CPS Earnings Supplement. See text for more details

Figure 5: Profiles of the Earnings-Schooling Relationship, 1979-2002
(A) Intercept

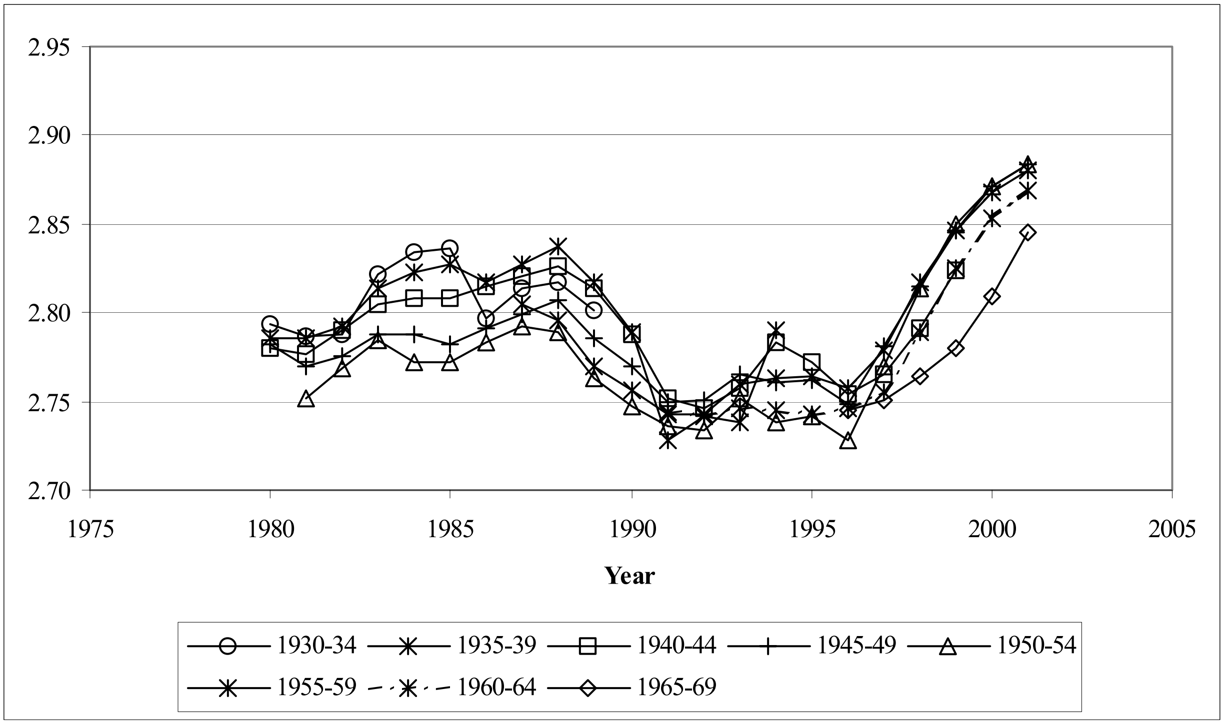


FIGURE 5: PROFILES OF THE EARNINGS-SCHOOLING RELATIONSHIP, 1979-2002
(B) LINEAR SCHOOLING COMPONENT

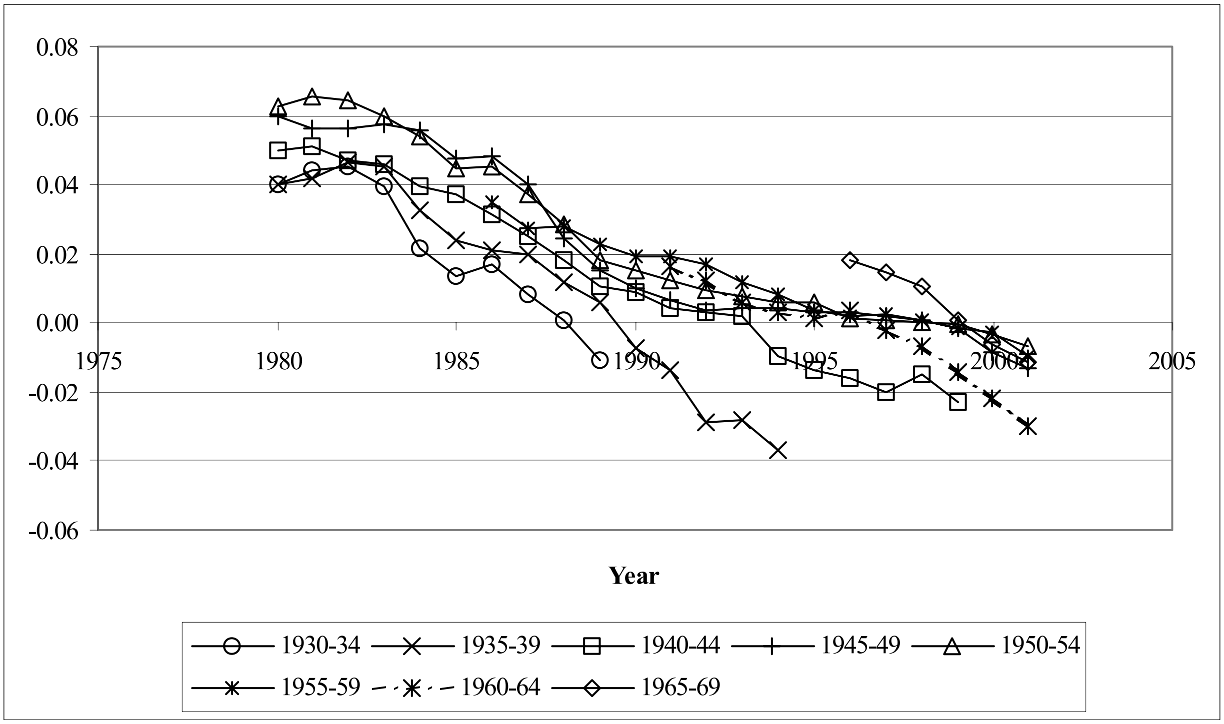


FIGURE 5: PROFILES OF THE EARNINGS-SCHOOLING RELATIONSHIP, 1979-2002
(C) QUADRATIC SCHOOLING COMPONENT

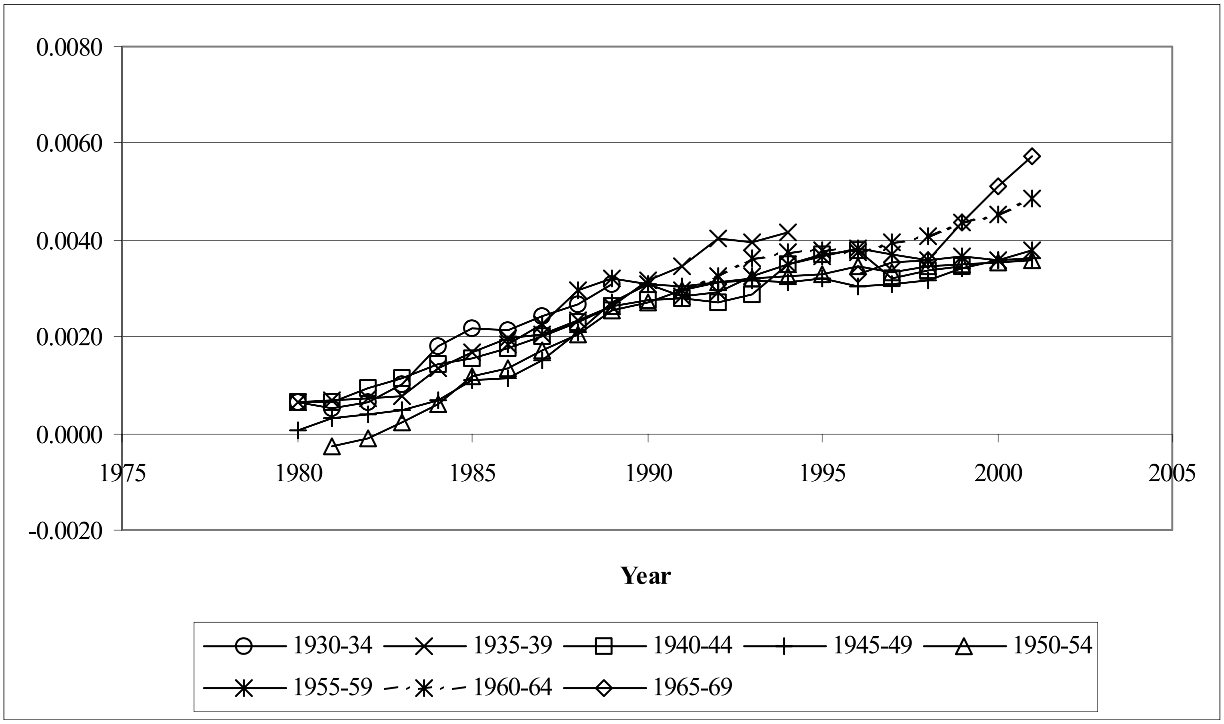
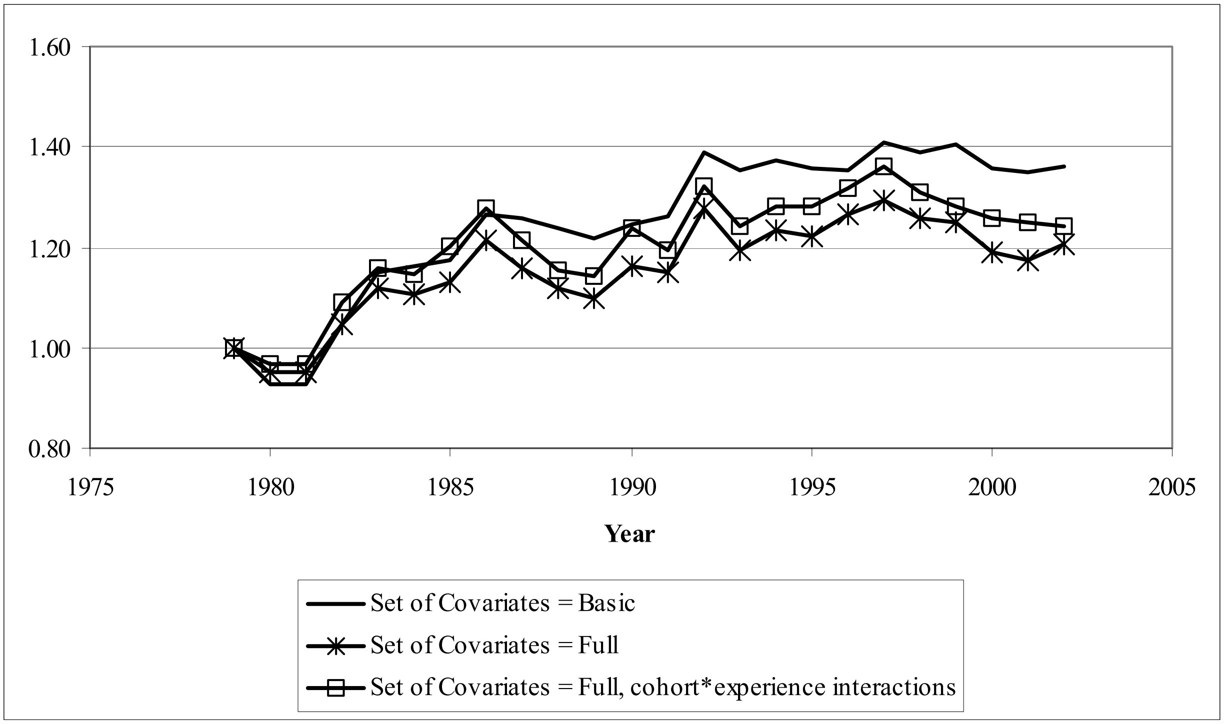


FIGURE 6: PARTIAL-INFORMATION ESTIMATES OF YEAR-SPECIFIC PARAMETERS

(A) CAUSAL RETURN TO EDUCATION (δ)



(B) RETURN TO UNOBSERVED ABILITY (Ψ)

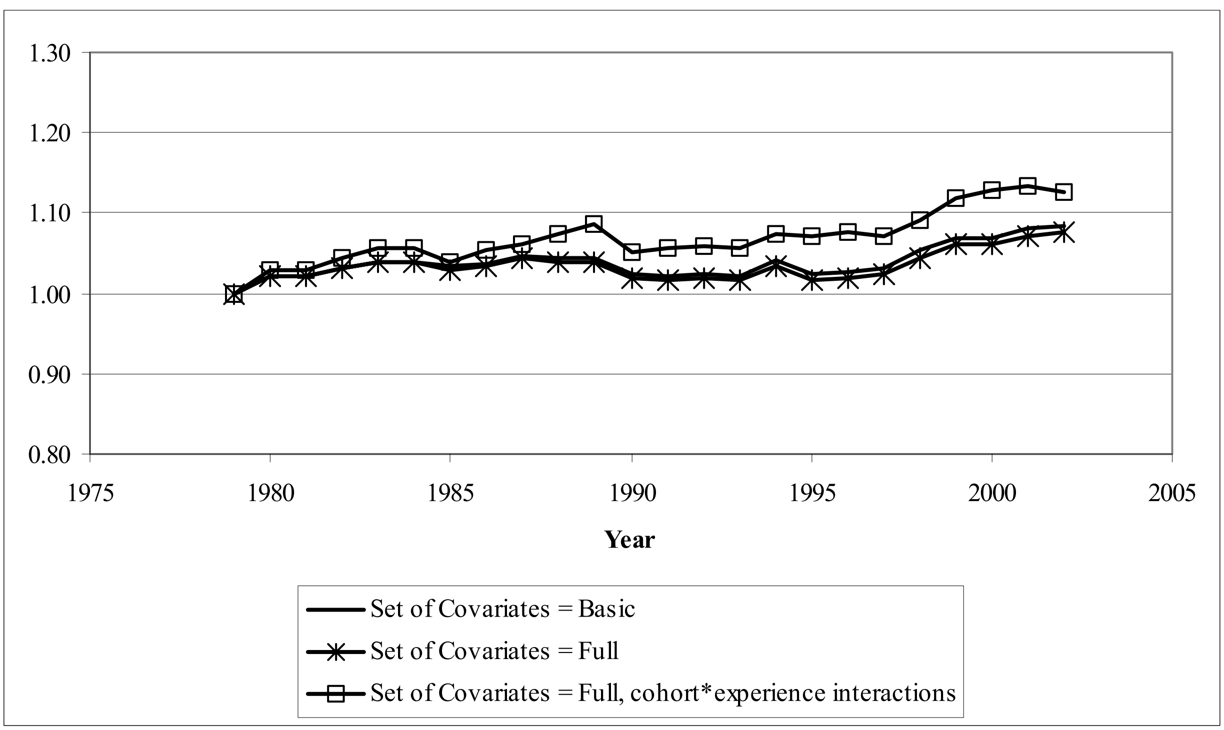
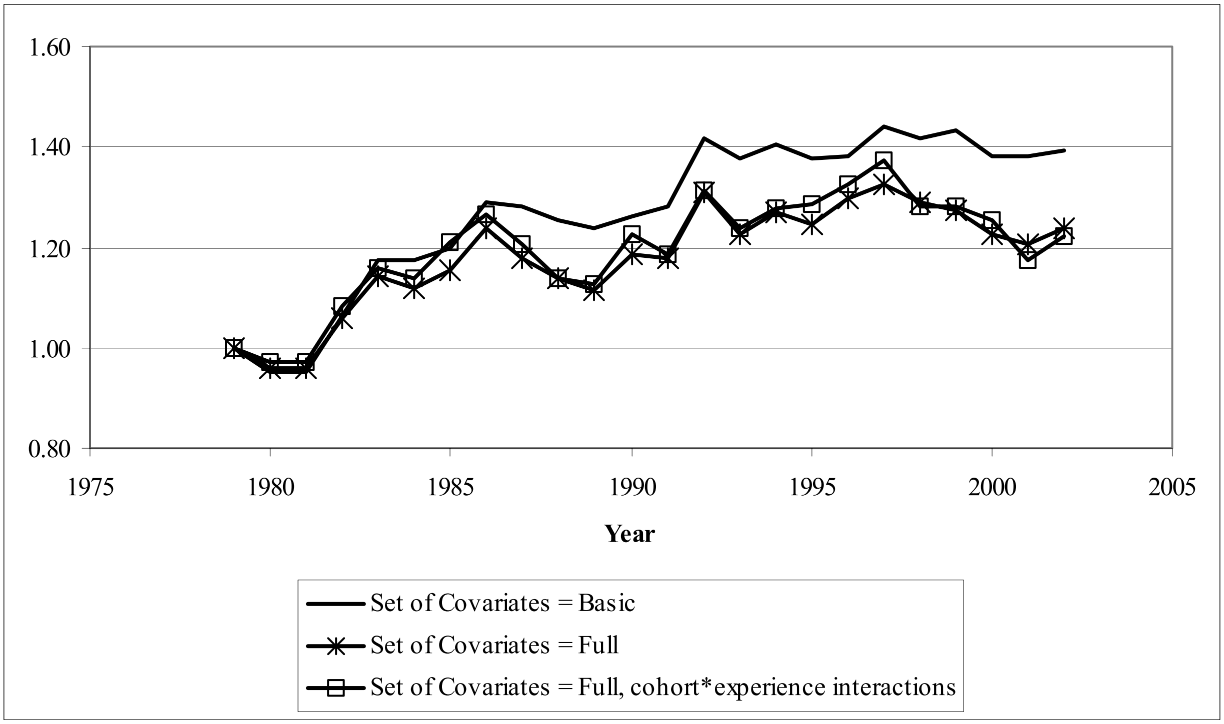


FIGURE 7: FULL-INFORMATION ESTIMATES OF YEAR-SPECIFIC PARAMETERS

(A) CAUSAL RETURN TO EDUCATION (δ)



(B) RETURN TO UNOBSERVED ABILITY (Ψ)

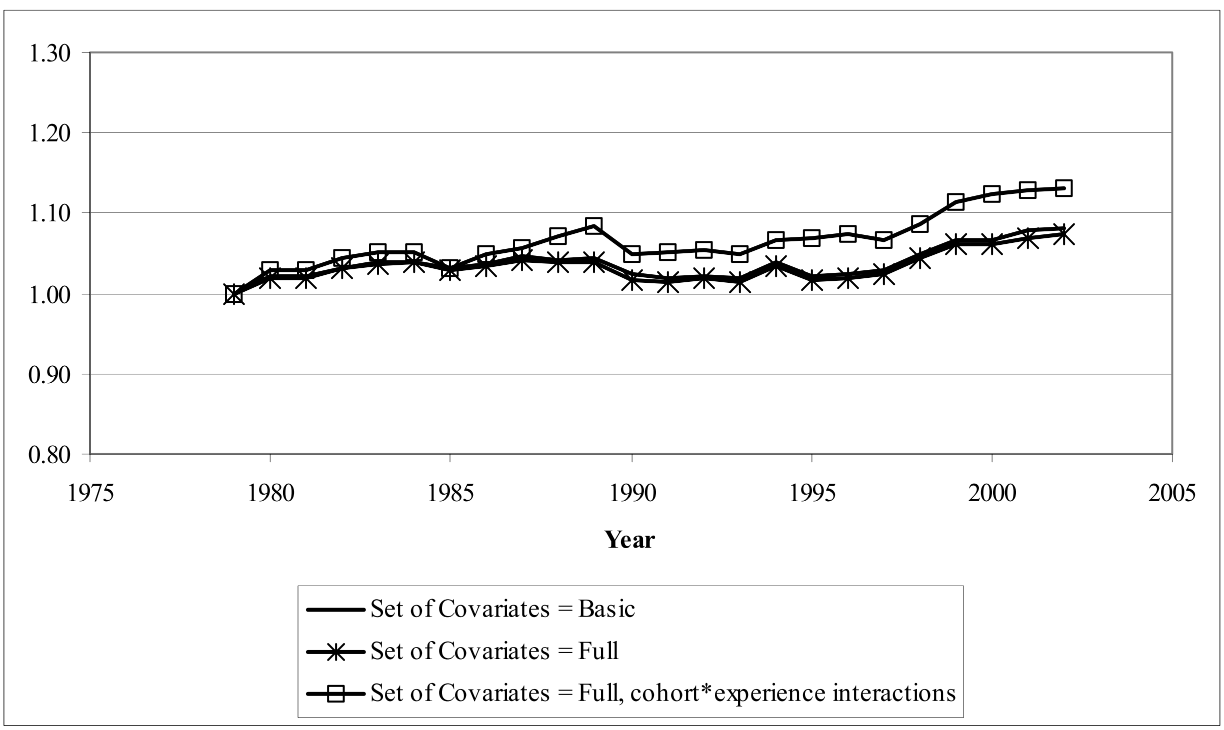


FIGURE 8: LEVEL OF COMPARATIVE ADVANTAGE BIAS FOR VARIOUS DEGREES OF CURVATURE IN THE EARNINGS FUNCTION

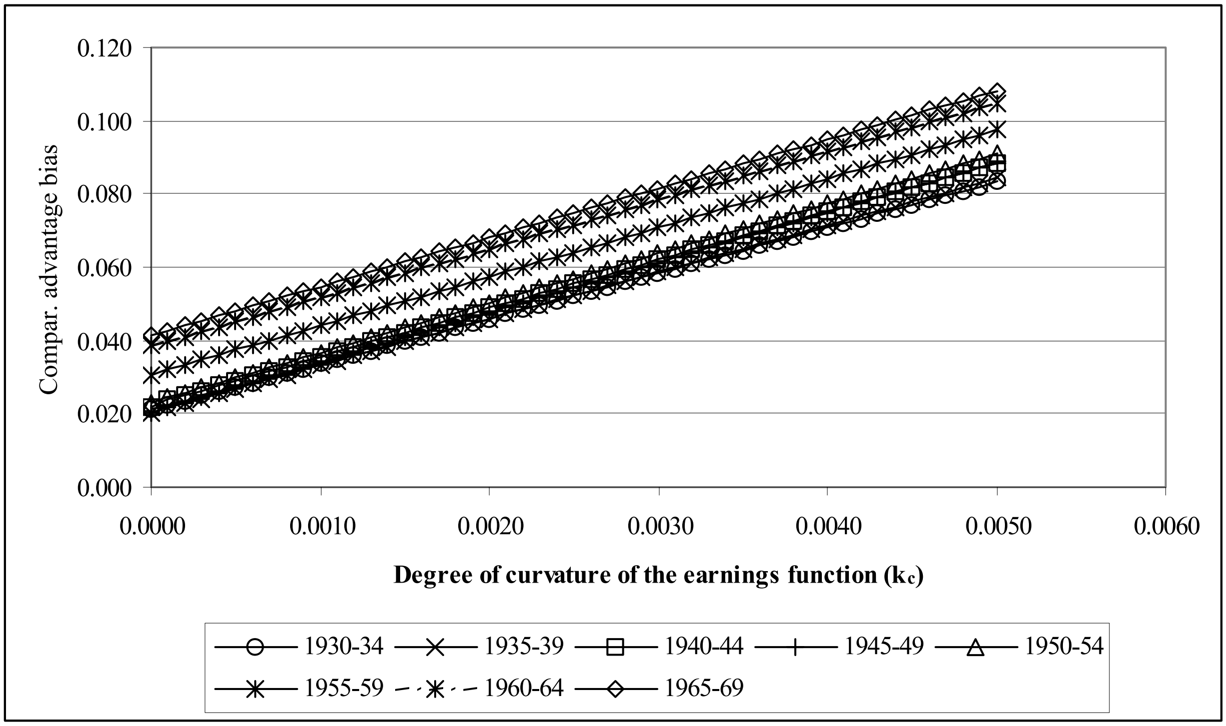
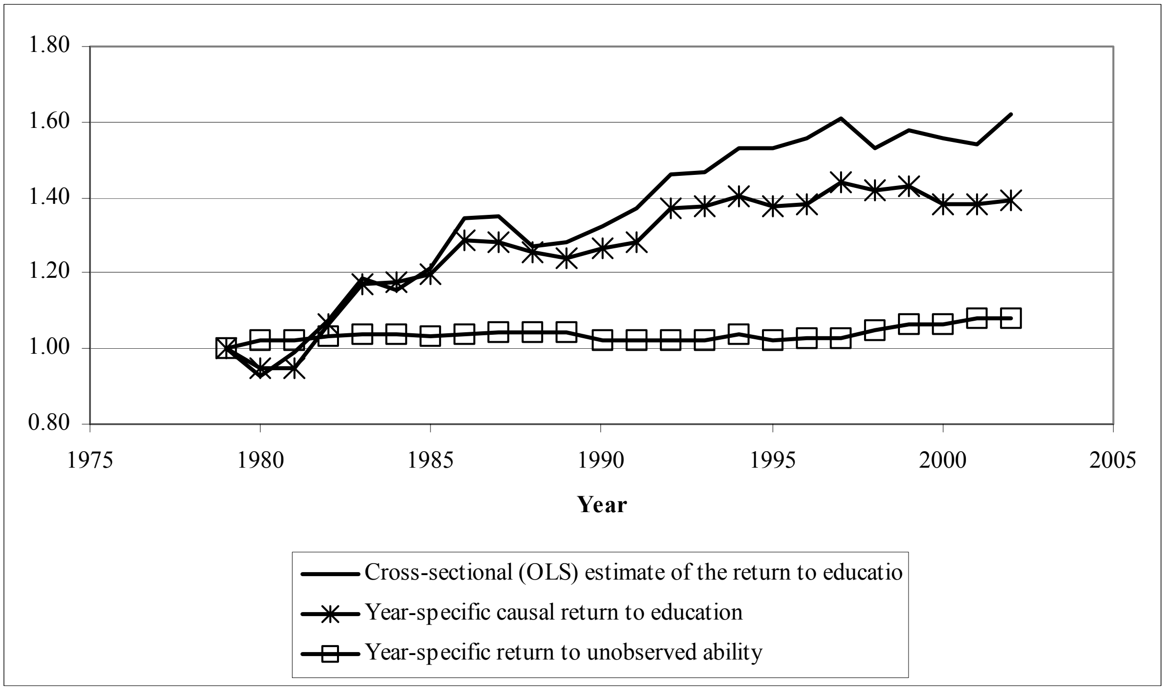


FIGURE 9: DECOMPOSING THE CHANGES IN THE EDUCATIONAL WAGE STRUCTURE

(A) CROSS-SECTIONAL MEASURE OF THE RETURN TO EDUCATION



(B) RESIDUAL WAGE DISPERSION (STD DEV OF LOG HOURLY WAGE RESIDUALS)

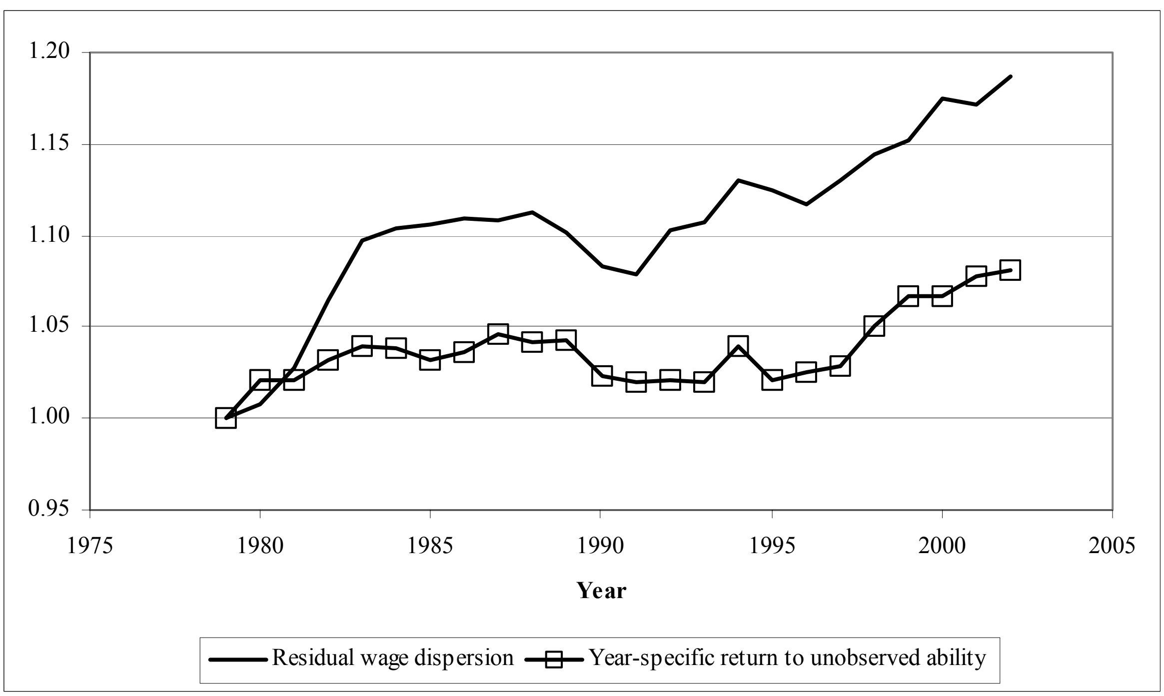
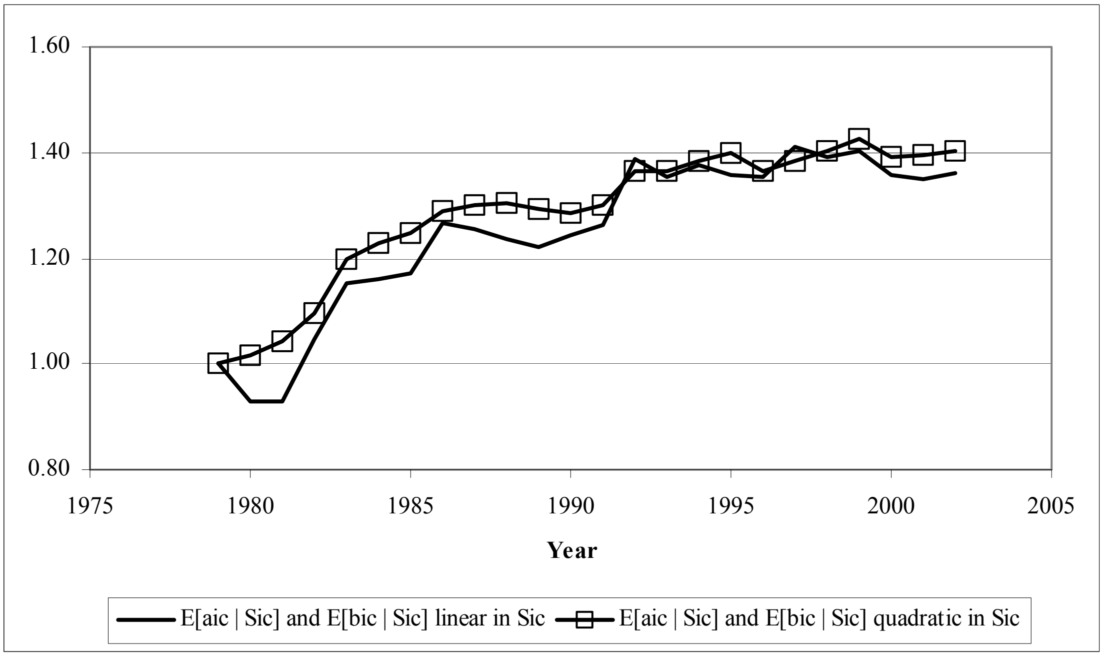


FIGURE 10: FULL-INFORMATION ESTIMATES OF YEAR-SPECIFIC PARAMETERS, ALLOWING A NONLINEAR RELATIONSHIP BETWEEN THE UNOBSERVABLES AND SCHOOLING

(A) CAUSAL RETURN TO EDUCATION (δ)



(B) RETURN TO UNOBSERVED ABILITY (Ψ)

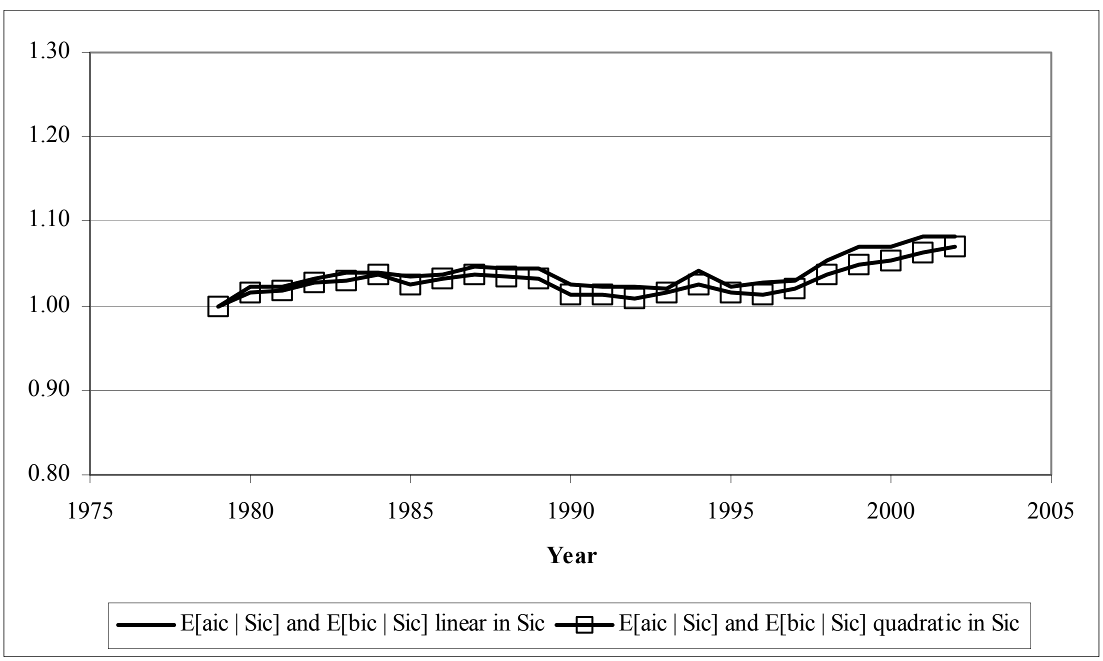


TABLE 1: SUMMARY STATISTICS

	Birth Cohort: 1930-34	1935-39	1940-44	1945-49	1950-54	1955-59	1960-64	1965-69
Real Hourly Wage	19.80	19.93	20.23	19.88	18.54	18.06	17.26	16.70
	(11.60)	(12.01)	(12.39)	(12.01)	(11.28)	(11.41)	(11.26)	(11.02)
Years of Education	12.50	12.83	13.27	13.70	13.63	13.42	13.34	13.43
	(3.32)	(3.12)	(2.99)	(2.85)	(2.67)	(2.57)	(2.56)	(2.60)
Fraction Black	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.09
Fraction Hispanic	0.06	0.07	0.07	0.07	0.08	0.08	0.11	0.14
Entry Year in Sample	1979	1979	1979	1979	1980	1985	1990	1995
Exit / Last Year in Sample	1990	1995	2000	2002	2002	2002	2002	2002
Age at Entry	47.01	41.93	36.87	31.85	28.00	28.00	28.08	28.09
Experience at Entry	28.68	23.24	17.83	12.44	8.61	8.73	8.95	8.69
Age at Exit	57.88	57.83	57.76	54.77	49.91	45.01	40.01	35.00
Experience at Exit	39.18	38.81	38.35	34.85	30.06	25.43	20.53	15.59
Observations	38,176	53,982	80,537	114,292	126,564	103,705	72,131	38,990

Standard deviations in parentheses
The sample include men born between 1930-1969, aged 26-60 in the survey year, with real hourly wages ranging between 5-100 ($2002).
Only observations in the first rotation out of the CPS samples are included. All statistics weighted by hours worked in the survey week.

TABLE 2: OMD ESTIMATES OF THE COHORT-SPECIFIC PARAMETERS

	(1) Average Marg. Return (b)	(2) Rel. Ability Bias (λ_1)	(3) Compar. Adv. Bias ($\lambda_2\overline{S}$)	(4) Average Marg. Return (b)	(5) Rel. Ability Bias (λ_1)	(6) Compar. Adv. Bias ($\lambda_2\overline{S}$)	(7) Average Marg. Return (b)	(8) Rel. Ability Bias (λ_1)	(9) Compar. Adv. Bias ($\lambda_2\overline{S}$)
Birth Cohort:									
1930-34	0.0365	-0.0021	0.0215	0.0311	-0.0110	0.0307	0.0411	-0.0214	0.0282
	(0.0031)	(0.0007)	(0.0062)	(0.0033)	(0.0057)	(0.0085)	(0.0042)	(0.0056)	(0.0081)
1935-39	0.0382	0.0018	0.0208	0.0268	0.0003	0.0270	0.0334	-0.0073	0.0250
	(0.0020)	(0.0005)	(0.0054)	(0.0027)	(0.0039)	(0.0068)	(0.0037)	(0.0053)	(0.0066)
1940-44	0.0428	0.0012	0.0225	0.0287	0.0025	0.0292	0.0319	-0.0005	0.0259
	(0.0018)	(0.0003)	(0.0053)	(0.0024)	(0.0033)	(0.0068)	(0.0035)	(0.0048)	(0.0062)
1945-49	0.0442	-0.0002	0.0211	0.0255	0.0096	0.0264	0.0283	0.0084	0.0266
	(0.0016)	(0.0003)	(0.0046)	(0.0013)	(0.0042)	(0.0057)	(0.0022)	(0.0057)	(0.0059)
1950-54	0.0416	0.0007	0.0237	0.0251	0.0094	0.0265	0.0244	0.0100	0.0235
	(0.0018)	(0.0003)	(0.0048)	(0.0023)	(0.0030)	(0.0054)	(0.0033)	(0.0045)	(0.0051)
1955-59	0.0382	0.0002	0.0313	0.0287	0.0042	0.0298	0.0265	0.0060	0.0278
	(0.0016)	(0.0003)	(0.0062)	(0.0025)	(0.0034)	(0.0059)	(0.0035)	(0.0051)	(0.0058)
1960-64	0.0322	0.0013	0.0387	0.0341	-0.0080	0.0358	0.0318	-0.0029	0.0352
	(0.0017)	(0.0004)	(0.0077)	(0.0028)	(0.0039)	(0.0071)	(0.0037)	(0.0055)	(0.0073)
1965-69	0.0303	0.0031	0.0415	0.0334	-0.0070	0.0365	0.0228	0.0077	0.0307
	(0.0020)	(0.0007)	(0.0084)	(0.0032)	(0.0049)	(0.0075)	(0.0037)	(0.0062)	(0.0070)
Goodness-of-Fit		628.5			484.7			503.2	
[d.f.]		[341]			[341]			[341]	
Set of Covariates		Basic			Full			Full	
Cohort*Experience		No			No			Yes	

Robust standard errors in parentheses. The standard errors in columns (3), (6), (9) computed using a Taylor series expansion.

TABLE 3: FULL-INFORMATION ESTIMATES OF YEAR-SPECIFIC PARAMETERS, CORRECTED FOR CENSORING

	Fraction topcoded	**Year-Specific Return to Education (δ)** (1)		(2)		**Year-Specific Return to Unobserved Ability (Ψ)** (1)		(2)	
Year:		Estimate	Std Err	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err
1979	0.01	1.000	---	1.000	---	1.000	---	1.000	---
1980	0.01	0.974	0.034	0.969	0.034	1.021	0.003	1.017	0.003
1981	0.02	0.998	0.035	1.004	0.036	1.023	0.003	1.019	0.003
1982	0.03	1.016	0.037	1.015	0.038	1.029	0.003	1.027	0.003
1983	0.04	1.135	0.039	1.121	0.040	1.034	0.003	1.033	0.003
1984	0.05	1.159	0.040	1.095	0.040	1.038	0.003	1.035	0.003
1985	0.05	1.183	0.039	1.105	0.038	1.031	0.003	1.021	0.003
1986	0.07	1.276	0.041	1.174	0.040	1.036	0.003	1.031	0.003
1987	0.08	1.330	0.042	1.189	0.041	1.045	0.003	1.037	0.003
1988	0.09	1.264	0.041	1.137	0.040	1.038	0.003	1.033	0.003
1989	0.01	1.281	0.043	1.113	0.041	1.040	0.003	1.034	0.003
1990	0.01	1.281	0.039	1.152	0.038	1.024	0.003	1.016	0.003
1991	0.01	1.321	0.041	1.146	0.039	1.021	0.003	1.016	0.003
1992	0.01	1.390	0.046	1.210	0.045	1.021	0.003	1.015	0.003
1993	0.02	1.379	0.044	1.168	0.041	1.021	0.003	1.014	0.003
1994	0.03	1.411	0.045	1.206	0.042	1.035	0.003	1.026	0.003
1995	0.03	1.403	0.045	1.221	0.043	1.019	0.004	1.011	0.003
1996	0.02	1.439	0.045	1.306	0.044	1.025	0.003	1.018	0.003
1997	0.03	1.415	0.045	1.260	0.043	1.026	0.003	1.018	0.003
1998	0.01	1.437	0.045	1.226	0.043	1.049	0.003	1.039	0.003
1999	0.02	1.418	0.045	1.240	0.043	1.066	0.003	1.053	0.003
2000	0.02	1.391	0.044	1.194	0.041	1.062	0.003	1.055	0.003
2001	0.02	1.385	0.047	1.177	0.044	1.074	0.003	1.065	0.003
2002	0.02	1.379	0.046	1.191	0.044	1.078	0.003	1.068	0.003
Goodness-of-Fit		887.3		608.1		887.3		608.1	
Set of covariates included		Basic		Full		Basic		Full	

Each model has 341 degrees of freedom, with a corresponding 5% critical value of 385.1

TABLE 4: OMD ESTIMATES OF COHORT-SPECIFIC PARAMETERS, CORRECTED FOR CENSORING

	(1) Average Marg. Return (b)	(2) Rel. Ability Bias (λ_1)	(3) Compar. Adv. Bias ($\lambda_2 \overline{S}$)	(4) Average Marg. Return (b)	(5) Rel. Ability Bias (λ_1)	(6) Compar. Adv. Bias ($\lambda_2 \overline{S}$)
Birth Cohort:						
1930-34	0.0458	-0.0026	0.0137	0.0282	-0.0016	0.0286
	(0.0030)	(0.0009)	(0.0044)	(0.0031)	(0.0007)	(0.0080)
1935-39	0.0424	-0.0017	0.0170	0.0298	-0.0012	0.0301
	(0.0017)	(0.0007)	(0.0045)	(0.0019)	(0.0004)	(0.0076)
1940-44	0.0487	-0.0009	0.0163	0.0364	-0.0009	0.0277
	(0.0018)	(0.0003)	(0.0039)	(0.0017)	(0.0003)	(0.0064)
1945-49	0.0503	-0.0002	0.0147	0.0397	-0.0002	0.0240
	(0.0016)	(0.0001)	(0.0033)	(0.0015)	(0.0004)	(0.0052)
1950-54	0.0470	0.0008	0.0186	0.0380	0.0006	0.0239
	(0.0017)	(0.0003)	(0.0038)	(0.0016)	(0.0003)	(0.0049)
1955-59	0.0427	0.0003	0.0289	0.0373	0.0001	0.0291
	(0.0015)	(0.0003)	(0.0057)	(0.0016)	(0.0003)	(0.0059)
1960-64	0.0352	0.0011	0.0391	0.0300	0.0009	0.0385
	(0.0016)	(0.0006)	(0.0077)	(0.0017)	(0.0004)	(0.0076)
1965-69	0.0318	0.0032	0.0411	0.0284	0.0023	0.0386
	(0.0020)	(0.0007)	(0.0083)	(0.0021)	(0.0006)	(0.0079)
Goodness-of-Fit		887.3			608.1	
[d.f.]		[341]			[341]	
Set of Covariates		Basic			Full	

Robust standard errors in parentheses. Each model has 341 degrees of freedom, with a corresponding 5% critical value of 385.1

TABLE 5: GOODNESS-OF-FIT TESTS OF ALTERNATIVE MODELS

	OLS	CLAD
1. Two-Factor Model	628.5	887.3
[d.f.]	[341]	[341]
(p-value)	(0.00)	(0.00)
2. No cohort variation in schooling slope ($b_c = b$)	82.5	140.2
[d.f.]	[348]	[348]
(p-value)	(0.00)	(0.00)
3. No cohort variation in compar. adv. bias ($\lambda_2^c = \lambda_2$)	204.5	392.5
[d.f.]	[348]	[348]
(p-value)	(0.00)	(0.00)
4. No ability bias ($\lambda_1^c = 0$)	69.2	73.5
[d.f.]	[348]	[348]
(p-value)	(0.00)	(0.00)
5. No comparative advantage bias ($\lambda_2^c = 0$)	3360.8	2802.7
[d.f.]	[349]	[349]
(p-value)	(0.00)	(0.00)
6. Stationary return to education ($\delta_t = 1$)	662.1	711.3
[d.f.]	[364]	[364]
(p-value)	(0.00)	(0.00)
7. Stationary return to unobserved ability ($\psi_t = 1$)	1511.0	1419.2
[d.f.]	[364]	[364]
(p-value)	(0.00)	(0.00)
8. Single return to skill ($\delta_t = \psi_t$)	597.1	659.6
[d.f.]	[364]	[364]
(p-value)	(0.00)	(0.00)

The entries in row 1 are the goodness-of-fit statistics associated with the two-factor model presented in Section 5. The entries in rows 2-8 are the differences in the goodness-of-fit statistics between each row and the two-factor model of row 1.

APPENDIX TABLE 1: PARTIAL-INFORMATION OMD ESTIMATES OF YEAR-SPECIFIC PARAMETERS

	Year-Specific Return to Education (δ)						**Year-Specific Return to Unobserved Ability (Ψ)**					
	(1)		(2)		(3)		(1)		(2)		(3)	
Year:	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err
1979	1.000	---	1.000	---	1.000	---	1.000	---	1.000	---	1.000	---
1980	0.928	0.035	0.950	0.040	0.967	0.048	1.022	0.003	1.021	0.003	1.029	0.006
1981	0.928	0.035	0.950	0.040	0.967	0.048	1.022	0.003	1.021	0.003	1.029	0.006
1982	1.047	0.040	1.050	0.040	1.091	0.055	1.032	0.003	1.032	0.003	1.045	0.007
1983	1.151	0.043	1.120	0.040	1.158	0.058	1.040	0.003	1.038	0.003	1.056	0.007
1984	1.162	0.043	1.100	0.040	1.148	0.057	1.039	0.003	1.038	0.003	1.055	0.007
1985	1.173	0.041	1.130	0.040	1.202	0.057	1.034	0.003	1.030	0.003	1.038	0.008
1986	1.267	0.044	1.210	0.040	1.279	0.060	1.037	0.003	1.035	0.003	1.053	0.008
1987	1.256	0.043	1.160	0.040	1.214	0.058	1.047	0.003	1.043	0.003	1.061	0.008
1988	1.237	0.043	1.120	0.040	1.155	0.057	1.043	0.003	1.039	0.003	1.073	0.008
1989	1.219	0.045	1.100	0.040	1.142	0.059	1.045	0.004	1.040	0.003	1.086	0.009
1990	1.245	0.042	1.160	0.040	1.236	0.057	1.024	0.003	1.019	0.003	1.052	0.009
1991	1.262	0.043	1.150	0.040	1.194	0.056	1.022	0.003	1.017	0.003	1.057	0.008
1992	1.388	0.048	1.280	0.050	1.321	0.063	1.023	0.003	1.019	0.003	1.058	0.009
1993	1.353	0.047	1.190	0.050	1.243	0.060	1.021	0.003	1.017	0.003	1.055	0.008
1994	1.375	0.047	1.230	0.050	1.281	0.062	1.041	0.003	1.034	0.003	1.074	0.009
1995	1.359	0.047	1.220	0.050	1.282	0.063	1.023	0.004	1.017	0.004	1.072	0.011
1996	1.355	0.047	1.270	0.050	1.316	0.063	1.027	0.003	1.020	0.003	1.077	0.009
1997	1.409	0.048	1.290	0.050	1.362	0.064	1.031	0.003	1.024	0.003	1.071	0.009
1998	1.390	0.048	1.260	0.050	1.314	0.062	1.053	0.004	1.044	0.003	1.092	0.009
1999	1.404	0.048	1.250	0.050	1.283	0.062	1.069	0.004	1.062	0.003	1.118	0.009
2000	1.356	0.046	1.190	0.050	1.258	0.059	1.069	0.004	1.062	0.003	1.129	0.009
2001	1.349	0.049	1.180	0.050	1.252	0.062	1.081	0.004	1.072	0.003	1.134	0.009
2002	1.362	0.049	1.210	0.050	1.247	0.063	1.083	0.004	1.076	0.003	1.127	0.009
Goodness-of-Fit	939.6		529.6		523.4		939.6		529.6		523.4	
Set of covariates included	Basic		Full		Full		Basic		Full		Full	
Cohort*Experience	No		No		Yes		No		No		Yes	

Each model has 212 degrees of freedom, with a corresponding 5% critical value of 247.0

APPENDIX TABLE 2: FULL-INFORMATION OMD ESTIMATES OF YEAR-SPECIFIC PARAMETERS

	Year-Specific Return to Education (δ)						Year-Specific Return to Unobserved Ability (Ψ)					
	(1)		(2)		(3)		(1)		(2)		(3)	
Year:	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err
1979	1.000	---	1.000	---	1.000	---	1.000	---	1.000	---	1.000	---
1980	0.950	0.035	0.958	0.040	0.971	0.044	1.021	0.003	1.020	0.003	1.029	0.007
1981	0.950	0.035	0.958	0.040	0.971	0.044	1.021	0.003	1.020	0.003	1.029	0.007
1982	1.064	0.039	1.057	0.044	1.083	0.049	1.032	0.003	1.031	0.003	1.043	0.007
1983	1.173	0.042	1.142	0.047	1.159	0.052	1.039	0.003	1.037	0.003	1.052	0.008
1984	1.174	0.042	1.117	0.046	1.137	0.051	1.038	0.003	1.038	0.003	1.052	0.007
1985	1.198	0.041	1.153	0.045	1.208	0.053	1.032	0.003	1.028	0.003	1.032	0.009
1986	1.288	0.043	1.236	0.047	1.267	0.055	1.036	0.003	1.034	0.004	1.048	0.009
1987	1.283	0.043	1.179	0.046	1.206	0.054	1.046	0.003	1.042	0.004	1.056	0.009
1988	1.254	0.043	1.138	0.045	1.140	0.052	1.042	0.003	1.038	0.003	1.070	0.009
1989	1.238	0.044	1.114	0.047	1.127	0.055	1.043	0.004	1.039	0.004	1.083	0.010
1990	1.263	0.041	1.186	0.045	1.225	0.053	1.023	0.003	1.017	0.004	1.049	0.009
1991	1.283	0.043	1.178	0.046	1.188	0.053	1.020	0.003	1.015	0.004	1.051	0.008
1992	1.417	0.048	1.310	0.051	1.313	0.060	1.021	0.003	1.018	0.004	1.053	0.010
1993	1.376	0.046	1.227	0.049	1.238	0.057	1.020	0.003	1.015	0.004	1.049	0.009
1994	1.403	0.047	1.268	0.050	1.277	0.059	1.039	0.003	1.033	0.004	1.067	0.010
1995	1.379	0.047	1.244	0.050	1.284	0.061	1.021	0.004	1.016	0.004	1.068	0.012
1996	1.382	0.046	1.298	0.051	1.327	0.061	1.025	0.003	1.019	0.004	1.073	0.010
1997	1.439	0.048	1.325	0.051	1.373	0.062	1.029	0.003	1.024	0.005	1.067	0.010
1998	1.418	0.048	1.290	0.051	1.280	0.060	1.050	0.004	1.043	0.004	1.087	0.010
1999	1.432	0.048	1.273	0.050	1.280	0.059	1.067	0.004	1.060	0.004	1.113	0.010
2000	1.382	0.046	1.224	0.047	1.253	0.057	1.067	0.004	1.060	0.004	1.123	0.009
2001	1.380	0.049	1.205	0.051	1.175	0.060	1.078	0.004	1.069	0.004	1.129	0.009
2002	1.393	0.049	1.239	0.052	1.220	0.061	1.081	0.004	1.073	0.004	1.131	0.009
Goodness-of-Fit	628.5		484.7		503.2		628.5		484.7		503.2	
Set of covariates included	Basic		Full		Full		Basic		Full		Full	
Cohort*Experience	No		No		Yes		No		No		Yes	

Each model has 341 degrees of freedom, with a corresponding 5% critical value of 385.1

APPENDIX TABLE 3: SIMULATION RESULTS AND MEASUREMENT ERROR CORRECTIONS

	Actual (From Table 2)	**Design 1** [π_1=0.08, π_2=0.001]	**Design 2** [π_1=0.06, π_2=0.002]	**Design 3** [π_1=0.04, π_2=0.003]	**Design 4** [π_1=0.02, π_2=0.004]
[A] Simulation Results					
Intercept	---	0.97	0.97	0.98	0.99
Linear Schooling Term	---	1.08	1.03	0.94	0.76
Quadratic Schooling Term	---	1.25	1.25	1.25	1.21
[B] OMD Estimates Corrected for Measurement Error					
Year-Specific Parameters:					
Causal Return to Education					
1979	1.000	1.000	1.000	1.000	1.000
1990	1.263	1.370	1.410	1.489	1.655
2002	1.393	1.539	1.593	1.703	1.932
Return to Unobserved Ability					
1979	1.000	1.000	1.000	1.000	1.000
1990	1.023	1.024	1.024	1.024	1.023
2002	1.081	1.079	1.077	1.075	1.071
Cohort-Specific Parameters:					
Born 1945-49					
Avg. Marginal Product	0.0442	0.0479	0.0457	0.0417	0.0340
Rel. Ability Bias	-0.0002	-0.0001	-0.0001	-0.0001	0.0000
Comparative Advantage Bias	0.0211	0.0250	0.0245	0.0236	0.0211
Born 1945-49					
Avg. Marginal Product	0.0303	0.0363	0.0353	0.0341	0.0304
Rel. Ability Bias	0.0031	0.0030	0.0030	0.0030	0.0029
Comparative Advantage Bias	0.0415	0.0470	0.0454	0.0426	0.0364

(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 correction factors were obtained by the simulations and are reported in the top panel of the table.