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

Olivier Deschênes

Department of Economics
2127 North Hall
University of California, Santa Barbara
Santa Barbara, CA 93106-9210

This version: June 2006

Abstract: This paper analyzes the changes in the educational wage structure in the United States between 1979 and 2002. In order to impose some structure on the observed relationship between schooling and earnings, I develop a model for earnings and schooling that incorporates heterogeneity in absolute and comparative advantage across individuals. 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 effect of 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 approximately 40% between 1979 and 2002, after controlling for the confounding effects of time-varying unobserved ability and comparative advantage biases across cohorts. I conclude that the rise in the causal effect of education, and not the rise in the return to unobserved ability explains most of the observed change in the educational wage structure in the U.S. between 1979 and 2002.

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

1. Introduction

The widening of the wage structure since the early 1980s is one of the better-documented facts about the U.S. labor market (see e.g., Katz and Autor 1999, Acemoglu 2002 for surveys). Both the between-group component and the within-group (or residual) component of wage inequality have been increasing. The early consensus explaining this phenomenon was that rising inequality stemmed from an increased demand to all dimensions of skills (education, unobserved ability, etc), combined with a slowdown in the relative supply of educated workers. In addition, the decline in the minimum wage and unionization further contributed to increased wage inequality.

In recent years, however, several pieces of evidence pertaining to the timing, magnitude, and distributional location of the evolution of wage inequality have challenged the early consensus. For example, Lemieux (2006a) showed that residual wage inequality practically remained constant in the 1990s once composition effects were accounted for. Similarly, Autor, Katz and Kearny (2006) found that over the last fifteen years, residual wage inequality growth was not uniform over its distribution. They report that residual inequality increased at the top of the distribution while it actually declined in the bottom, a pattern they attribute to the polarization of the U.S. labor market. The clear implication of these findings is that the returns to the various dimensions of skills have not been increasing uniformly over the last decades.

In this paper, I propose a framework for interpreting the observed changes in the educational wage structure, the central dimension of *between-group* wage inequality. The purpose is to assess whether the causal effect of education on labor market earnings has changed or whether the rise in the educational wage differentials is confounded by time-varying ability biases. This information has tremendous importance in the context of the current debate about the trends in inequality. For example, it has been argued that rising return to unobservable skills correlated with education is the main explanation behind the increased education wage differentials (see e.g., Taber 2001).[1]

The interpretation of the observed education wage differentials is guided by a simple extension of Card's (1999) causal model of earnings and schooling to the case where repeated cross-sectional observations on individuals from different birth cohorts are available. As a consequence, and unlike most recent studies of the wage structure, two factors of unobserved ability are considered: absolute ability and heterogeneous returns to education. It follows that changes in the conventional measure of the return to

[1] See also Blackburn and Neumark (1993), and Murnane, Willett and Levy (1995).

education can be decomposed into three components: (i) changes in the causal effect of education over time; (ii) changes in the return to unobserved ability over time; and (iii) changes in the mapping between the two factors of ability and completed education across cohorts.

Another central idea in the model is 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 introduce a year-specific component in the causal effect of education and in the return to unobserved ability. Second, the education wage differentials may vary over time because of composition or cohort effects, and thus the model will account for cohort-specific components. Combined, both dimensions entail a time-varying causal effect of education and time-varying ability biases.

The key empirical prediction 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 log 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 effect of 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 earnings regression.

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 the "convexification" of the educational wage structure also occurs within cohort.[2] I then make use of these changes in the U.S. wage structure to identify the parameters of the structural 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 two decades. 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.

My estimates indicate that most of the recent increase in the educational wage differentials is attributable to a rise in year-specific component of the causal effect of education. In particular, I find that the year-

[2] I borrow the term "convexification" from Lawrence Katz's Albert Rees Lecture at the 2006 Society of Labor Economists meetings.

specific component of the causal effect of education increased by 40-48% between 1979 and 2002, which amounts to 2/3 of the increase in the conventional measure of the return to education over that time period. At the same time, the return to unobserved ability increased by 4-10%. The estimates also indicate an important increasing inter-cohort trend in the correlation between educational attainment and individual-specific returns to education. This finding is confirmed by the evidence in Angrist, Chernozhukov, and Fernandez-Val (2006) who found increased heterogeneity in the Mincerian returns to education across quantiles of the log earnings distribution.[3]

Taken as a whole, the evidence in this paper greatly 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. In addition, this finding has implications for the literature studying wage inequality. The results in this paper imply that an increase in causal return to education, and not an increase in the return to unobserved skills is the main driving force behind the growth in residual wage inequality. The rest 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. Sections 4 and 5 discuss the identification, estimation and interpretation of the results of this paper. Section 6 concludes.

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

This section describes the data and 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. The main advantage of the NLSY is its longitudinal structure that can allow for individual fixed-effect estimation. 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

[3] See Buchinsky (1994) for an earlier analysis of changes in the wage structure using quantile regressions.

a better-suited data source. In addition, the NLSY is not a representative sample of the U.S. workforce as it follows young workers only.

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 (2006b): For workers paid by the hour, which represent approximately 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. Finally, I also follow the literature and adjust top-coded earnings by a factor of 1.5 (see e.g., Katz and Autor 1999). Since only 3% of the sample has top-coded earnings, this consideration is inconsequential.

Following DiNardo, Fortin and Lemieux (1996) and others, 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. 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 wage below $5 or above $200 in 2002 constant dollars. 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.[4] None of these considerations affect the main conclusions of this paper.

Table 1 presents summary statistics for the sample 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 that were subjected 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

[4] 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.

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 on average, 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. This pattern is documented and analyzed by Card and Lemieux (2001). 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. Their inclusion or exclusion from the sample does not affect the results significantly.

A. Changes in the earnings-schooling relationship

The standard human capital earnings function specifies log-linear relationship between wages and years of education, with a constant coefficient on education. 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). In order to obtain some simple evidence on the form of the 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. The regressions also included a quartic in experience, and dummy variables for race, ethnicity, marital status, metropolitan area and census division. The estimated dummy variables are shown in Figure 1 for 3 time periods: 1979-1981, 1989-1991 and 1999-2002.[5]

Figure 1 shows the "convexification" of the log earnings-schooling relationship between 1979 and 2002. 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 log earnings-schooling relationship

[5] 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.

not 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.

This phenomenon also has implication for within-group inequality. Reflecting the patterns showed in Figure 1, Lemieux (2006b) finds that within-group inequality increased substantially among college-educated workers, but mostly remained constant for other groups. Similarly, Autor, Katz and Kearny (2006) show that residual wage inequality growth was not uniform over its distribution: it grew more or less continuously over the last two decades in the top half of the distribution of wage residuals and actually declined at the bottom half. In addition, they find a similar "convexification" in employment growth patterns by education over the 1980s and 1990s. More precisely they report that the share of total hours worked increased the most in high-skill jobs, the least in middle-skill jobs, and was somewhat in between for low-skill jobs.[6] Finally, a recent paper by Angrist, Chernozhukov, and Fernandez-Val (2006) also documents a related increased convexity in the U.S. educational wage structure. They use data from the 1980, 1990, and 2000 U.S. Census and find that changes over time in the Mincerian return to schooling was not uniform across quantiles of the log earnings distribution. In particular, they show that the return to schooling was practically constant across quantiles in 1980 (at about 7.5%). However, in 1990 and 2000, the returns to education displayed vast heterogeneity across quantiles, and grew the most above the 80th quantile of the log wage distribution.

Figure 1 and the related evidence in the literature suggest that the prominent 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 ORG for 1979-2002. The models included the same set of controls listed above and were estimated for 5 specific time periods, spanning 1979 to 2002. Panel A reports the coefficients on the linear and quadratic terms in education from models that pool all cohorts. The estimates confirm the evidence shown in Figure 1: Throughout the 1980s and 1990s there was a remarkable increase in the contribution of schooling squared to log wages, from 0.0007 to 0.0039, corresponding to a 450% increase. The rise in quadratic education coefficient was accompanied by an offset in the linear term of -120%. Panel B reports the same estimates for selected cohorts: 1935-39, 1945-49, 1955-59. The same pattern of "convexification" of the log wage-schooling function is observed within cohort as well. For all three cohorts considered, the quadratic education coefficient increased over the 1980s and 1990s, while the linear effect of schooling on wages declined. The similarity of the trends in the linear and quadratic schooling coefficients across cohorts points out to a

[6] Autor, Katz, and Kearny (2006) use an occupation's percentile in the 1980 education distribution as proxy for skills.

common underlying mechanism affecting all cohorts equally. At the same time, differences across cohorts in the levels of these 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 model 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. As previously discussed, such changes will arise through time effects and composition effects. To this end, I introduce a time dimension and cohorts into Card's (1999) cross-sectional causal model for earnings and schooling. This allows me to characterize the concept of changing causal effect of education and show how the parameters of the model can be identified from a series of augmented human capital earnings regressions.

To begin, consider a potential outcomes model of life-cycle log earnings as a function of schooling:

$$(1)\ \log Y_{ict}(S) = \Psi(t)a_{ic} + \delta(t)[b_{ic}S - 0.5k_cS^2]$$

where i denotes individuals, c denotes cohorts and t denotes time. This model describes the evolution over time of log earnings for person i, born in cohort c, if that person were assigned schooling level S. Following Rosen (1974), I refer to this equation as the "structural income-generating function". There are two factors representing individual heterogeneity in earnings capacity: a_{ic} is the unobserved ability (i.e. the part of earnings capacity that does not interact with schooling) and b_{ic} is the part of the education slope specific to person i in cohort c. The parameter k_c allows for diminishing returns in the production of earnings capacity through schooling and is assumed to be the same for individuals from the same cohort. While this assumption may seem restrictive, intuitively, the first-order effect of heterogeneity in the return to schooling must appear through the intercept of the marginal benefit equation (which is given by b_{ic}), not its slope (which is in part determined by k_c).[7] In addition, I show below through simulations that allowing for heterogeneity in the curvature parameter and allowing for a different time effect on the curvature parameter has virtually no impact on the estimates. At a given point in time, the model predicts that variation in log earnings among individuals with the same level of schooling, and belonging to same

[7] In a utility-maximizing model of schooling determination the parameter k_c corresponds to the sum of the slopes of the marginal benefit and marginal cost equations.

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").

Equation (1) has a causal interpretation as it provides clear answers to counterfactual questions regarding the effect of an exogenous change in a person's educational attainment on log earnings. The objective of this paper is to obtain information about the parameters underlying equation (1). In particular, the model describes the structural relationship between log earnings and schooling as being determined by a combination of individual/cohort specific parameters (a_{ic}, b_{ic}, k_c) and year-specific parameters ($\psi(t)$ and $\delta(t)$). Optimizing models of schooling determination will lend an economic interpretation to the individual/cohort-specific parameters, while the year-specific parameters reflect the impact of market forces (e.g., the demand for skill) in determining the causal effect of education.

The economic interpretation of the year-specific parameters, $\psi(t)$ and $\delta(t)$, originates from an aggregate model of demand for skills and skill wage premia (see e.g., Katz and Murphy 1992, Acemoglu 2002). The canonical model used to study between- and within-group inequality is the CES production function, where there is a single index of skills, and where education and unobserved ability are perfect substitutes. In that context, growth in the demand for skills, arising in part from skilled-biased technical change, would lead to an increase in the "return to skill", unless the increasing demand is offset by an increase in the supply of skills. This simple model then predicts that between- and within-group inequality should move together, a prediction rejected by U.S. data from the 1970s onwards. My approach here is to be more agnostic and allow the two components of "skills", education and unobserved ability, to be imperfect substitutes and thus have a different return in the labor market. In addition, Acemoglu (2002) shows that a simple "two-factor" model of skills can be used to explain the salient changes in between- and within-group inequality in the U.S. since 1970.

I follow the literature (see e.g., Willis 1986) and assume that individuals choose schooling to maximize the present discounted value of future earnings using a constant discount rate and earning nothing while in school. Assuming that individuals have correct expectations about their own parameters and the future values of the year-specific parameters, this simple model implies that optimal schooling satisfies:

(2) $$S_{ic}^{*} = \frac{b_{ic} - r_{ic}}{k_c}$$

where r_{ic} is a person-specific discount rate and $k_c \geq 0$ is the curvature parameter that appeared in (1). This shows that variation across cohorts in distribution of b_{ic} and r_{ic}, and in the value of k_c will result in differences across cohorts in average educational attainment. Substituting optimal schooling in equation (1) and re-arranging in terms of a constant coefficient model we obtain:

$$(3)\ \log Y_{ict} = \Psi(t)a_c + \delta(t)[b_c S_{ic} - 0.5 k_c S_{ic}^2] + \eta_{ict}$$

where $\eta_{ict} = \Psi(t)(a_{ic} - a_c) + \delta(t)(b_{ic} - b_c)S_{ic}$. In this model, the average causal effect of education for a given cohort at a given point in time is $\beta(c,t) = \delta(t)[b_c - k_c \overline{S}_c]$, where $b_c = E[b_{ic}]$. According to this specification, changes over time in the average causal effect of education in the population will be due to cohort-specific factors (entry and exit of cohorts with different values of b_c, k_c, and $\overline{S}_c$) and to year-specific factors (changes in $\delta(t)$). In what follows, I will refer to these as the cohort-specific and the year-specific components of the average causal effect of education.

However, the error term in (3) is correlated with schooling and thus cross-sectional regressions will not be informative about the average causal effect of education. As (2) shows, the correlation between the error term and schooling arises from the optimizing behavior of individuals. In order to embody this correlation that stems from optimal schooling decisions, and to derive its implications for the population regression function it is useful to consider the following theoretical linear regressions relating the heterogeneity components to years of education:

$$(4.1)\ (a_{ic} - a_c) = \lambda_{1c}(S_{ic} - \overline{S}_c) + u_{1ic}$$

$$(4.2)\ (b_{ic} - b_c) = \lambda_{2c}(S_{ic} - \overline{S}_c) + u_{2ic}$$

where $E[u_{1ic}|S_{ic}]=0$, $E[u_{2ic}|S_{ic}]=0$, $\lambda_{1c}=Cov(a_{ic},S_{ic})/Var(S_{ic})$, and $\lambda_{2c}=Cov(b_{ic},S_{ic})/Var(S_{ic})$. If individuals with higher return to schooling respond to the incentives of comparative advantage and acquire more schooling, the slope λ_{2c} will be positive. The absolute ability slope λ_{1c} can be positive or negative depending on the model describing the optimizing behavior of individuals.[8] Throughout I assume that a_c, b_c, λ_{1c}, and λ_{2c} 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 members made a permanent

[8] See Griliches (1977) for a discussion of this point. The evidence of non-hierarchical sorting in Willis and Rosen (1979) and Garen (1984) suggest that $\lambda_{1c}<0$ and $\lambda_{2c}>0$.

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.

It is straightforward to show that in the model described by (3), (4.1), and (4.2), the conventional measure of the return to education for cohort c observed in year t (which I label $\beta_{ols}(c,t)$) corresponds to:[9]

$$(5)\ \beta_{ols}(c,t) = \delta(t)(b_c - 0.5k_c\overline{S}_c) + \Psi(t)\lambda_{1c} + \delta(t)\lambda_{2c}\overline{S}_c$$

Thus, the conventional estimate of the return to education equals the sum of three components: $\delta(t)(b_c - 0.5k_c\overline{S}_c)$, the average causal effect of education for cohort c in period t, $\Psi(t)\lambda_{1c}$ a time-varying ability bias specific to cohort c and $\delta(t)\lambda_{2c}\overline{S}_c$, 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 across cohort in λ_{1c}) and on the return to unobservable ability in period t ($\Psi(t)$).[10] 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 across cohort in $\lambda_{2c}\overline{S}_c$), and on the return to education specific to period t ($\delta(t)$). Equation (5) highlights the econometric difficulties associated interpreting changes in the educational wage structure: changes in the conventional estimate of the average return to education may be driven by changes in λ_{1c}, λ_{2c} and $\Psi(t)$ as opposed to changes in b_c and $\delta(t)$.

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 λ_{2c}=0) and $\lambda_{1c}=\lambda_1$). As such, our current 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.

4. Identification and Estimation of the Year-Specific Parameters

[9] This result was obtained under the assumption that the third central moment of S_{ic} is 0.

[10] Even with stationary returns to unobserved ability ($\Psi(t)$=1 in all periods), an increase in the ability bias component $\Psi(t)\lambda_{1c}$ can arise if absolute ability sorting is more important for younger cohorts (e.g., if λ_{1c} is rising across cohorts).

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. Unfortunately, credible instrumental variables for schooling are not readily available in large repeated cross-sectional data sets like the CPS. 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., 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. Unfortunately, in a sample of prime-aged adults who have completed their educational investments, fixed-effects estimates of the return to education are not identified without an instrument for education (Hausman and Taylor 1980). However, Angrist and Newey (1991) show that in a sample of young workers, whose educational attainment is still changing, the return to schooling can be estimated in a fixed-effect models. Since such data sets, like the NLSY are small and unrepresentative of the U.S. workforce, this approach is not undertaken here.[11]

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 structural earnings generating function. I begin by showing the that year-specific component of the causal effect of education and the return to unobserved ability ($\delta(t)$ and $\Psi(t)$) are identified up to a normalization under the assumption that the conditional expectation functions of the heterogeneity components are linear in schooling, as expressed in (4.1) and (4.2).[12]

To proceed, substitute equations (4.1) and (4.2) into the structural earnings function (3):

$$(7)\ \log Y_{ict} = \Psi(t)[a_c + \lambda_{1c}(S_{ic} - \overline{S}_c)] + \delta(t)[b_c + \lambda_{2c}(S_{ic} - \overline{S}_c)]S_{ic} - \delta(t)0.5k_c S_{ic}^2 + \varepsilon_{ict}$$

[11] In addition, Angrist (1995) shows that one can estimate the *change* in the return to education using individual fixed-effect model. Because of the limitations of the common panel data sets on wages and schooling, this approach is not undertaken here. It is also unclear how to extend this approach to the correlated random coefficient model.

[12] This linearity assumption would follow from the joint normality assumptions commonly used in the literature (see e.g., Willis and Rosen 1979), but only linearity of the conditional mean is required here.

where $\varepsilon_{ict} = \Psi(t)u_{1ic} + \delta(t)u_{2ic}S_{ic} + v_{ict}$ is an heteroskedastic error term, and where v_{ict} represents sampling error and is presumed orthogonal to schooling. Interestingly, the expression for ε_{ict} illustrates that in a two-factor model of ability, rising residual wage dispersion and rising return to unobserved ability are not mechanically linked, as it is the case in the standard one-factor model of ability (see e.g., Juhn, Murphy and Pierce 1993). Lemieux (2006a) exploits this fact to analyze changes in residual wage inequality.

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

$$(8)\ E[\log Y_{ict} \mid S_{ic}] = \pi_0(c,t) + \pi_1(c,t)S_{ic} + \pi_2(c,t)S_{ic}^2$$

where:

$$(8.1)\ \pi_0(c,t) = \Psi(t)[a_c - \lambda_{1c}\overline{S}_c]$$

$$(8.2)\ \pi_1(c,t) = \delta(t)b_c + \Psi(t)\lambda_{1c}\overline{S}_c - \lambda_{2c}\overline{S}_c$$

$$(8.3)\ \pi_2(c,t) = \delta(t)[\lambda_{2c} - 0.5k_c]$$

A simple intuition regarding the source of identifying variation for $\Psi(t)$ and $\delta(t)$ arise from equations (8.1) and (8.3) above. Variation over time in the quadratic term $\pi_2(c,t)$ identifies the year-specific causal return to education $\delta(t)$ up to a normalization ($\delta(t)$=1 in 1979). Similarly, variation over time in the intercept term $\pi_0(c,t)$ identifies the return to unobserved ability $\Psi(t)$ up to a normalization ($\Psi(t)$=1 in 1979).[13] Figure 1 in the appendix 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 $\delta(t)$ and $\Psi(t)$ are innocuous. While the time-series variation in $\pi_1(c,t)$ is not necessary to identify $\delta(t)$ and $\Psi(t)$, I discuss below a model which exploits this variation as well.

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:

$$(6)\ \log Y_{ict} = \pi_0(c,t) + \pi_1(c,t)S_{ic} + \pi_2(c,t)S_{ic}^2 + g(X_{ict}, \gamma_t) + \varepsilon_{ict}$$

[13] The connection between the price of unobserved skills and the intercept of the log earnings function has also been emphasized in other papers (see e.g., Heckman and Sedlacek (1985)).

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_{ic} 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 $g(X_{ict}, \gamma_t)$ includes a quartic in potential labor market experience, a race indicator, an Hispanic ethnicity indicator, a marital status indicator, dummy variables for the census divisions and a dummy variable for metropolitan areas.[14] I refer to this set of covariates as the "Basic" set of covariates. In addition, in some specifications, the experience profile is allowed to vary by year and by cohort. Otherwise, the effect of each of the 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, $\pi_0(c,t)$ and $\pi_2(c,t)$, which can be used to estimate $\Psi(t)$ and $\delta(t)$ with a minimum-distance estimator. For a given cohort-year pair, the relationship between the estimated regression coefficients and the parameters of the structural earnings function in equations (8.1) and (8.3) can be written as $\hat{\pi}(c,t) = f(\theta_{ct}) + \omega_{ct}$, where:

$$\hat{\pi}(c,t) = \begin{bmatrix} \hat{\pi}_0(c,t) \\ \hat{\pi}_2(c,t) \end{bmatrix} \quad \text{and} \quad f(\theta_{ct}) = \begin{bmatrix} \Psi(t)[a_c - \lambda_{1c}\overline{S}_c] \\ \delta(t)[\lambda_{2c} - 0.5k_c] \end{bmatrix}$$

and where ω_{ct} is a combination of sampling and specification errors. Thus, the parameters of interest $\delta(t)$ and $\Psi(t)$ can be readily estimated using an optimal minimum distance approach (OMD). The combination of cohort-specific parameters $(a_c - \lambda_{1c}\overline{S}_c)$ and $(\lambda_{2c} - k_c)$ are not interpreted as this stage. Below I consider a model that interprets both the year-specific and cohort-specific parameters.

C. Results

Figures 2 and 3 displays the OMD estimates of the year-specific causal return to education and return to unobserved ability ($\delta(t)$ and $\Psi(t)$) for every year between 1979 and 2002. Since these estimates only use information on $\pi_0(c,t)$ and $\pi_2(c,t)$, I label them "partial-information" estimates. Estimates corresponding to two different specifications of the log wage regressions are showed. The solid line corresponds to the regressions including the basic set of covariates and squared line corresponds to the regressions that include cohort-specific quadratic experience profiles in addition to the basic set of covariates. The

[14] 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).

estimates show that the year-specific component of the causal effect education increased by 40-44% between 1979 and 2002, though the increase was not uniform over the whole time period. All the point estimates are statistically significant, with standard errors ranging from 0.03-0.05. Figure 2 reveals important information about the timing of the change: Almost all of the increase is concentrated in specific time periods, namely 1981-1987 and 1989-1992.

Figure 3 shows that the return to unobserved ability increased by 4-10% over the 1980s and 1990s. Again, the growth was 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 results of Card and DiNardo (2002), and Lemieux (2006b) who found that most of the increase in the standard deviation of log wage residuals between 1970 and 2002 can be accounted for by increases in the early 1980s. In fact, most "overall" measures of residual wage inequality like the 90-10 percentile differential of log wage residuals and the standard deviation of log wage residuals grew at differential rates over the 1980s and 1990s. However, as showed by Autor, Katz, and Kearny (2006), inequality in the upper half of the wage distribution grew almost continuously over the 1980s and 1990s.

A clear implication of Figures 2 and 3 is 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.[15] This is in sharp contrast with the previous literature, for example Taber (2001), who found attributed all of the increase in the conventional measure of the return to college over the 1980s to an increase in the return to unobserved ability.

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

A. Identification

More assumptions and normalizations are required to identify the cohort-specific parameters. 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

[15] The conventional estimate of the return to education (i.e. from a Mincerian regression) increased by about 60% between 1979 and 2002.

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 also simplifies the identification problem.

As shown earlier, $\Psi(t)$ and $\delta(t)$ can be identified from the time-series variation in the regression coefficients $\pi_0(c,t)$, $\pi_0(c,t)$, and $\pi_2(c,t)$. Similarly, variation across cohorts in the levels of regression coefficients identifies cohort-specific parameters. As illustrated by equations (8.1)-(8.3), 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_{1c}, λ_{2c} 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), Kling (2001)). In addition, Deschenes (2006) found that for males born in the United States between 1910 and 1950 the structural earnings function is practically linear. Thus, under the maintained assumption that $k_c=0$ the parameters capturing comparative advantage selection (λ_{2c}) are readily estimable from the levels of the quadratic schooling coefficients across cohorts and provide a lower bound on the extent of comparative advantage bias.[16]

This reduces the number of parameters by C, but C additional restrictions must be imposed. Provided that an estimate of λ_{2c} is available, the average marginal productivity of schooling for cohort c, b_c, is identified from the linear schooling coefficient in (8.2). Therefore, some restrictions must be imposed on a_c and λ_{1c}. As equation (8.1) makes clear, imposing less restriction on the shape of λ_{1c} 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_{1c} and λ_{2c} across cohorts, combined with the year-specific

[16] Since economic theory predicts that $k_c\geq0$, the estimates derived under the assumption $k_c=0$ will provide a lower bound on the degree of comparative advantage bias across cohorts. To see this, note that from equation (8.3), $\lambda_{2c}=\frac{\pi_2(c,t)}{\delta(t)}+0.5k_c$. Thus λ_{2c} will increase linearly with the curvature parameter k_c. Consequently, the degree of comparative advantage bias will rise as k_c rises if average schooling for a cohort remains constant as k_c changes.

parameters $\delta(t)$ and $\Psi(t)$ are required to make such an assessment. One normalization consistent with this objective is to impose that $\sum_c \lambda_{1c}=0$, and assume no differences across cohorts in the average level of absolute ability (i.e. that $a_c=a$ for all cohorts). This last assumption is not a major concern since a_c does not enter into the linear and quadratic schooling coefficients of the log earnings function (8). For these reasons, I refer to the estimates of λ_{1c} as "relative ability bias".

Given these identifying restrictions it is straightforward to estimate the parameters of the model using OMD. For a given cohort-year pair, the relationship between the estimated regression coefficients and the parameters of the structural earnings function in equations (8.1)-(8.3) can be written as $\hat{\pi}(c,t) = f(\theta_{ct}) + \omega_{ct}$, where:

$$\hat{\pi}(c,t) = \begin{bmatrix} \hat{\pi}_0(c,t) \\ \hat{\pi}_1(c,t) \\ \hat{\pi}_2(c,t) \end{bmatrix} \quad \text{and} \quad f(\theta_{ct}) = \begin{bmatrix} \Psi(t)[a - \lambda_{1c}\overline{S}_c] \\ \delta(t)b_c + \Psi(t)\lambda_{1c}\overline{S}_c - \lambda_{2c}\overline{S}_c \\ \delta(t)\lambda_{2c} \end{bmatrix}$$

and where ω_{ct} is a combination of sampling and specification errors.

B. Results

Table 3 reports the OMD estimates of the year-specific parameters of the structural earnings function for every year between 1979 and 2002. Since these estimates of $\delta(t)$ and $\Psi(t)$ use information on $\pi_0(c,t)$, $\pi_1(c,t)$ and $\pi_2(c,t)$, I label them "full-information" estimates. Again, the estimates corresponding to two different specifications of the log wage regressions are showed. In general, the full-information estimates of $\delta(t)$ and $\Psi(t)$ are very similar to the partial-information estimates displayed in Figures 2 and 3. Again, these show that the year-specific causal return to education (top panel) increased by 42-48%, while the return to unobserved ability (bottom panel) increased by 4-10% between 1979 and 2002.

The estimates of the cohort-specific parameters are reported in Table 4. In addition, the goodness-of-fit statistics associated with each model are reported. Each panel contains a separate series of estimates, corresponding to the two 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_{1c}), while the third column presents the comparative advantage bias ($\lambda_{2c}\overline{S}_c$). 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 2.95% to 4.46%. 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 component $\delta(t)$.

The estimates of the relative ability bias (λ_{1c}) indicate no inter-cohort trend in the correlation between unobserved ability and completed education. Again, because of the normalization the levels of the estimates are not meaningful, only the differences across cohorts are. The entries of Table 4 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 4 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_{2c}, the slope in the cohort-specific relationship between person-specific returns to education and years of education, and $\overline{S}_c$, the average schooling for cohort c. As shown by equation (5), 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) and Garen (1984) who found that individuals pursued comparative advantage in their educational investments. As Table 4 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.[17] In fact, for the cohorts born in the mid-1950s, comparative advantage bias explains approximately 40% of the conventional measure of the return

---

[17] The increasing trend in $\lambda_{2c}\overline{S}_c$ across cohort is solely due to larger correlation between educational attainment and person-specific return to education for younger cohorts (larger λ_{2c}) and not to an increase in average educational attainment across cohorts. As Table 1 showed, average educational attainment actually decreased for cohorts born after 1950.

to schooling.[18] The finding of the greater role of heterogeneity and self-selection in schooling decisions is also confirmed by Angrist, Chernozhukov, and Fernandez-Val (2006) who found that the schooling coefficient in quantile regressions grew the most between 1980 and 2000 at the top of the distribution of log earnings.

C. 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 changes in the wage structure. For convenience, the goodness-of-fit statistics associated with the two-factor model presented in Tables 3 and 4 is reported in row 1. The estimates are from the regressions with the basic set of covariates. 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 less restrictive model in row 1 is reported, as well as the corresponding degrees of freedom and p-values.[19]

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 result indicates that there are across-cohort differences in the causal effect of education at a given point in time. Such differences could arise from difference in school quality across cohorts. The third row tests a model imposing no variation across cohorts in the mapping between educational 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 (λ_{1c}=0) and no comparative advantage bias (λ_{2c}=0). Clearly, these restrictions are easily rejected. This is particularly important since the assumption of homogeneous returns to schooling has been used in several 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

[18] This figure is obtained by comparing the comparative advantage bias for a cohort, $\lambda_{2c}\overline{S}_c$, to the conventional measure of the return to schooling, $\beta_{ols}(c,t)$ (see equation (5)) , averaged over all years in which a cohort is observed in the sample.
[19] The difference between the goodness-of-fit statistics has an asymptotic χ^2 with degrees of freedom equal to the number of additional restriction being tested. See Chamberlain (1984).

skill, in which $\delta(t)=\Psi(t)$ in all time periods. Variants of this model have been used in other studies of the wage structure (see e.g., Card and Lemieux 1996). However, the restriction of a single price of skill is rejected in this application.

In summary, the evidence in Table 5 clearly indicates that the data does not support any further restrictions imposed on the two-factor model. As a consequence, the two-factor model provides a useful starting benchmark for future analyses of changes in the educational wage structure.

D. Interpretation

Taken as a whole, the evidence in this paper points to important conclusions regarding changes in the 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 component of the causal return to education, and not to an increase in the return to unobserved ability. This is shown graphically in Panel A of Figure 4, which displays the conventional estimate of the return to education, from linear log wage regressions for 1979-2002. In addition estimates of $\delta(t)$ and $\Psi(t)$ are also displayed. 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 widening of the educational wage differentials, but the scope of its contribution is limited. As such, the rise in the year-specific component of the causal return to education explains approximately 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.

Panel B of Figure 4 examines the connection between the estimated return to unobserved ability and the patterns of residual wage dispersion. 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 $\Psi(t)$ from the model. The similarly of the trends in the two series is remarkable. The evidence confirms the central role of rising return to unobserved ability 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 4.B reinforces the credibility of the estimates of return to unobserved ability reported in this paper.

Finally, Table 6 shows the observed and predicted changes in the "conventional" measure of the return to education for each of the cohorts in the sample. The observed change is calculated as the difference in return to education in the entry year and exit year from the sample for each cohort. Those years are listed in the bottom of the table. As is evident in the table, the return to schooling increased for all cohorts, but unequally. Cohorts born after 1940 have gained the most, in particular those born between 1940 and 1954. The predicted change is calculated from equation (5) using the estimated parameters of the model. As shown before, the model predicts that for each cohort changes in the conventional measure of the return to schooling over time will arise because of: (i) changes in the causal effect of education, (ii) changes in the ability bias, and (iii) changes in the comparative advantage bias. Table 6 reports each of these components for each cohort.

It is evident that the quality of the model fit varies by cohort. For cohorts born 1940-54 and 1960-64, the model predicts changes remarkably well. For example, the 1940-44 cohort is observed in the sample between 1979 and 2000. During that time period, the cohort-specific Mincerian estimate of the return to education increased by 50%, from 0.058 to 0.087. For this cohort, the model predicts a change of 0.0268 log points, which is relatively close to the observed change of 0.0289 log points. In addition, for the 1940-44 cohort, 60% of the predicted change is attributable to a change in the average causal effect of education ($[\delta(2000)-\delta(1979)]b_{1940-44}$) and 40% is due to a change in the comparative advantage bias ($[\delta(2000)-\delta(1979)]\lambda_{2,1940-44}\overline{S}_{1940-44}$). However, the model performs more poorly for the older and younger cohorts. Part of the problem is that these cohorts are observed in the sample for shorter periods and relatively old or young, and the experience profile may be poorly approximated by the quartic used in the specifications.

E. Sensitivity analysis

In this section, I shortly describe the results of an extensive sensitivity analysis. A fuller discussion, including the tables with the results of the sensitivity analysis, is available in the companion appendix to this paper. Particular attention is devoted to: (i) measurement error in reported schooling; and (ii) functional form assumptions in the structural earnings function.

(i). Measurement error in reported schooling

It is well known that measurement error in the independent variable generally leads to inconsistency of the estimated regression coefficients. For example, in a linear regression, classical measurement error in the independent variable leads to downward bias in its estimated coefficients. 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 (see e.g., Table 9 in Angrist and Krueger 1999). 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.

In the companion appendix I show 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 some knowledge on the sign of the biases and evaluate the effect of measurement error bias on the OMD estimates. I simulated data on schooling and log earnings assuming that the reliability ratio for schooling of 90%. Regression estimates from these simulated data indicated 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 downward 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 the parameters of the model. The main conclusion from this estimation was that in presence of classical measurement error in schooling, the OMD estimate of $\delta(t)$ is biased downward, while none of the other parameter estimates are seriously affected. Thus, the OMD estimates reported in the paper can be viewed as lower bounds.

(ii). Functional form assumptions

The structural earnings function considered in Section 3 assumes that the main effect of education on wages is person-specific, while the quadratic effect is the same for all individuals from a given cohort. In addition, it is assumed that the same time loading factor applies to schooling and schooling squared. I now describe the results of a simulation analysis of the robustness of the approach proposed in the paper to departures from the functional form assumptions embodied in the structural earnings function.

In particular, I simulated data from the following structural earnings function:

(1') $\log Y_{ict}(S) = \Psi(t)a_{ic} + \delta(t)b_{ic}S - 0.5\gamma(t)k_{ic}S^2$

where the curvature parameter k_{ic} is now allowed to be person-specific and correlated with schooling in the same way as a_{ic} and b_{ic}.[20] In addition, a change in the effect of schooling squared to log earnings over time is not restricted to be the same as the effect of a change in schooling. In the simulations I considered different degrees of self-selection (which corresponds to different values for the λ parameters in equation (4)), and different value for the parameters $\Psi(t)$, $\delta(t)$, and $\gamma(t)$.

While the data was generated from model (1'), I estimated the parameters following the same approach as in the paper, that is, assuming the true model was (1). The goal was to gauge how robust this approach is to departures from the assumptions in the paper. The results indicated that the year-specific parameters are basically unaffected from such departures. In particular, I considered models where $\gamma(t)$ was stationary, equal to $\delta(t)$, and larger than $\delta(t)$. Across all specification, the estimates of $\delta(t)$ were always downward biased, by 1-8%. The estimates of $\Psi(t)$ were practically unaffected. The magnitude of the bias increased as the degree of self-selection increased. As such, the results of these simulations provided no clear evidence that the results obtained in the paper are driven by the functional form assumptions employed to identify the models.

6. Conclusion

This paper provides a framework to interpret and analyze changes in the educational wage structure in the United States between 1979 and 2002. A structural model for earnings, with absolute and comparative advantage in the schooling decision of individuals is used to 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 causal effect of education rises. This feature of the changing U.S. wage structure is 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

[20] That is, $(k_{ic} - k_c) = \lambda_{3c}(S_{ic} - \overline{S}_c) + u_{3ic}$, where λ_{3c} is assumed to be positive.

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 40% between 1979 and 2002. By comparison, the conventional measure of the return to education---which combines both the causal and the biases components---increased by about 60%.

The empirical background for this finding is the striking rise in the convexity of the earnings-schooling relationship that affected all cohorts uniformly over the 1980s and 1990s. The results also indicate that all of the increase in the average causal effect of education is due to 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 4-10% between 1979 and 2002, thereby explaining at most 20% of the observed increase in the conventional estimate of the return to schooling. In the context of traditional models of demand for skills that are the foundation of the literature on wage inequality, these findings suggest that education and unobserved ability are imperfect substitutes.

The estimates also imply that there are no differences across cohorts in the mapping between absolute ability and schooling. Therefore, this 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, the 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. 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

Acemoglu, Daron (2002): "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, 40(1): pp. 7-72

Angrist, Joshua D. and Whitney K. Newey (1991): "Over-Identification Tests in Earnings Functions with Fixed Effects," *Journal of Business and Economic Statistics*, 9: 317-323

Angrist, Joshua D. (1995): "The Economic Returns to Schooling in the West Bank and Gaza Strip," *American Economic Review*, 85: pp. 1065-1087

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.

Angrist Joshua D., Victor Chernozhukov, and Ivan Fernandez-Val (2006): "Quantile Regression Under Misspecification, with an Application to the U.S. Wage Structure," *Econometrica*, 74: pp. 565-574

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.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearny (2006): "The Polarization of the U.S. Labor Market," *American Economic Review, Papers and Proceedings*, 96: pp. 189-194

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.

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

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.

Deschenes, Olivier (2006): “Estimating the Effect of Family Background on the Return to Education,” Working Paper, Department of Economics, University of California, Santa Barbara.

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.

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

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

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.

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.

Katz, Lawrence F. and Kevin M. Murphy (1992): “Changes In Relative Wages, 1963-1987: Supply and Demand Factors,” *Quarterly Journal of Economics*, 107: pp. 35-78

Kling, Jeffrey (2001): “Interpreting Instrumental Variables Estimates of the Returns to Schooling,” *Journal of Business and Economic Statistics*, 19: 358-364

Lemieux, Thomas (2006a): “Postsecondary Education and Increasing Wage Inequality,” *American Economic Review, Papers and Proceedings*, 96: pp. 195-1999

Lemieux, Thomas (2006b): “Composition Effects, Wage Measurement, and the Growth in Within-Group Wage Inequality,” *American Economic Review*, Forthcoming

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.

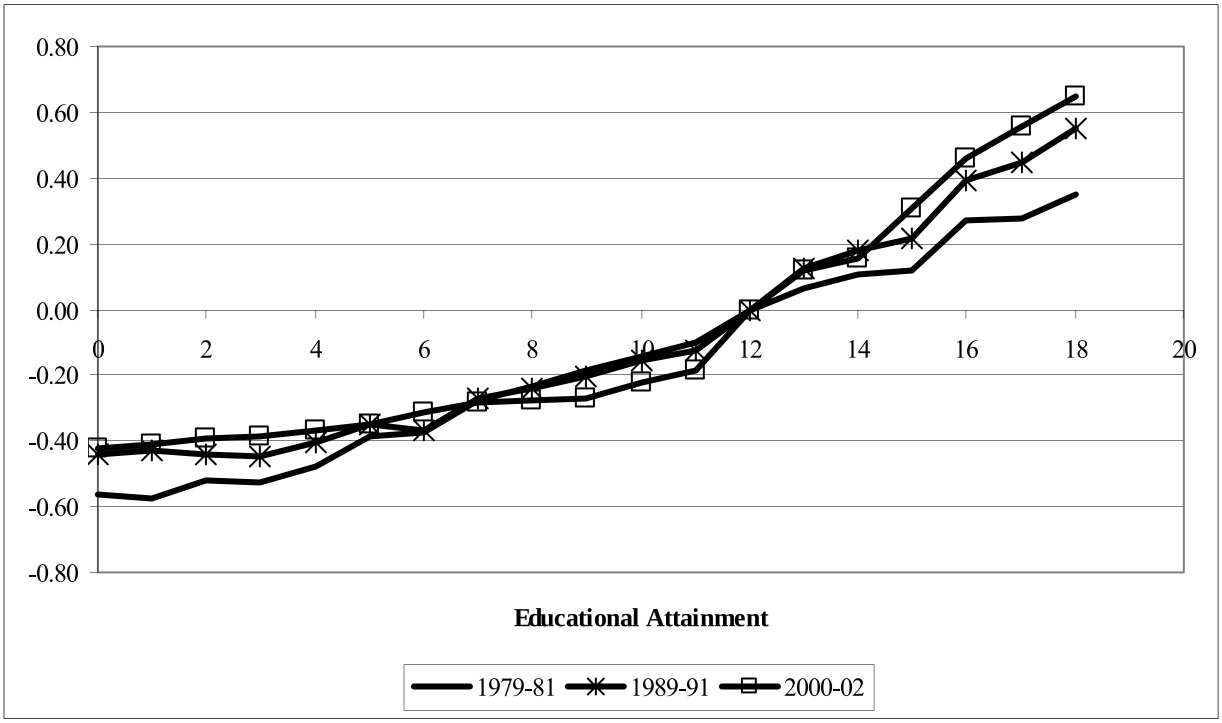
Rosen, Sherwin (1977): "Human Capital: A Survey of Empirical Research," in Ronald Ehrenberg, editor, Research in Labor Economics, volume 1, Greenwich Connecticut: JAI Press.

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

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.

Figure 1: Regression-Adjusted Mean Log Real Hourly Wage by Year of Education, 1979-2002



Note: Mean log real hourly wages are relative to that of high school graduates

Figure 2: Estimate of the Year-Specific Component of the Causal Effect of Education (1979=1)

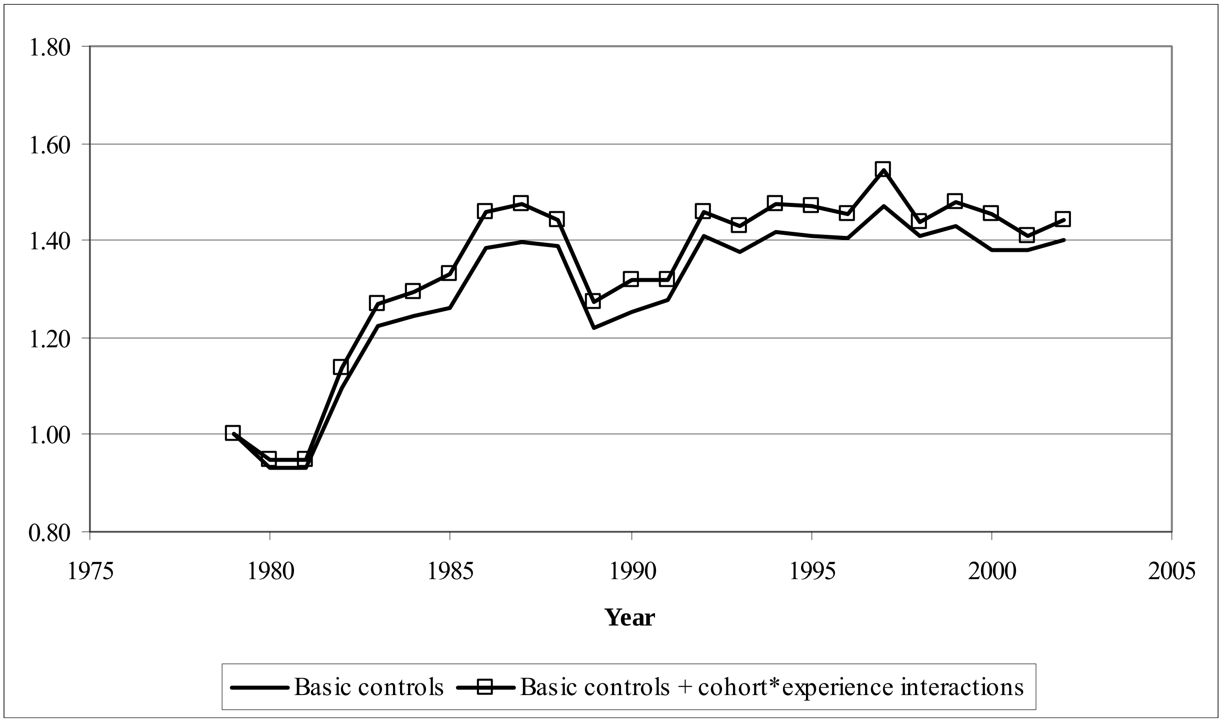


Figure 3: Estimates of the Return to Unobserved Ability (1979=1)

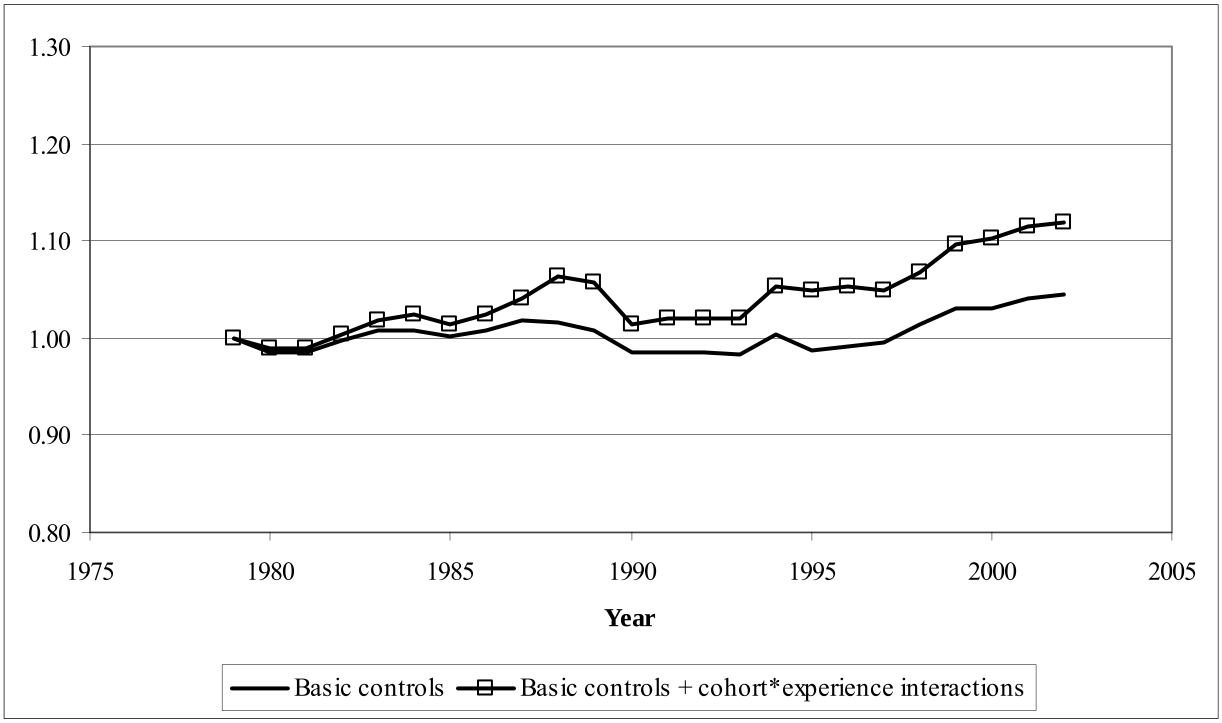
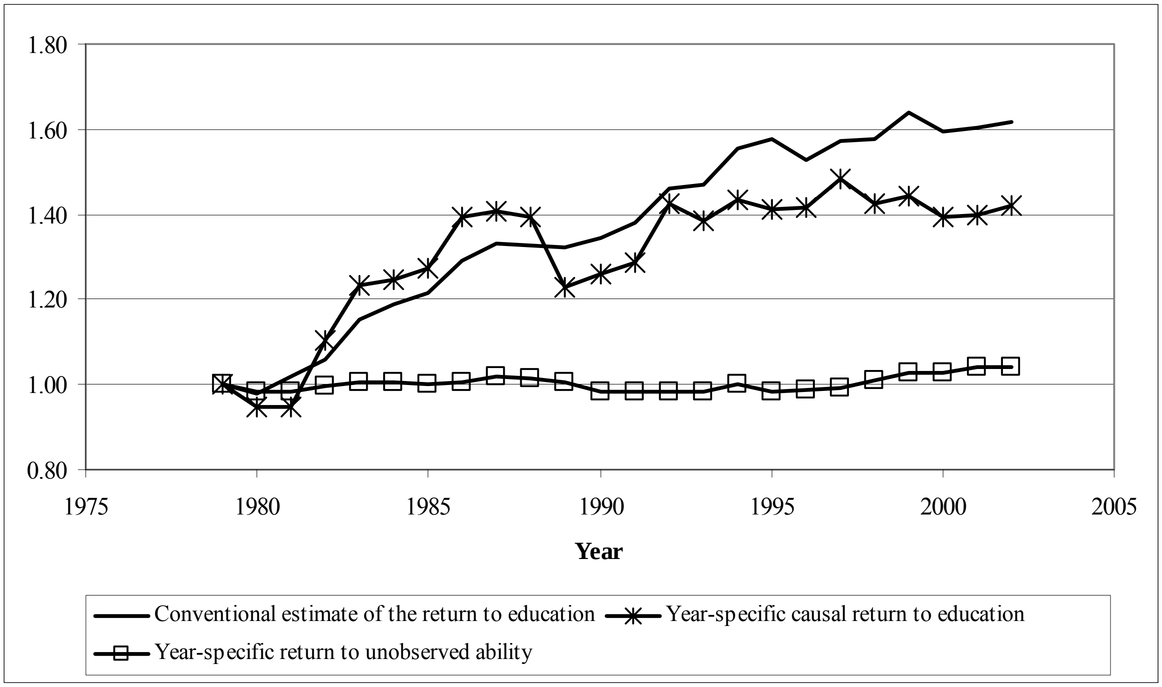


Figure 4: Decomposing the changes in the educational wage structure

A. Conventional measure of the return to education



B. Standard deviations of log hourly wage residuals

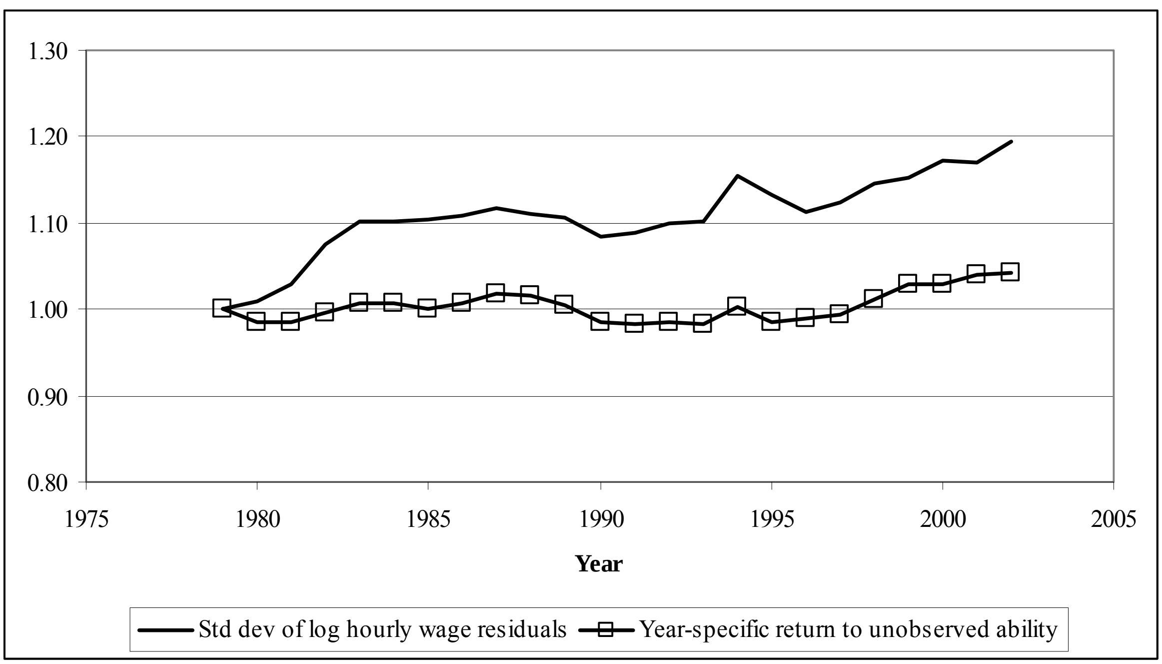


Table 1: Summary Statistics

	Birth Cohort: 1930-34	1935-39	1940-44	1945-49	1950-54	1955-59	1960-64	1965-69
Real Hourly Wage	21.24	21.27	21.49	20.92	19.23	18.67	17.68	17.06
	(15.83)	(16.13)	(16.42)	(15.64)	(14.39)	(14.42)	(13.84)	(13.45)
Years of Education	12.50	12.83	13.27	13.70	13.63	13.42	13.34	13.43
	(3.31)	(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 are parentheses. The sample include men born between 1930-1969, aged 26-60 in the survey year, with real hourly wages ranging between 5-200 ($2002). Only observations in the first rotation out of the CPS samples are included. All statistics are weighted by hours worked in the survey week.

Table 2: Estimates from Quadratic Log Hourly Wage Regressions

	Year: 1979-1981	1984-1986	1989-1991	1994-1996	2000-2002
[A] All cohorts:					
Education	0.0416	0.0259	0.0043	-0.0012	-0.0084
	(0.0028)	(0.0030)	(0.0030)	(0.0032)	(0.0027)
Education squared	0.0007	0.0019	0.0029	0.0035	0.0039
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
[B] Selected cohorts:					
1935-39					
Education	0.0424	0.0208	-0.0096	-0.0018	---
	(0.0121)	(0.0140)	(0.0160)	(0.0283)	
Education squared	0.0006	0.0019	0.0031	0.0031	---
	(0.0005)	(0.0005)	(0.0006)	(0.0011)	
1945-49					
Education	0.0580	0.0458	0.0150	0.0342	0.0206
	(0.0103)	(0.0130)	(0.0149)	(0.0175)	(0.0140)
Education squared	0.0003	0.0015	0.0026	0.0024	0.0029
	(0.0004)	(0.0005)	(0.0006)	(0.0006)	(0.0005)
1955-59					
Education	---	0.0341	0.0226	0.0065	0.0040
		(0.0181)	(0.0149)	(0.0156)	(0.0119)
Education squared	---	0.0017	0.0024	0.0034	0.0036
		(0.0007)	(0.0006)	(0.0006)	(0.0004)

Robust standard errors are in parentheses. The models also include a quartic in experience, and dummy variables for race, ethnicity, marital status, metropolitan area and census division.

Table 3: OMD Estimates of Year-Specific Parameters ("Full-Information" Estimates)

	Year-Specific Return to Education, $\delta(t)$				Year-Specific Return to Unobserved Ability, $\Psi(t)$			
	(1)		(2)		(1)		(2)	
Year:	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err
1979	1.000	---	1.000	---	1.000	---	1.000	---
1980	0.948	0.036	0.959	0.046	0.984	0.003	0.989	0.007
1981	0.948	0.036	0.959	0.046	0.984	0.003	0.989	0.007
1982	1.103	0.041	1.145	0.053	0.997	0.003	0.996	0.008
1983	1.234	0.045	1.288	0.059	1.006	0.003	1.006	0.009
1984	1.245	0.045	1.300	0.059	1.006	0.003	1.012	0.009
1985	1.273	0.044	1.369	0.061	1.001	0.003	1.000	0.011
1986	1.393	0.047	1.483	0.066	1.007	0.003	1.011	0.011
1987	1.409	0.048	1.497	0.067	1.018	0.003	1.025	0.012
1988	1.394	0.048	1.458	0.065	1.016	0.003	1.051	0.011
1989	1.230	0.045	1.283	0.060	1.005	0.003	1.048	0.011
1990	1.261	0.042	1.347	0.058	0.985	0.003	1.006	0.011
1991	1.288	0.044	1.352	0.060	0.983	0.003	1.006	0.010
1992	1.424	0.049	1.499	0.068	0.984	0.003	1.007	0.012
1993	1.385	0.048	1.473	0.067	0.982	0.003	1.004	0.011
1994	1.432	0.049	1.515	0.068	1.002	0.003	1.037	0.012
1995	1.411	0.049	1.520	0.070	0.985	0.004	1.035	0.014
1996	1.417	0.049	1.512	0.068	0.989	0.003	1.039	0.012
1997	1.484	0.050	1.609	0.071	0.993	0.003	1.033	0.012
1998	1.424	0.049	1.491	0.067	1.011	0.003	1.052	0.011
1999	1.441	0.049	1.523	0.068	1.028	0.004	1.083	0.012
2000	1.393	0.047	1.511	0.066	1.028	0.004	1.087	0.011
2001	1.396	0.051	1.444	0.068	1.039	0.004	1.098	0.011
2002	1.419	0.051	1.481	0.069	1.042	0.004	1.104	0.011
Goodness-of-Fit	651.8		586.2		651.8		586.2	
Set of Controls	Basic		Basic		Basic		Basic	
Cohort*Experience	No		Yes		No		Full	

Robust standard errors are reported. Each model has 341 degrees of freedom, with a corresponding chi-square critical value of 385.1

Table 4: OMD Estimates of Cohort-Specific Parameters

	(1) Average Marginal Return (b_c)	(2) Relative Ability Bias (λ_{1c})	(3) Comparative Adv. Bias ($\lambda_{2c}\bar{S}_c$)	(4) Average Marginal Return (b_c)	(5) Relative Ability Bias (λ_{1c})	(6) Comparative Adv. Bias ($\lambda_{2c}\bar{S}_c$)
Birth Cohort:						
1930-34	0.0325	-0.0021	0.0280	0.0446	-0.0202	0.0264
	(0.0031)	(0.0008)	(0.0079)	(0.0043)	(0.0060)	(0.0077)
1935-39	0.0353	-0.0018	0.0266	0.0369	-0.0048	0.0240
	(0.0020)	(0.0005)	(0.0068)	(0.0036)	(0.0055)	(0.0063)
1940-44	0.0404	-0.0011	0.0278	0.0339	0.0073	0.0229
	(0.0018)	(0.0004)	(0.0065)	(0.0034)	(0.0050)	(0.0056)
1945-49	0.0422	0.0000	0.0256	0.0318	0.0000	0.0262
	(0.0016)		(0.0056)	(0.0022)		(0.0058)
1950-54	0.0398	0.0007	0.0273	0.0288	0.0124	0.0241
	(0.0016)	(0.0003)	(0.0055)	(0.0031)	(0.0047)	(0.0051)
1955-59	0.0367	0.0001	0.0340	0.0324	0.0041	0.0295
	(0.0016)	(0.0004)	(0.0067)	(0.0034)	(0.0056)	(0.0061)
1960-64	0.0314	0.0012	0.0408	0.0389	-0.0091	0.0383
	(0.0017)	(0.0005)	(0.0081)	(0.0037)	(0.0061)	(0.0078)
1965-69	0.0295	0.0031	0.0430	0.0317	-0.0020	0.0364
	(0.0020)	(0.0007)	(0.0087)	(0.0036)	(0.0069)	(0.0079)
Goodness-of-Fit		651.8			586.2	
[d.f.]		[341]			[341]	
Set of Controls		Basic			Basic	
Cohort*Experience		No			Yes	

Robust standard errors are in parentheses. Each model has 341 degrees of freedom, with a corresponding chi-square critical value of 385.1

Table 5: Goodness-of-Fit Specification Tests for Alternative Models

1. Two-Factor Model	651.8
[d.f.]	[341]
(p-value)	(0.001)
<u>Specification Tested:</u>	
2. No cohort variation in schooling slope ($b_c = b$)	68.4
[d.f.]	[348]
(p-value)	(0.001)
3. No cohort variation in compar. adv. bias ($\lambda_2^c = \lambda_2$)	136.2
[d.f.]	[348]
(p-value)	(0.001)
4. No ability bias ($\lambda_1^c = 0$)	58.2
[d.f.]	[348]
(p-value)	(0.001)
5. No comparative advantage bias ($\lambda_2^c = 0$)	4003.4
[d.f.]	[349]
(p-value)	(0.001)
6. Stationary return to education ($\delta_t = 1$)	742.6
[d.f.]	[364]
(p-value)	(0.001)
7. Stationary return to unobserved ability ($\psi_t = 1$)	1298.5
[d.f.]	[364]
(p-value)	(0.001)
8. Single return to skill ($\delta_t = \psi_t$)	685.3
[d.f.]	[364]
(p-value)	(0.001)

Note: Under the null hypothesis, the test statistics have a chi-square distribution with 7 degrees of freedom (Rows 2-4), 8 degrees of freedom (Row 5), and 23 degrees of freedom (Rows 5-8). The p-values for the specification tests are in parentheses.

Table 6: Decomposition of the Changes in the Conventional Estimate of the Return to Education, by Cohort

	Birth Cohort: 1930-34	1935-39	1940-44	1945-49	1950-54	1955-59	1960-64	1965-69
Conventional return to education:								
At entry in sample	0.0569	0.0614	0.0577	0.0674	0.0532	0.0750	0.0901	0.0895
At exit from sample	0.0642	0.0628	0.0866	0.0982	0.0893	0.0952	0.1017	0.1182
Change	**0.0073**	**0.0014**	**0.0289**	**0.0308**	**0.0361**	**0.0202**	**0.0116**	**0.0288**
Percent change	*12.8*	*2.3*	*50.2*	*45.6*	*68.0*	*26.9*	*12.8*	*32.1*
Decomposition of change predicted by model:								
Change in causal effect	0.0085	0.0145	0.0159	0.0177	0.0187	0.0054	0.0050	0.0002
Change in absolute ability bias	0.00003	0.00003	-0.00003	0.00000	0.00004	0.00000	0.00007	0.00018
Change in comparative advantage bias	0.0073	0.0109	0.0109	0.0107	0.0129	0.0050	0.0064	0.0003
Predicted change	**0.0158**	**0.0255**	**0.0268**	**0.0284**	**0.0316**	**0.0103**	**0.0115**	**0.0008**
Unexplained change	**0.0085**	**0.0241**	**-0.0022**	**-0.0024**	**-0.0045**	**-0.0099**	**-0.0001**	**-0.0280**
Survey years	1979-1990	1979-1995	1979-2000	1979-2002	1980-2002	1985-2002	1990-2002	1995-2002