Ability Bias and the Rising Education Premium in the United States: A Cohort-Based Analysis

Barış Kaymak
Université de Montréal

I use differences in educational attainment by birth cohorts to estimate the rise in the return to education in the United States. If average ability is similar among nearby cohorts, then differences in educational attainment lead to differences in earnings only if education is productive. The results reveal that (i) the return to a year of schooling increased from 4.8 percent to 8.4 percent between 1964 and 2003, (ii) the ability bias rose from 1.8 percent to 4.7 percent during the same period, and (iii) the acceleration in the education premium after 1980 is explained almost entirely by the rise in the ability bias.

The rise in the education premium has been one of the most striking trends in the United States in the second half of the twentieth century. Between 1964 and 2003, the standard estimates of the return to a year of education, obtained by least squares, doubled from around 6 percent to over 12 percent. The observed education premium is widely argued to reflect not only the causal effect of education on productivity but also the unobserved ability differences among workers that are correlated with education. This correlation creates a wedge, referred to as ability bias, between the actual effect of schooling on a worker's earnings and the observed earnings gap among workers with different education levels (Griliches 1977; Willis and Rosen 1979). Therefore, it is not clear whether the rising education premium reflects a higher return to formal education or a larger ability bias. The distinction is essential for human capital policy. If the higher education premium simply reflects the growing importance of unobserved ability factors acquired much earlier in life, it may be better to divert resources from formal education to preschool education and child care.

This paper estimates the return to education by exploiting variation in educational attainment by year of birth. If average ability is similar among nearby cohorts, then differences in educational attainments of

This paper is based on the first chapter of my thesis at the University of Rochester. I am indebted to Mark Bils and Gordon Dahl for their advice. I thank Joseph Altonji, Ronni Pavan, Uta Schoenberg, Lance Lochner, and two anonymous referees for their valuable comments. I acknowledge support from the W. A. Wallis Institute of Political Economy at the University of Rochester.

[*Journal of Human Capital,* 2009, vol. 3, no. 3]

cohorts lead to differences in earnings only to the extent that education is socially productive.[1] This permits the estimation of the rate of return to education by projecting the average earnings of cohorts on their average educational attainments. I apply this method to accomplish three objectives: (*a*) estimating the effect of schooling on earnings, (*b*) assessing the portion of the rate of return to a year of schooling that should be ascribed to ability bias, and (*c*) evaluating the role of ability bias in the surge of education premium in the United States.

The findings reveal that the return to education increased significantly, from 4.8 percent to 8.4 percent, on average per year of schooling between 1963 and 2003. The corresponding ability bias grew by about 3 percentage points from 1.8 percent to 4.7 percent. In addition, controlling for supply of skills accounts for the decline in the return to education in the 1970s, which displays a steady increase otherwise. The results also suggest that the increase in the return to education due to marketwide conditions, such as rising demand for skill, is partly curbed by decreasing cohort quality, conditional on education. When these cohort trends are accounted for, the return to education increases almost linearly, whereas the acceleration in the observed education premium after 1980 is driven mostly by a larger ability bias.

Year of birth is an important determinant of educational attainment. Since workers usually complete education early in their life before entering the labor market and merely go back to school, the year of birth naturally describes the relative cost and benefit conditions for schooling decisions. For instance, a fall in the cost of education would increase the educational outcomes of only more recent cohorts.

In any case, different sources of variation must be distinguished in the educational attainment by cohort. If all of the variation is caused by changes in the cost of education, estimation of the return to education is straightforward. This certainly is not the case here since the change in the return to education over time, if anticipated, would also affect educational choices, presumably in a different way, for each cohort. This does not prevent identification, however, since at any given year, time-specific changes in the return to education have a symmetric impact on all birth cohorts that is proportional to their educational attainment.

In addition, the cohort differences in educational attainment vary considerably by geographical region. Changes in state-level education policies, education finance, or the differences in the timing of such events are typical factors that generate variation in education levels across cohorts by location. The additional variation generated by the state-birth cohorts also produces more robust estimates and substantiates the estimation design.

[1] For instance, in a pure signaling model in which education is not productive, the cohort differences in average education would not produce differences in earnings.

Do workers of different birth cohorts have similar average ability or potential earnings? Probably not, but this does not necessarily prevent the estimation of the return to education. Smooth, long-term changes in average ability, due to better preschool education or parental care, for instance, are allowed to the extent that they can be controlled by a smooth trend in the year of birth. While year-to-year movements may also occur in potential earnings by cohort, the identification requires that short-term movements in average ability are orthogonal to the average educational attainment.

The literature provides two plausible scenarios in which this requirement could be violated. Labor market conditions at the time a worker graduates are documented as having a somewhat persistent effect on earnings (Freeman 1981; Beaudry and DiNardo 1991; Raaum and Røed 2006). This would prevent identification of the return to education if average educational attainment of a cohort responds to cyclical market conditions. Alternatively, workers of a larger cohort may experience lower wages early in their career if they are not perfectly substitutable for the existing workforce. Welch (1979) and Berger (1985) argue that members of the baby boom generation suffered from reduced earnings growth at the onset of their careers. Similarly, this poses a problem for the estimation method in this paper only if their educational attainment is affected, for example, because of reduced effective school quality due to large classes.[2] To account for these scenarios, I control for the effects of cohort size by education and time and the unemployment rates around the year of graduation. These variables also help control for changes in the supply of skills, which can generate equilibrium effects on our estimates. The addition of these controls has some effect on the estimated return to education for the late 1970s but otherwise has little quantitative consequence, mostly because these variables explain a small fraction of the variation in educational attainment by cohorts.

Other complications may potentially arise when the return to education varies by cohort, conditional on time. Just as for cohort effects in earnings, this is less of a concern for improvements in school quality or changes in cohort quality. Since these developments are likely to be long-term phenomena, they can be controlled by a smooth cohort trend in the return to education. What the estimation strategy does not allow is the year-to-year cohort-specific shifts in the return to education, conditional on time, since they would probably be correlated with average educational attainment. If this correlation is positive, ignoring this variation would yield upward-biased estimates of the return to education, whereas its effect on the estimated changes over time would be ambig-

[2] Higher educational attainment by an average size cohort could also lead to a lower return for the additional years of education due to congestion effects in higher levels of education. I provide a bound on this effect that cannot be directly captured by the size of a cohort.

uous.[3] This is different from a possible correlation between the return to education and educational attainment at the individual level. The latter does not pose a threat to estimation as long as it is constant across cohorts.

Accounting for a cohort trend in the return to education isolates the changes in the return to education that are caused by time-specific marketwide conditions. When skill supplies are controlled, these changes are more likely to be induced by changes in the demand for skill. Allowing for a cohort trend in the return to education yields a steady increase in the return to formal education since 1963. However, the estimates of the ability bias remain stable until 1980 and increase notably afterward.

This paper relates to at least two strands of the literature in labor economics. A number of studies have investigated relative changes in the return to ability and education using data from test scores (Blackburn and Neumark 1993; Murnane, Willett, and Levy 1995; Heckman and Vytlacil 2001). Among these, Heckman and Vytlacil demonstrate in depth that the results are sensitive to different specifications. The panel structure of the data used in these studies, coupled with the strong correlation between education and test scores, prevents a robust identification of the changes in the returns to education and ability, separately over time. On the basis of observed patterns in residual wage dispersion, Card and Lemieux (1996) and Chay and Lee (2000) found the rise in return to ability to be only partially responsible for the rise in the education premium during the 1980s. In an environment with multiple dimensions of ability, wage dispersions cease to be informative (Taber 2001; Gould 2002). Estimating a dynamic model of selection, Taber found that the rise in the education premium in the 1980s was driven almost entirely by changing return to ability. Similarly, Deschenes (2006) relied on a model of selection to infer the relative prices of ability and education using the convexity of earnings as a function of education. In contrast, he found absolutely no role for rising return to ability between 1980 and 2000. The difficulty of disentangling the roles of education and ability in earnings has led researchers to rely on indirect methods to uncover the changes in these roles. Since variation by birth cohorts can be used flexibly during different time periods, this paper contributes to the literature by providing direct estimates of the return to education and how it evolved over time.

An even larger amount of work has been devoted to disposing of the ability bias of the standard estimates of the return to education. A common method is to use variation in educational attainment that is orthogonal to ability (see, among others, Angrist [1990], Angrist and Krue-

[3] If the correlation between cohort-specific return to education and educational attainment is stable over time, then the levels of estimated returns would be biased, but the relative changes would be consistently estimated.

ger [1991], and Card [1995]).[4] Almost all of these studies found much larger returns to education compared to the estimates here.[5] Under more structural restrictions (monotonic treatment response and monotonic treatment selection), Manski and Pepper (2000) derive an upper bound for the average annual return to college of about 9.9 percent. Estimating a dynamic structural model of school choice, Belzil and Hansen (2002) found the average return to education to be 4.7–6.4 percent. In this paper, I find the estimated return to education averaged over the entire sample to be around 5.0–6.5 percent, which is similar to their findings.

Section I discusses some of the factors that have contributed to the differences in educational attainment by birth cohorts. Section II lays out the main specification and discusses the potential complications related to the estimation strategy. Section III presents the benchmark estimation results and investigates the validity and sensitivity of the estimation design. Section IV presents the estimates of the return to education over time. Section V concludes with a discussion of the results.

I. Educational Attainment in the United States

The objective here is not to provide a justification for a single, specific event that separates workers into treatment and control groups with respect to their educational achievement. Provided that the decision environment for education is subject to variation, classifying workers by their year of birth naturally separates workers by the determinants of their schooling choices. In this sense, the variation in educational attainment by cohorts is more mechanical and crude than some of the more finely targeted instruments in the literature, such as changes in the compulsory schooling laws. An important advantage of the current approach, however, is that it can be applied to different time spans as well as different locations. I exploit this flexibility to estimate the changes in the return to education over time. Nevertheless, it may be useful to mention some examples of changes in potential determinants of earnings that are relevant for the workers in the sample.[6]

Figure 1 depicts the average tuition cost of education for the years 1919–95. Average real tuition cost per student increased from $777 in 1919 to over $4,000 in the early 1990s in postsecondary institutions. The time that an average worker had to work to meet the average tuition payment varied from 3 to 6 weeks between 1919 and 1995. Figure 2

[4] See Card (1999) and Heckman, Lochner, and Todd (2008) for a survey of the empirical literature on the rate of return to education.

[5] Although often criticized for relying on weak instruments, the estimates that use quarter of birth as an instrument for education (Angrist and Krueger 1991; Staiger and Stock 1997) are closest to the findings of this paper.

[6] The data used here contain cohorts that were born between the years 1910 and 1968, and the average educational attainment of these cohorts varied from 10 to 14 years in the sample (see table 1 below).

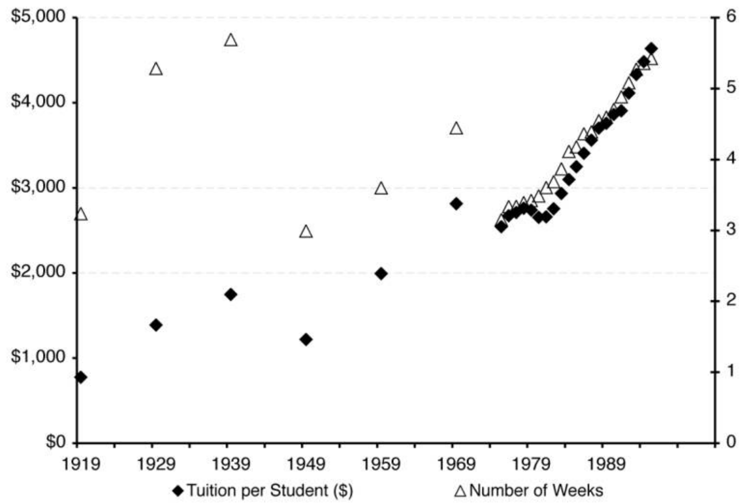

Figure 1.—Average tuition cost. The tuition cost per student is approximated by dividing the total tuition revenues of postsecondary education institutions by postsecondary enrollment. The right scale shows the number of weeks a worker with average earnings would spend to meet the average tuition payment. Tuition amounts are expressed in real 2003 dollars. Source: U.S. Department of Education.

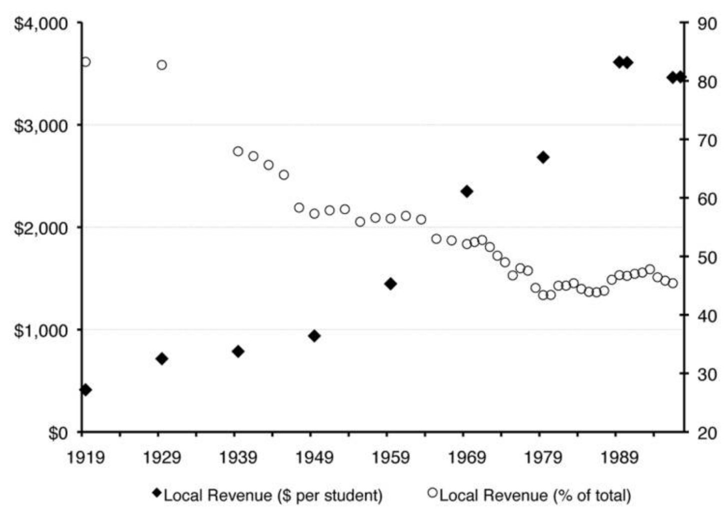

Figure 2.—Local revenues of public elementary and secondary schools. Revenues are expressed in real 2003 dollars. Source: Statistics of State School Systems, U.S. Department of Education.

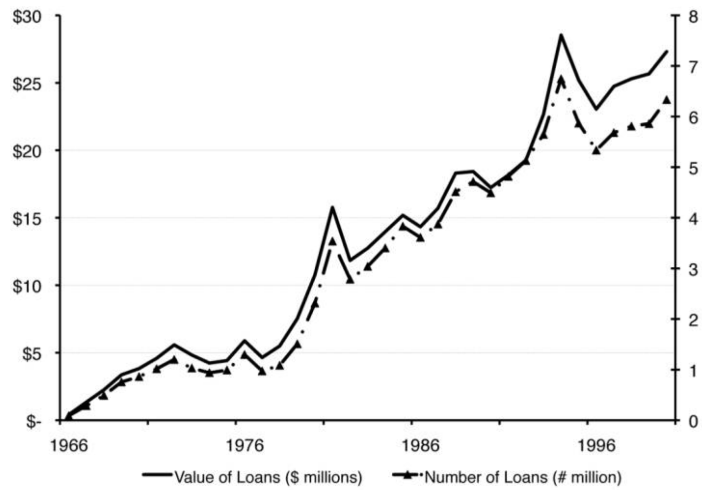

Figure 3.—Volume and accessibility of FFEP loans. Source: National Student Loan Data System, U.S. Department of Education.

shows the revenues of educational institutions obtained from local sources in the form of tuition payments, gifts, and donations. Local revenue per student increased throughout the period, and the total share of local funding for secondary education decreased from around 80 percent to 45 percent between 1919 and 1969, indicating a rise in state and federal funding for education.

The Federal Family Education Program (FFEP) was initiated in 1966, thus available to roughly a quarter of the cohorts in the sample. Since these loans were not awarded on a merit basis, the associated changes in the cost of education are not directly related to ability. Figure 3 displays the total number and value of loans provided under the FFEP. Total subsidies to college increased considerably since the early stages of the program and were made widely accessible.

The importance of institutional factors in educational attainment has been mentioned in the literature. Goldin (1999) emphasized the role of compulsory education laws in the "high school movement" of the early twentieth century.[7] Bound and Turner (2002) argue that the G.I. Bills for the veterans of World War II raised college attendance rates in the mid-twentieth century.

A change in the return to education over time could also generate fluctuations in educational attainment. Shifts in production technology, in favor of skilled workers, would make higher education more attractive.

[7] In fact, variation in educational attainment due to changes in the compulsory schooling laws has been used to estimate the return to education. See, e.g., Lang and Kropp (1986), Angrist and Krueger (1991), or Acemoglu and Angrist (2000) for estimations of the spillover effects of education.

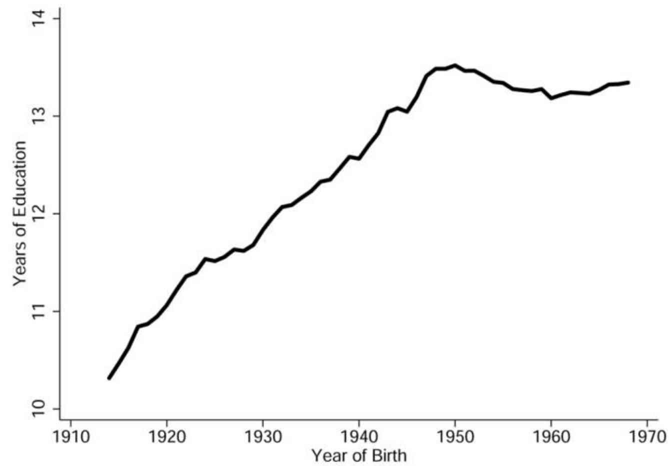

Figure 4.—Educational attainment by year of birth, March CPS

Such changes in skill prices may be less apparent to earlier cohorts. Even when these changes are fully anticipated, educational attainment would respond differently across cohorts because of discounting.

Figure 4 depicts the average educational attainment of birth cohorts among male workers between the ages of 24 and 60 using data from the March supplements to the Current Population Survey (CPS). The average level of education increased steadily earlier in the last century and then stagnated beginning with cohorts born in the 1950s.

The estimation method identifies the rate of return to education by short-term movements. To see the short-term variations in education and earnings by cohort, the variables were projected on a set of cohort indicators, with controls for survey year indicators, race, and a quartic trend in age. The estimated cohort effects were then detrended using a quartic trend in year of birth. Figure 5 shows log weekly earnings and years of schooling by cohort in deviations from the estimated trend. The two variables move closely across cohorts in figure 5*A*, suggesting a significant return to education. Figure 5*B* shows a scatter plot of education-earnings pairs, which outlines the identification of the return to education for the baseline model. The solid line summarizes the fitted values from a linear regression. The slope of the fitted line is 4.70 percent, which is the estimate of the average return to a year of education in this simple specification.

II. Model

Consider the following basic relationship between education and earnings:

$$\ln w_{ict} = \alpha_{ct} + \beta_{ct} s_{ict} + \gamma_{ct} a_{ict} + u_{ict}, \quad (1)$$

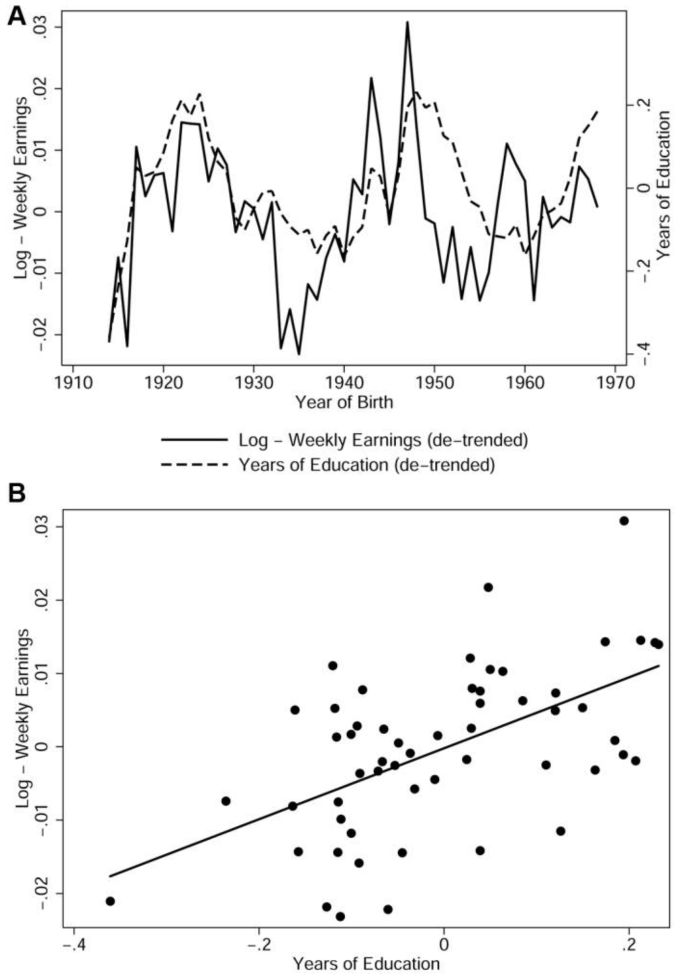

Figure 5.—Estimated cohort effects in education and earnings, March CPS. A quartic cohort trend is used to detrend the series. The reported values also control for a quartic trend in age, race, and survey year.

where i denotes an individual worker, c denotes birth cohort, and t denotes time period. The amount of schooling for individual i from cohort c in year t is denoted by s_{ict}. The term β_{ict} is the return to education and may vary by cohort and time.[8] A worker's age is denoted by a_{ict}. Since α_{ct} captures the variation in log earnings across cohorts over time, the residual term, u_{ict}, is assumed to have a zero mean for each cohort-time cluster. A worker's unobserved ability is reflected in the random error term, u_{ict}, which is potentially correlated with educational attain-

[8] Possible individual variation in β_{ct} is ignored at the moment. Section II.D provides a discussion in greater detail.

ment. This leads to the classical ability bias in the standard estimates of the return to education (Griliches 1970).

To highlight the main issues regarding the identification of the return to education using birth cohorts, fix a year t and consider the average wages of a cohort. Equation (1) gives

$$\overline{\ln w_{ct}} = \alpha_{ct} + \beta_{ct}\bar{s}_{ct} + \gamma_{ct}a_{ct} + \bar{u}_{ct}. \tag{2}$$

Variation in average earnings across cohorts could arise from differences in predetermined earnings potential, captured by the intercept term, α_{ct}; from differences in educational attainment, $\bar{s}_{ct}$; or from differences in the return to education, β_{ct}, for instance, as a result of changes in school quality. Additional differences could be driven by age-related differences in productivity, $\gamma_{ct}a_{ct}$, such as differences in on-the-job human capital accumulation.

At this level of generality, none of the above parameters can be individually identified by data on earnings, education, and age. For instance, in any given year t, cohort effects are empirically indistinguishable from age-related factors unless one is willing to impose functional forms on these effects. The primary concern in this paper is the identification of the return to education and how it changes over time, using variation in educational attainment and earnings by cohorts. This comes at a price for it entails some restrictions on how much the parameters above can vary by cohort. In the following, I begin by assuming that the parameters of (2) are identical across cohorts but may change over time, and then I show how the ability bias can be eliminated to identify the return to education at any time. I provide some plausible scenarios in which these parameters would change by cohorts, and I discuss how far one can go in controlling for these variations without impeding the identification of the return to education. Finally, subsection D evaluates the implications of individual-level heterogeneity in the return to education for our estimates.

A. *Changes in the Return to Education over Time*

Assuming that α_{ct}, β_{ct}, and γ_{ct} are all identical across cohorts, the average earnings of a cohort in year t is

$$\overline{\ln w_{ct}} = \alpha_{t} + \beta_{t}\bar{s}_{ct} + \gamma_{t}a_{ct} + \bar{u}_{ct}. \tag{3}$$

The first thing to notice in equation (3) is that while the individual ability/error term, u_{ict}, is correlated with educational attainment, *average* ability, $\bar{u}_{ct}$, does not show any variation across cohorts (other than the variation due to sampling error); that is, it is orthogonal to *average* educational attainment. Therefore, in the absence of cohort effects on earnings, the return to education, β_t, can be identified by evaluating (3) using data grouped by birth year.

Second, since (3) can be evaluated for each year, the change in the return to education can also be identified. Aggregate variation in the return to education or in average earnings over time caused by the changing supply of or demand for skills does not interfere with the estimation of the return to education. Such economywide changes have a symmetric effect on all cohorts in a given year, in proportion to their educational attainment.

A few words are in order to explain the actual implementation of our estimation. First, while the group-based estimator is consistent, a mathematically equivalent but more efficient method is to estimate (3) at the individual level by using binary year of birth indicators as instruments for educational attainment (Angrist 1988). Moreover, estimation of changes in the return to education, with respect to observable covariates, has been extensively analyzed in the econometrics literature in models with correlated random coefficients (Wooldridge 1997, 2003; Heckman and Vytlacil 1998). The estimators provided in the literature can be applied to our case with proper adjustments. A relatively more robust procedure is outlined in Wooldridge (2003) and employed in this paper. In particular, let $\mathbf{x}$ be a set of covariates for the return to education, such as a time trend or dummy variables for the survey year, that we wish to interact with education. Let $\mathbf{Z}$ be a set of year of birth indicators. The three-step estimator first estimates a reduced form for s_{ict} by regressing it on $\mathbf{x}$ and $\mathbf{Z}$ along with other control variables in the model to obtain the fitted values, $\hat{s}_{ict}$. The return to education and its interaction with $\mathbf{x}$ then can be estimated by two-stage least squares (TSLS), where $\hat{s}_{ict}$ and $\hat{s}_{ict}(\mathbf{x} - E[\mathbf{x}])$ are used as instruments for s_{ict} and $s_{ict}(\mathbf{x} - E[\mathbf{x}])$, respectively (Wooldridge 2003).

This procedure can be applied directly when $\mathbf{x}$ includes a smooth time trend or binary time indicators. In the next section I estimate both of these alternatives: first with a quartic function of time and then with 10-year dummy variables. Bunching several years requires the estimation to be reweighted so that the estimate of the return to education for each decade is consistent for the average return during these 10 years, $\sum_{t_0}^{t_0+10}\beta_t/10$. This is necessary because birth cohorts enter and exit the sample in different years, causing the average return to be calculated over different time spans for each cohort. In addition, the relative sampling weights of cohorts change over time leading to a weighted average of β_t's that is cohort specific. Appendix A shows that a properly reweighted estimation delivers the *unweighted* average return over time. This essentially requires, for each 10-year interval, that only those cohorts that appear in each and every one of the 10 years are used, and each observation is weighted so that the average 10-year return calculated for each of these cohorts is the same.

B. *Cohort Effects in Earnings and the Return to Education*

Eliminating the ability bias in the previous section relies on the assumption that average predetermined earnings potential, α_{ct}, does not vary across cohorts, given t. This assumption could be too restrictive, however, since improvements in early child care and parental education and changes in the quality of preschool programs would lead to differences in ability across cohorts, which would be reflected in the cohort-specific intercept. Since the average education of cohorts increases over time, taking the average of earnings by cohort would not eliminate the ability bias. Nevertheless, if such developments happen slowly, over an extended period, then they are of less concern for identifying the return to education. In the next section, a smooth trend in year of birth is included in the regression to capture these changes. Therefore, the return to education is identified by *short-term* movements in average earnings of a cohort in response to their average education.

One could argue that year-to-year movements take place in earnings potentials of cohorts, which cannot be captured by a smooth trend. For instance, workers who belong to a particularly large cohort may experience lower earnings (Welch 1979). Similarly, Freeman (1981) points out that workers who enter the labor market during an economic downturn experience unfavorable starting salaries. While plenty of reasons could be considered for the short-term cohort variation in earnings, these effects do not necessarily impede the identification of the return to education unless they are correlated with educational attainment. Therefore, a more precise statement for the identification requirement is that *year-to-year changes in cohort effects around a smooth trend are not correlated with average educational attainment of a cohort.* In Section III.C, I investigate the plausibility of this requirement by explicitly controlling for these effects. Overall, the cohort effects mentioned in the literature turn out to be quantitatively unimportant, albeit statistically significant.

Cohort effects may be present not only in the intercept. Potential improvements in the quality of the education system and schools would be reflected in the return to education, β_{ct} (Card and Krueger 1992; Heckman, Layne-Farrar, and Todd 1996). In general, arbitrary variations in the return to education by cohort could possibly lead to a correlation between average educational attainment, $\bar{s}_{ct}$, and β_{ct}, which would undermine the consistency of the estimates presented here. One can control for the slow changes in the quality of schools by interacting a smooth trend in year of birth with educational attainment. This way, a more subtle distinction can be seen between the rise in the return to education that is due to higher school quality and the rise therein due to changing economywide conditions, such as higher demand for skill. Even in the absence of any economywide shifts in the demand for skilled labor, better schools would lead to a rise in the measured return to education over time. This would happen slowly, not only because these developments usually occur over a long period but also because each entering cohort of workers constitutes only a small fraction of the labor force. Controlling for a cohort trend in the return to education allows us to separate relatively sudden changes in the return to education that are likely due to changing market conditions from one year to another.

If the average return to education varies by cohort, conditional on time, but is not correlated with average educational attainment of a cohort, then the estimation by cohort means is still consistent for a weighted average of these returns. The assigned weights are higher for the returns of the cohorts with larger deviations from the overall average educational attainment. This particular average corresponds to the average realized contribution of education to the marginal product.

When average educational attainment of a cohort is correlated with the cohort-specific return to education, estimation by cohort means is no longer consistent.[9] If the concern is that cohorts with higher returns to education attain higher levels of education, the estimates of the return to education provided here are biased upward and should be considered as bounds. Whether or not the differences between these estimates over time are biased is a less trivial question. If the correlation between a cohort's average educational attainments and the cohort-specific return is constant over time, one would expect the bias to be constant over time as well. This would imply that the bias in the estimates of the *changes* in the return to education over time is of second order.

C. *The Age-Cohort-Time Triangle*

The fact that one's (current) age is the difference between the current year and the year of birth has haunted studies that attempt to empirically distinguish between the effects of these three variables on earnings. Any earnings regression will limit some of these effects by assumption, even when we do not particularly care about estimating them. This paper is not an exception in this regard, but the way in which this issue manifests itself is somewhat different. This subsection aims to pinpoint where this paper stands in this Bermuda triangle of earnings regressions.

Because of the linear dependence, these three variables span a two-dimensional space. Upon observing two dimensions, one cannot infer measures about the third dimension. Technically, for the current estimation strategy to work, it is sufficient to assume that these three effects in earnings span a *strict* subset of the two-dimensional plane. In other words, when one allows for full age and time effects, all possible cohort variation is absorbed. As long as some cohort variation is left out of the earnings equation (1), one meets the necessary rank conditions to es-

[9] A correlation between the return to education and educational attainment at the individual level does not imply that this correlation will also be reflected in the grouped data. See Sec. II.D for a discussion.

timate the cohort variation in educational attainment and, hence, the return to education. Note that we are not interested in separately identifying the effects of age, time, and birth year on earnings here. The particular restriction imposed here is that the age and cohort effects on earnings can be approximated by smooth functions. No restrictions are imposed on time effects since they are more likely to display short-term variations.

A similar problem arises when we estimate changes in the return to education, β_{ct}. Nevertheless, the focus is now on estimating the time effects per se. Potentially, one could imagine that unrestricted age, cohort, and time effects are also present in the return to education. Of course, these are not separately identifiable; therefore, it is necessary to restrict the way these variables affect β_{ct}. This paper first assumes that β_{ct} varies only by year, and then this assumption is relaxed by allowing cohort or age effects in the return to education with a smooth function. If, instead, one is interested in short-term changes in the return to education over the life cycle (caused by age and not time), one would at least need to assume that cohort and time effects are smooth. Nevertheless, I see no reason to believe that age effects, for instance, due to human capital accumulation on the job, would display sudden changes.[10]

D. Individual Heterogeneity in the Return to Education

Consider arbitrary variations in the return to education across individuals. Suppose that the individual return to education can be decomposed into a cohort-time mean and an individual deviation from this mean: $\beta_{ict} = \bar{\beta}_{ct} + b_i$. One could also express the educational attainment in individual deviations from the cohort-time averages: $s_{ict} = \bar{s}_{ct} + \nu_{ict}$. When these are substituted into equation (1), average earnings of a cohort in year t is

$$\overline{\ln w_{ct}} = \alpha_{ct} + \rho_{ct} + \bar{\beta}_{ct}\bar{s}_{ct} + \gamma_{ct}a_{ct} + \bar{u}_{ct}, \tag{4}$$

where $\rho_{ct} = E[\beta_{ict}\nu_{ict}|c, t]$. The identification of the return to education is granted if the restrictions outlined earlier for α_{ct} can be extended to include ρ_{ct}, because that is the only difference between equations (2) and (4). This new term essentially reflects the correlation between the individual variation in the return to education and the workers' schooling choices. If workers with higher educational attainment have higher (marginal) returns to education, then this term is positive. This correlation is not a problem by itself since it would simply be reflected in the constant term. Moreover, it is not a concern even if this correlation

[10] Sudden changes might exist at the worker level as a result of a change of jobs, occupations, employment status, etc., but these are averaged out when the data are grouped by cohorts.

changes over time, just as time variations in α_t are allowed. Time variation in ρ_{ct} can occur if workers of a given cohort are selected differently over time into the labor force. Given the discussion in the previous section, even smooth changes in ρ_{ct} by cohorts do not pose a significant threat to estimating the return to education. As long as ρ_{ct} and $\bar{\beta}_{ct}$ do not display sudden swings between nearby birth cohorts, the return to education can be estimated as before.

Although the identification requirement is the same, new complications emerge when individual heterogeneity exists in the return to education. In particular, changes in the *distribution* of educational attainment from one cohort to another now become critical for consistency and the interpretation of our estimates. A difficulty brought about by individual heterogeneity is that the estimation by cohort means can result in cohort-specific returns to education even when the distribution of β_{ict} among workers is the same for all cohorts. This could happen, for instance, when the change in the distribution of educational attainment across cohorts is led by a particular group of workers. Assume that workers with higher ability, u_{ict}, have on average higher β_{ict}. This alone is not a problem. But now suppose that the changes in education across cohorts are caused by the top 10 percent of workers in the ability distribution obtaining progressively higher levels of education. In this scenario, evaluation of (4) using cohort means would yield an estimate of the average return to education, conditional on u_{ict} being in the top 10 percent. Essentially, this is a local average treatment effect (LATE; Angrist and Imbens 1995). Estimation of a meaningful LATE in our case requires that the changes in schooling across cohorts are monotonic in the sense that the cumulative distribution of educational attainment must be well ordered across cohorts. In the next section, I show that the change in educational attainment over the sample was indeed monotonic, and I discuss the underlying population that identifies the estimates presented here.

Since the focus here is on the change in the return to education, one must be careful about possible changes in the underlying population that identify the return to education. For instance, if a strong positive correlation exists between the return to education and the level of education, the changes in educational attainment among recent cohorts are likely to be driven by workers with particularly low returns. In this case, the cohort-based estimation would underestimate the rise in the education premium. However, if significant diminishing returns to education exist for each individual (as in Card [2001]), a negative correlation would be seen between marginal returns and educational choices at the optimum.[11] The conclusion above would then be reversed, and one would overestimate the rise in the return to education.

[11] This requires that the return to education is strongly diminishing in the level of education relative to the variance of β_{ict} within a cohort.

I handle these concerns in two ways. In the next section, I first explore the possibility of diminishing returns to education. I find that the return to education is approximately linear in the level of education. Second, the method here can account for smooth changes in the return to education by cohort. In Section IV, I find some support for the positive sorting story above and that the rise in the return to education is in fact even larger.

III. Data and Estimation Results

The data were taken from the annual March supplements to the CPS for the years 1964–2003 and from the decennial census surveys (DCS) for the years 1960–2000. The sample is restricted to men between 25 and 60 years of age. A birth cohort is defined by the group of workers born in the same year. For each birth cohort, at least 10 years of observations for the CPS and two observations for the DCS are required. Hence, the sample includes cohorts born between 1914 and 1968 in the March CPS and between 1910 and 1965 in the DCS.

Educational attainment is measured by the number of years of schooling. Wages are measured by weekly earnings, calculated by dividing the annual wage and salary earnings by the total number of weeks worked during the year. I drop all workers with less than half the minimum weekly earnings, based on a 40-hour week and the minimum wage in 2003. Workers who spent less than 13 weeks at work during the year before the survey are also dropped. Further details on the data are provided in Appendix C.

Table 1 shows the descriptive statistics. Column 1 displays the frequency of cohorts in the sample. Owing to the revolving nature of cohorts, the earliest and the latest cohorts constitute smaller fractions of the sample. The average age of the youngest cohort in the sample is 30, and the average age of the oldest cohort is 57. The average level of education increases steadily until the latest cohorts in the sample. Column 4 displays the average log weekly earnings. Since the age compositions of cohorts in the sample are different, the earnings of early cohorts are bound to include the return to experience. For a meaningful comparison, weekly earnings are projected on a complete set of age, time, and race dummies. Reported earnings are the predicted values, assuming that all observations are white males of age 40 and that they are all subject to an average time effect.[12] A preliminary comparison of the average weekly earnings with the average educational attainment in column 3 suggests a positive relation between education and productivity. Each pair of cohorts could be used to obtain a Wald estimate of the rate of return to education. For instance, the two oldest cohorts

[12] This procedure was applied only to weekly earnings reported in table 1 and not to the dependent variable in the regressions.

TABLE 1
DESCRIPTIVE STATISTICS

Birth Cohort	Percent Frequency (1)	Average Age in Sample (2)	Average Education (3)	Log Weekly Earnings[a] (4)
	A. CPS March Supplement, 1964–2003			
1959–68	18	33	13.26	6.91
1949–58	26	37	13.39	6.96
1939–48	24	41	13.03	7.01
1929–38	17	44	12.14	6.97
1919–28	12	49	11.42	6.92
1914–18	3	53	10.65	6.87
Total	100	40	12.70	7.01
	B. Decennial Census, 1960–2000			
1960–65	11	33	13.40	6.89
1950–59	23	37	13.51	6.93
1940–49	23	41	13.18	7.00
1930–39	18	42	12.11	6.98
1920–29	16	45	11.27	6.94
1910–19	9	50	10.31	6.88
Total	100	40	12.53	6.95

Note.—Data come from the March CPS and the DCS. Sample includes men of ages 25–60. Average wage inflation is used to calculate real income in 2003 prices.
[a] Reported weekly earnings are corrected for age, survey year, and race differences.

yield a return of 6.4 percent in the CPS sample, the next-oldest pair yields a return of 7.3 percent, and so forth. These simple Wald estimates are not conclusive since an underlying trend in average ability would make it harder to separate the marginal effect of education.

A. *Benchmark Estimates*

I now turn to the benchmark specification. A quartic trend in year of birth is included to capture the long-run trends in ability and other predetermined components of wages.[13] Other controls are a quartic function of age to capture the return to experience and dummies for race and the survey year. The working assumption is that the average ability is constant across cohorts, conditional on a smooth trend, and that no other cohort effects are present on wages that are correlated with educational attainment. We can then use indicators for the year of birth as instruments for educational attainment.[14]

Table 2 displays the benchmark estimation results. In the March CPS

[13] On the basis of fig. 4, a trend break in the linear component was allowed beginning with the workers born in 1947. The benchmark estimate changed by less than 0.04 percentage points in this case. Using a quadratic time trend instead of a quartic trend produces an estimate of 4.7 percent.

[14] Since linear age and cohort components are in the main equation of interest, I drop the first two time indicators. Then three cohort dummies were dropped in the first stage to avoid collinearity. The first three cohorts are combined in the intercept; however, the results are robust to alternative combinations.

TABLE 2
RETURN TO EDUCATION: BENCHMARK SPECIFICATION

	March CPS		Decennial Census			
Explanatory Variable	LS (1)	IV (2)	LS (3)	IV (4)	LS (5)	IV (6)
100 × educ	8.18	4.40	7.50	3.99	7.10	4.60
	(.16)	(1.32)	(.17)	(.83)	(.17)	(.34)
Instruments	Year of birth		Year of birth		Year of birth × state of birth	
First-step Wald test	31.82		202.65		72.26	
Wu-Hausman test	8.32*		18.56*		76.91*	
Observations	943,760		5,227,984		5,227,984	

Note.—All specifications control for a quartic age trend, race, and survey year. Columns 1–4 also include a quartic trend in year of birth. Columns 5 and 6 include a quadratic trend and allow the linear component to vary by state. Additional controls in cols. 5 and 6 are indicators for state of birth and state of residence. The instruments are indicators for year of birth (interacted with state of birth in col. 6). All standard errors are clustered by cohort and are reported in parentheses. Data come from the March CPS (1964–2003) and the DCS (1960–2000). The sample includes men of ages 25–60.
* Significant at the 1 percent level.

sample, the least-squares (LS) estimate of the return to a year of schooling is 8.2 percent, which is comparable to the estimates found in the literature. However, the instrumental variable (IV) estimate of the return to education, reported in column 2, is 4.4 percent with a standard error of 1.3 percent. This implies an ability bias of 3.8 percent. The model is estimated under three different structural specifications of the covariance matrix, where observations are clustered only by cohorts, only by survey year, and by cohort and survey year interactions, respectively. The standard errors reported in table 2 are obtained under the specification that allows arbitrary correlations within cohorts and are the highest of the three clustering methods.[15] Even with these conservative errors, the Wu-Hausman test statistic for the difference between the two estimators is 8.3, suggesting that ability bias is statistically significant.

The estimates from the DCS in columns 3 and 4 further confirm the result. These results are reported to check the robustness of results across data sets, improve the efficiency that is brought about by the higher sample size in the DCS, and provide a basis of comparison between the benchmark specification and the more general specification in columns 5 and 6. The estimated rate of return is 4.0 percent per year of education, which is close to the March CPS estimate. The standard error of the IV estimate is 0.8 percentage points.

The educational choices of individuals are likewise affected by regional factors such as local institutional arrangements on education finance, policies on compulsory education at the state level, or simply the availability of school options. Changes in these local factors would

[15] The standard errors obtained with survey year clusters and cohort × survey year interaction clusters are 0.77 and 0.56, respectively.

affect schooling choices of a subgroup of workers in a birth cohort and create additional variation in education across cohorts by state or region. If one's state of birth captures these local conditions, then the return to education can be identified using birth cohorts by states.[16]

The state of birth need not capture the cost and benefit conditions of a cohort in *that* state. In a setting in which individuals are mobile, state of birth may affect the set of educational choices for a cohort. Differences in costs of mobility or taste would generate an association between an individual's place of birth and educational outcome.

In this more general specification, the regional labor market conditions are captured by indicators for the state of residence. The specification includes a quadratic trend in the year of birth with a varying linear term across states to allow for regional time trends in the predetermined component of wages.[17] Finally, the dummy variables for the state of birth are included in the main regression; therefore, the estimate of the return to education is still identified only by differences across birth cohorts. Instruments are interactions of state of birth indicators with the year of birth indicators.

The estimation results for the DCS are reported in columns 5 and 6.[18] The IV estimate of the return to schooling in this case is 4.60 percent, which is very close to the estimate obtained in the March CPS. The standard error of the estimate is 0.34, which is substantially lower than the previous specifications, thanks to the additional variation captured by state cohorts. The Wu-Hausman test rejects the hypothesis that the IV and the LS estimates are similar at the 1 percent significance level.

The benchmark estimates indicate that the ability bias is significantly positive. This is in contrast with the earlier IV studies that usually found much higher returns to education, arguably with the exception of Angrist and Krueger (1991), who found the return to education to be similar to the LS estimate. While their estimates were discounted for relying on weak instruments and hence being centered around the LS estimate, this is certainly not a valid point here. First, the estimates are considerably lower than the LS estimate; therefore, the presence of a bias would mean that the true return is even lower than the reported estimates. Second, the *F*-statistics for the first-stage regression are significantly high. Even for the least precise sample in table 2, the *F*-statistic for the cohort dummies is 31.82. The critical values reported in Stock and Yogo (2005) confirm that any potential TSLS bias is far less than 5 percent of the ordinary least squares bias. To show that the results here do not come from differences in the sample, Appendix B compares the estimates here with those that use variation by quarter of birth to estimate the return to education.

[16] With a similar motivation, Angrist and Krueger (1991) exploit differences in age-related schooling requirements across states to improve the precision of their estimates.

[17] Allowing for varying curvature does not alter the results.

[18] The state of birth is not available for the March CPS data.

B. *Return Heterogeneity and Local Average Treatment Effects*

Could heterogeneity in the return to education explain the low estimate obtained here? This is possible, for instance, if the return to education varies across workers and the differences in educational attainment between cohorts are driven by workers with particularly low marginal returns to education. Another possibility is a changing skill premium over time. Since birth cohorts are observed during different time periods, the estimates using earlier cohorts as instruments will give different results relative to later cohorts if the return to education is not constant over time. In fact, Hausman's *J*-statistic for overidentification is 192.4 (188.2 for the DCS sample) for the benchmark specification; that is, different (sub)sets of instruments yield significantly different estimates of the return to education. Next, I address the individual heterogeneity in the return to education and then discuss the time variation in the return to education in Section IV.

1. Response Monotonicity

When the return to education varies across individuals or schooling levels, the IV estimate is identified by the underlying subpopulation that is affected by the variation in the instrument. A necessary condition for estimation of a sensible LATE is that the changes in the distribution of educational attainment are monotone across cohorts.

Figure 6*A* plots the empirical cumulative distribution functions (CDF) of educational attainment by birth cohort. The distribution functions are well ordered; that is, the educational choices change in the same direction within a cohort. Figure 6*B* displays the difference in the cumulative distributions of two consecutive birth cohorts. None of the lines cross the zero line, indicating a monotonic change in educational attainment across groups. To better test the robustness of the model for selection of cohorts based on monotonicity, the exercise in figure 6 was replicated on an annual basis, controlling for the covariates and including a smooth trend in birth year. Excluding all cohorts that experience a nonmonotonic change in educational attainment[19] resulted in an estimate of 4.3 percent, which is similar to the benchmark estimate. This is mainly because these cohorts are assigned a lower weight by the IV estimate since they also display relatively small changes in the educational attainment.

Figure 6 is essential for further understanding the workers that iden-

[19] Let $F(s, c)$ denote the empirical cumulative density for schooling level s and cohort c. A cohort c was excluded if

$$\frac{\sum_s \mathbf{1}(F(s,\ c) - F(s,\ c-1) > 0) \times (F(s,\ c) - F(s,\ c-1))}{\sum_s |F(s,\ c) - F(s,\ c-1)|} \in [0.1,\ 0.9],$$

where $\mathbf{1}(x)$ is an indicator function that equals one if x is satisfied, i.e., if the change in one direction failed to constitute more than 90 percent of the total absolute change.

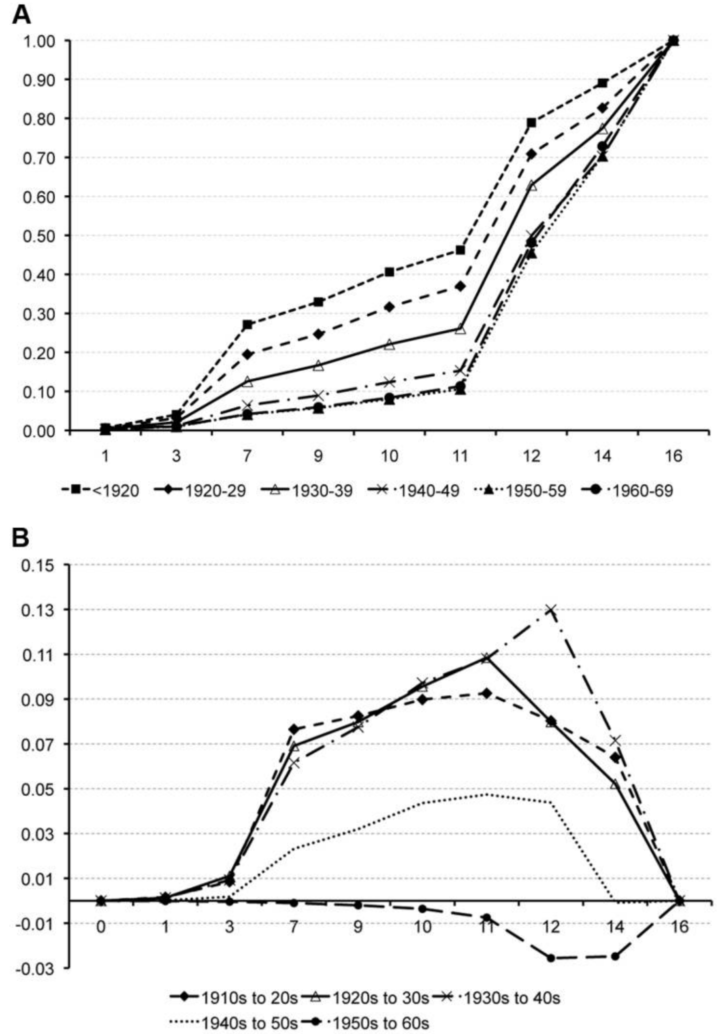

Figure 6.—Distribution of educational attainment by birth cohorts: men born between 1914 and 1968. *A*, Schooling CDF. *B*, Differences in schooling CDF. Data come from the March CPS.

tify the IV estimate. The differences between the cumulative distributions in figure 6*B* are similar for cohorts that were born between 1914 and 1939. A Wald estimate of the return to education using these cohorts would effectively be identified by workers with 7–16 years of education. This indicates that the estimated return has little to say about less than 7 years of education, but at the same time, it is not concentrated on a specific level of education either. For cohorts born in the 1950s and the

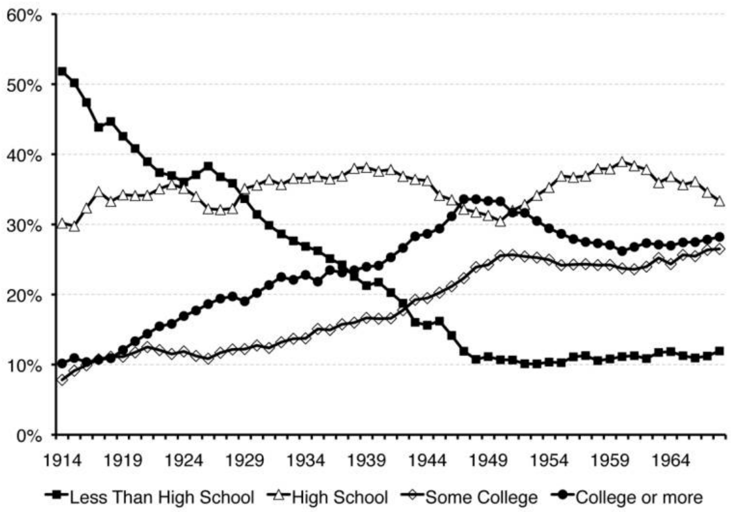

Figure 7.—Educational attainment by year of birth: percent of workers. Source: March CPS.

1960s, the change in educational attainment is concentrated on workers with 14–16 years of education. If significant scale effects were present in the return to education, the estimate based on these cohorts would be different; however, excluding these cohorts does not change the results reported in table 2. The reason is that the return to education is mostly constant across schooling levels, as will be discussed in more detail shortly.

Could it be that the change in educational attainment was led by low-ability workers? Figure 7 displays the change in the educational attainments of cohorts by four schooling categories.[20] For the earliest cohort in the sample, 52 percent completed at most eleventh grade, 30 percent are high school graduates, and only 10 percent are college graduates. The fraction of workers who have less than a high school education declines to around 10 percent in 35 years, accompanied by a rise of 23 percentage points in the share of college graduates and 16 percentage points in the share of workers with some postsecondary education. Assuming that marginal return and ability are strongly positively correlated, the top 10 percent of the ability distribution has always been college graduates, and the bottom 10 percent has always attained less than 12 years of schooling. The benchmark estimates are identified by the middle 80 percent of the ability distribution. Unless the estimates

[20] The four categories are defined as less than high school, high school if the worker completed the twelfth grade, some college for 13–15 years of completed education, and college or more if the worker has 16 or more years of education.

in the literature correspond particularly to the top 10 percent of the ability distribution, the difference in estimates cannot be explained by differences in the underlying populations.

2. Quadratic Specification

Although Mincer (1974) found a linear relationship between log earnings and years of education, the observed wage-schooling profiles in the data have become more convex over the years (Heckman et al. 2008). However, cross-country evidence suggests a negative relationship between the LS return to schooling and years of schooling, indicating a concave production function. The curvature of the human capital production function has important implications for the effect of schooling on differences in growth of income across countries, for differences in educational attainment, and for wage-experience profiles. I extend the benchmark specification by adding a quadratic term in years of schooling. Consider the version of the model in Section II.D with individual heterogeneity in β_{ict} and add a quadratic term in education (deviated from mean education), which carries a constant coefficient for all workers. The mean return can be estimated if (i) the mean and the variance of β_{ict} are independent of the year of birth and if (ii) s_{ict} is linear in b_i, the individual deviation from the mean return, conditional on covariates. Furthermore, the coefficient of the quadratic term would estimate the curvature of the return to education (Wooldridge 1997). These conditions are more restrictive than those used to estimate the baseline model; therefore, the estimates in this section should be taken with caution.

Table 3 reports the estimation results. When a quadratic term is included, the LS estimate of the return to education increases by about a percentage point to 9.1 percent (in comparison to table 2) in the March CPS. Note that educational attainment is first demeaned and then squared. The IV estimate of the return to education at the mean is 4.5 percent, which is similar to the benchmark estimate.

The coefficients of the quadratic term in columns 1 and 3 are significantly positive, indicating a convex relationship between earnings and schooling. Approximately 3 years of education above the mean is associated with 1 percentage point of additional return in the March CPS sample. The IV estimates of the quadratic term, in contrast, are not significantly different from zero in all cases. This indicates the absence of strong scale effects in the return to education.

The LS estimate of the curvature parameter is identified *across* workers given their optimal education choices and does not necessarily indicate an increasing return to scale for individual workers. The IV estimate of the curvature instead is identified by the movements in the average educational attainment, corresponding to the curvature of the average

TABLE 3
Return to Education: Quadratic Specification

Explanatory Variable	March CPS LS (1)	March CPS IV (2)	Decennial Census LS (3)	Decennial Census IV (4)	Decennial Census LS (5)	Decennial Census IV (6)
100 × educ	9.14	4.47	8.98	2.41	8.32	4.85
	(.15)	(1.43)	(.17)	(1.04)	(.15)	(.66)
100 × educsq	.36	−0.47	.44	.03	.47	.06
	(.02)	(.53)	(.02)	(.29)	(.02)	(.09)
Instruments	Year of birth		Year of birth		Year of birth × state of birth	
First-step Wald test:						
educ	46.07		326.80		72.26	
educsq	6.15		16.92		57.42	
Wu-Hausman test	10.78*		41.00*		29.15*	
Observations	943,760		5,227,984		5,227,984	

Note.—All specifications control for a quartic age trend, a quadratic trend in year of birth, race, and survey year. Columns 5 and 6 allow the cohort trend to vary by state and further control for indicators for state of birth and state of residence. The instruments are indicators for year of birth (interacted with state of birth in col. 6). The variable educsq is demeaned and squared years of education. All standard errors are clustered by cohort and are reported in parentheses. Data come from the March CPS (1964–2003) and the DCS (1960–2000). The sample includes men of ages 25–60.
* Significant at the 1 percent level.

production function (or of the average worker if the production function permits aggregation).

C. *Cohort Effects in Wages*

In general, one might be concerned about potential cohort effects in wages that would violate the exclusion restriction. The IV estimator in cohort-based studies may be misleading when the treatment group is a single cohort of workers and if a strong cohort effect is present on the treatment group relative to the comparison cohorts (see Card 2001). Such cohort effects are not of concern here. Since the estimation method uses several cohorts, random cohort effects that are outside the model would be close to zero on average unless they are systematically related to average educational attainment.[21] The other cohort effects mentioned in the literature can be grouped into two main categories: those related to a cohort's size and those related to the economic conditions around the time a cohort enters the market. I address these in the following subsections.

1. Cohort Size and Congestion Effects

The benchmark specification implicitly assumes that workers in different age and education groups are perfect substitutes for each other. If work-

[21] These effects can be serious when the instruments are weak (e.g., Bound, Jaeger, and Baker 1995).

ers of different cohorts are imperfect substitutes, then a change in the labor supply of an age or education group has an equilibrium effect on the distribution of wages. For instance, Card and Lemieux (2001) drew attention to the imperfect substitution of workers of different age groups to interpret the changing skill premia, whereas Heckman, Lochner, and Taber (1998) emphasized the supply effects across different skill levels. Welch (1979) argues that members of the baby boom generation had to accept lower wages at the onset of their career since the market was overwhelmed with younger workers. Omission of changes in the labor supply from the wage regression would bias our estimates only if they were related to educational attainment in a systematic way.

The literature on cohort effects in earnings points out two plausible scenarios in which a cohort's educational attainment can be correlated with its size. Following Welch (1979), large cohorts may prefer to stay longer at school to avoid potential congestion in the entry-level labor market (Falaris and Peters 1992). This would lead to a negative correlation between cohort intercepts and educational attainment and, hence, to a downward bias in the estimate of the return to education. However, educational attainment of larger cohorts may be adversely affected by the limited availability of public funds (Bound and Turner 2007) or the reduced quality of education due to smaller class size. Similarly, Stapleton and Young (1988) pointed out that if substitutability between young and old workers diminishes with education, workers with higher educational attainment in a large cohort will experience suppressed wages for a longer period, which in turn will reduce the incentive to acquire education. All these arguments would cause the estimates in the previous section to overstate the return to education.

To account for the effects of cohort size, the control set is broadened to include measures of labor supply by age and education. Let N_{cst} denote the total labor supply of all workers in cohort c and education group s in year t.[22] A worker's earnings are the amount of efficiency units the worker governs, $h_i = \exp(\alpha_i + \beta_i s_i)$, priced at the wage per efficiency unit, $\omega_{cst} = \partial Q_t / \partial N_{cst}$, where Q_t is the production function. Equation (1) becomes

$$\ln w_i = \alpha_{ct} + \ln \omega_{cst} + \beta_{ct} s_{ict} + \gamma_{ct} a_{ict} + u_{ict},$$

where changes in labor supply will be reflected in ω_{cst}. To capture movements in ω_{cst}, a second-order trans-log functional form in cohort and skill supplies by year is included in the regression. This form is preferred for its flexibility to approximate different production functions:

$$\ln \omega_{cst} = \psi_0 \ln N_{ct} + \psi_1 \ln N_{st} + \psi_2 \ln N_{ct} \times \ln N_{st} + \psi_3 \ln^2 N_{ct} + \psi_4 \ln^2 N_{st}.$$

Table 4 displays the estimation results. Each year of age and completed

[22] Since age, time, and birth year are linearly dependent, one can equivalently use age instead of cohort in the notation: $N_{cst} = N_{ast}$.

TABLE 4
RETURN TO EDUCATION AND COHORT SIZE

	March CPS		Decennial Census			
Explanatory Variable	LS (1)	IV (2)	LS (3)	IV (4)	LS (5)	IV (6)
100 × educ	8.29	6.38	7.86	4.88	7.50	4.85
	(.19)	(2.50)	(.21)	(1.72)	(.21)	(.43)
Instruments	Year of birth		Year of birth		Year of birth × state of birth	
First-step Wald test	58.24		171.9		72.26	
Wu-Hausman test	.59		3.05***		4.55*	
Observations	943,760		5,227,984		5,227,984	

Note.—All specifications control for a quartic age trend, a quadratic trend in year of birth, race, survey year, and number of workers in each cohort by education and survey year. Columns 5 and 6 allow the cohort trend to vary by state and further control for indicators for state of birth and state of residence. The instruments are indicators for year of birth (interacted with state of birth in col. 6). All standard errors are clustered by cohort and are reported in parentheses. Data come from the March CPS (1964–2003) and the DCS (1960–2000). The sample includes men of ages 25–60.
* Significant at the 1 percent level.
*** Significant at the 10 percent level.

schooling is considered a separate category in the estimations. In all specifications, the IV estimate of the return to education increases in comparison to the benchmark estimation, getting closer to (but still less than) the LS estimate. The return to education is 4.88 percent in the DCS sample and 6.38 percent in the CPS. The data support a positive covariance between education and cohort size and a negative covariance between wages and cohort size, supporting the entry deferral argument suggested by Falaris and Peters (1992). The change in the IV estimate is particularly due to the omission of cohort size effects in the benchmark model because the identification of the return to education essentially depends on cohort-based comparisons of education and earnings. The LS estimate only slightly decreases when accounting for the supply effects.

The quality of education might also decrease if a cohort of average size stays longer at school, hence crowding the classes at higher levels of education. This would decrease the effective quality of college education independent of a cohort's size relative to other cohorts. Although this channel is not directly captured by the cohort size controls, given the change in the IV estimate with and without the cohort size variables, one could infer a bound on the magnitude of this effect. Since controlling for cohort size raises our estimates by less than 2 percentage points and the average education is 12.7 years, the marginal effect of an additional year of education on the quality of education should be worth less than 0.16 percentage points. Therefore, the change in the quality of education due to congested classes is unlikely to be an important factor in explaining the difference between the IV estimate and the LS estimate.

Column 6 in table 4 reports the estimation results using interactions of birth state with birth year as instruments for education. Here, the IV estimate of the return to education is 4.85 percent, which is similar to the benchmark specification. This is not surprising since the cohort size effects on wages are partly equilibrium effects and operate at a national level because of worker mobility across states. Because the identification with state cohorts comes from relative changes in educational attainment across cohorts at the state level, they are less prone to these effects. Moreover, the estimate here has a standard error that is one-fourth of the specification in column 4. This enables us to provide improved confidence for the Wu-Hausman test statistic, which rejects the similarity to the LS estimate at the 1 percent significance level.

The point estimates in table 4 indicate an ability bias of 2–3 percentage points, depending on the data set and specification. Since the LS return is still around 8.0 percent, the increase in actual marginal product of a worker induced by a year of education constitutes two-thirds of the observed wage gap.

2. Labor Market Conditions at Entry

Additional cohort effects in wages could exist if the fluctuations in the labor market upon entry have a lasting effect on workers' careers. Freeman (1981) emphasizes that workers who enter the labor market during a downturn may experience not only low entry-level wages but also lower wage growth. This is particularly important if educational choices respond to the market conditions. If the cohort effects are important, then workers might be willing to delay their market entry and wait for the economy to recover. This predicts a countercyclical educational attainment, which would bias the benchmark estimate downward. However, economic downturns might have an income effect if the workers are subject to liquidity constraints and cause a reduction in the average education. This would create an upward bias in the benchmark estimate.

The literature on the cyclicality of educational attainment in general finds a countercyclical pattern in enrollment decisions (Gustman and Steinmeier 1981; Sakellaris and Spilimbergo 2000; Dellas and Sakellaris 2003), with the exception of Corman (1983) and Card and Lemieux (2000), who do not find a significant effect. The quantitative effects estimated in these studies are usually small. Enrollment rates increase by 0–5 percent for a percentage point increase in the unemployment rate. In their study of managerial wages in a firm, Baker, Gibbs, and Holmstrom (1994) also dismissed any correlation between educational attainment and cohort effects in wages over the business cycle.

One way to capture the cohort effects that might potentially arise from persistent effects of market conditions is to explicitly control for the unemployment rate around the time of graduation. In particular, for each worker, we calculate the estimated graduation date (year of

TABLE 5
RETURN TO EDUCATION AND LABOR MARKET ENTRY CONDITIONS

	March CPS				Decennial Census			
	LS (1)	IV (2)	LS (3)	IV (4)	LS (5)	IV (6)	LS (7)	IV (8)
100 × educ	8.22	5.11	8.32	5.49	8.08	4.97	8.76	6.44
	(.16)	(1.11)	(.19)	(1.78)	(.15)	(.84)	(.20)	(.91)
Wu-Hausman test	8.02*		2.55		14.23*		6.83*	
Cohort size controls	. . .	. . .	Yes	Yes	. . .	. . .	Yes	Yes
Unemployment rate	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note.—All specifications control for a quartic age trend, a quadratic trend in year of birth, race, survey year, and the aggregate rate of unemployment at the time of graduation, a year before, and a year after graduation. Columns 3 and 4 and 7 and 8 also control for the number of workers in each cohort by education and survey year. The instruments are indicators for year of birth. All standard errors are clustered by cohort and are reported in parentheses. Data come from the March CPS (1964–2003) and the DCS (1960–2000). The sample includes men of ages 25–60.
* Significant at the 1 percent level.

birth + 6 + years of schooling of the worker) and include the unemployment rate in this year, the year before, and the year after.[23] In addition, the current unemployment rate is added to the regression as a control variable.

Table 5 displays the results. The estimated return to education is virtually the same in the LS model for both data sets. For the IV estimation, the return to education is 5.11 percent for the linear specification in the CPS data. This is half a percentage point higher than the benchmark estimation. Given that the standard error is around 1 percent, the two estimates are statistically similar. If we also include the cohort size variables discussed in the previous section to control for equilibrium effects, the estimated return to education is 5.49 percent. Therefore, we conclude that controlling for labor market entry conditions has a negligible effect on our estimates. The coefficients of the lagged unemployment variables (not reported) are not statistically different from zero in all specifications. The results for the DCS data are similar. The rate of return to education in the linear specification is 4.97 percent, which is slightly higher than the benchmark estimate.[24] Once the cohort size variables are included, the rate of return to education becomes 6.44 percent, which is very similar to that obtained in the CPS data. The small increase in the estimated rates suggests a countercyclical pattern in educational attainment, consistent with the studies mentioned above.

[23] Adding more lags and leads does not change the estimates but restricts the available variation in education by cohorts; therefore, they are excluded.

[24] If the unemployment rate is persistent, then these estimates might also be capturing some cohort size effects, which would explain the similarity of the estimates obtained here and in the previous section.

IV. The Rising Skill Premium

The widening wage structure since the 1980s is a well-documented fact of the U.S. labor market (Katz and Murphy 1992; Katz and Autor 1999; Autor, Katz, and Kearney 2008). The relative wages of college graduates have risen substantially, suggesting tighter markets for the educated, though this could also be the observed outcome of a rise in the return to unobserved ability. One of the advantages of using the year of birth as an instrument is that the estimation can be carried out for different periods or across countries, unlike some instruments based on one-time, institutional changes in the economy. In this section, I estimate the change in the return to education and ability bias over time. A relatively more straightforward way to capture this is to estimate an interaction of years of education with a time trend. I follow the procedure described in Section II to estimate the interaction of educational attainment with a quartic time trend. The control variables are indicators for race, survey year, a quadratic trend in year of birth, and variables for cohort size and skill supplies to capture the potential cohort effects, as described in the previous section. The observations are weighted, as before, by their cross-sectional weights provided by the U.S. Census.

Figure 8 plots the estimated trend for the return to education. The first specification (LS1 and IV1) controls only for the cohort size, by education and year; the second specification (LS2 and IV2) also controls for labor market entry conditions. The trends from the two specifications, estimated by LS, are virtually identical. The LS estimate of the

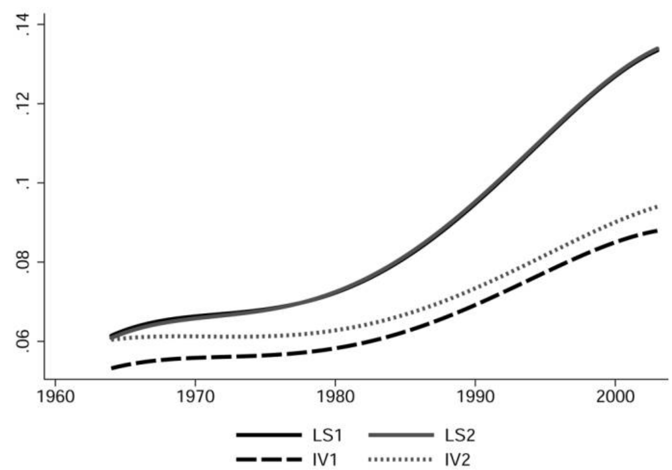

Figure 8.—Rising return to education, 1964–2003. The figure shows the estimated quartic trend for the return to education. All specifications control for a quartic function in age, a quadratic trend in year of birth, race, survey year, and cohort size by education and year. Specifications LS2 and IV2 also control for the unemployment rate at graduation and a year before and after graduation. Data come from the March CPS.

TABLE 6
RISING RETURN TO EDUCATION: 1964–2003

Years	LS (1)	IV (2)	LS (3)	IV (4)	LS (5)	IV (6)
1964–73	6.56	6.34	6.78	5.31	6.59	4.75
	(.05)	(.54)	(.05)	(.60)	(.08)	(.33)
1974–83	6.91	6.21	7.33	5.94	7.36	5.41
	(.04)	(.44)	(.05)	(.45)	(.06)	(.26)
1984–93	9.28	7.32	9.65	7.00	10.04	6.99
	(.05)	(.46)	(.06)	(.45)	(.07)	(.26)
1994–2003	12.03	8.75	12.29	7.90	13.09	8.39
	(.07)	(.54)	(.07)	(.55)	(.11)	(.32)
Observations	943,760	943,760	759,315	759,315	759,315	759,315

Note.—All specifications control for a quartic age trend, a quadratic trend in year of birth, race, survey year, cohort size by education and survey year, and unemployment rates at graduation, a year before, and a year after graduation. Instruments are indicators for the year of birth. Specifications 1 and 2 use sampling weights; 3 and 4 use only a subset of cohorts and new weights as depicted in fig. A1 and described in App. A to yield unweighted averages of returns within decades. Columns 5 and 6 add controls for a quartic cohort trend in the return to education. Reported figures in these columns reflect the average return for all cohorts. Standard errors are clustered by cohort and are reported in parentheses. Data are taken from the March CPS, 1964–2003.

return to education increases from 6 percent in 1964 to around 13 percent. The increase in the return to education accelerates after 1980, which is consistent with findings in the literature.

The dashed curves show the IV estimates of the trend for the return to education. The IV estimate in the first specification starts at about a percentage point below the LS estimate in 1964, rising only mildly to 6 percent until 1980, then picking up to 8 percent. The result from the second specification is consistently higher by a percentage point but displays a similar pattern over time. This implies that the rise in the wage gap across education levels must, at least in part, be due to the rise in the return to formal education. At the same time, the gap between the LS estimate and the IV estimate widens, especially after 1980. This suggests that the price of unobserved skill has also gone up or, alternatively, that the ability levels have become more dispersed relative to educational choices.

These findings are at odds with those of both Deschenes (2006), who finds nearly constant returns to unobserved ability, and Taber (2001), who does not find any evidence for a rise in the return to education. It is somewhat harder to compare the findings in Taber's study with ours because his focus is on college education per se, and the reported average and marginal treatment effects are not directly comparable to the IV estimates here.

In table 6, I report the estimates of the return to education in 10-year intervals. Columns 1 and 2 report the estimates for the specification in figure 8 with controls for both cohort size by education and labor market entry conditions. The standard errors for the IV estimates are around a half percentage point, much smaller than the errors in the specification in which the return to education is assumed constant (table 5). Columns 3 and 4 report the estimates using, for each decade, only those cohorts that appear in the sample in every year of the decade.[25] The remaining observations are reweighted to estimate unweighted averages of return to education during each decade (see Sec. II and App. A). The results show a similar pattern but are slightly lower than in columns 1 and 2. The return to education rises from 5.3 percent in the 1964–73 period to 7.9 percent during 1994–2003. The hypothesis that the return to education is constant over time is strongly rejected. The point estimate displays a rise of 2.6 percentage points, which is about a half of the observed rise in the LS estimate.

The Wu-Hausman test statistics for the difference between the LS and the IV estimates are 6.09, 9.75, 34.9, and 63.9 in sequence. They are all significant at a confidence level of 1 percent, suggesting that the ability bias is significantly positive and that it has risen significantly in the last four decades. The point estimate of the ability bias is 1.5 percentage points (22 percent of the LS estimate) in 1964–73, and it rises to 4.4 percentage points (36 percent of the LS estimate).

When the two specifications in columns 2 and 4 of table 6 are compared, the reweighting of the observations changes the estimates by less than 1 percentage point. This indicates that the correlation between the relative weights of cohorts and educational attainment is not strong. Therefore, the benchmark estimates reported in the previous section are not too far from the average return to education, though it masks the underlying trend.

When the controls for number of workers in cohort and education groups by year are excluded, the IV estimates of the return to education decline by 2.9 points from 8.8 percent in the 1964–73 period to 5.9 percent in the late 1970s and early 1980s. This is followed by 7.2 percent and 10 percent in the following two decades.[26] This highlights the role of the rise in the supply of educated workers in the 1970s in explaining the drop in the return to education in this period. If one ignores the supply-side considerations and restricts attention to the period since the 1980s, the rise in the education premium appears more sizable relative to that implied by the changes in demand for skill only.

Steady demand versus acceleration.—In the literature, a debate exists about whether the demand for skill increased steadily or at a faster pace after 1980 as a result of developments in electronics and information technologies (Katz and Murphy 1992; Autor, Krueger, and Katz 1998; Card and Lemieux 2001; Acemoglu 2002). The LS specifications in figure 8 display an apparent acceleration after 1980, and the specifi-

[25] Specifically, cohorts born in 1914–39 are used to identify the return to education during 1964–73, cohorts born in 1923–49 are used for 1974–83, cohorts born in 1933–59 are used for 1984–93, and cohorts born in 1943–68 are used for 1994–2003.

[26] These are comparable to col. 4 in table 6.

cations have been already conditioned on changes in the supplies of workers by education, cohort, and year. Therefore, the acceleration is not explained by supply-side considerations. This section attempts to determine whether this acceleration is led by the return to formal education, by unobserved ability, or by both.

Both specifications in figure 8 indicate that most of the rise in the return to formal education occurred in the last two decades. The return to ability has been increasing since the 1960s, but at a faster pace after 1980. While these observations suggest that acceleration occurred in demand for both observed and unobserved skills, this result should be taken with caution since the estimates ignore possible cohort trends in the return to education.

Advancements in the quality of education over extended periods could lead to an increasing return to education, conditional on a schooling level. These improvements would be slowly reflected in the cross-sectional measures of returns to education as better-educated, newer cohorts slowly diffuse into the existing workforce. Although this does not change the fact that there was a rise in the return to education, it matters for understanding *when* it happened.

Alternatively, the average ability, conditional on a level of schooling, may decline by cohort as newer cohorts attain higher and higher levels of education. Similarly, if the return to education varies with age, say, because of differential rates of human capital accumulation across cohorts, including cohort trends would detach these effects from the time-specific estimates of the return to education. Once these effects are controlled for, the change in the return to education over time is identified by relatively sudden, year-to-year movements that are more likely due to marketwide changes in the labor market and that apply to all cohorts at any given year.

Figure 9 shows the effect of controlling for cohort trends in the return to education. Specifications LS2 and IV2 are the same as before (fig. 8). Specifications LS3 and IV3 control for a quartic trend in year of birth in the return to education. Three changes are seen in the time patterns of the estimates. First, the time profile of the LS return slightly rotates and becomes steeper. This implies that the cohort trend has a negative slope. Although a continuously deteriorating schooling system may be hard to believe, this result can be reconciled by a decreasing student quality across cohorts, conditional on education. This is consistent with older workers having higher returns to education.

Second, the average estimate of the return to formal education is now lower. This signals that the cohort trend in return to education is positively correlated with average educational attainment of cohorts. This is akin to the nonclassical ability bias analyzed in Card (2001) and Heckman et al. (2008) at the individual level.

Third, controlling for cohort trends alters the time profile of the return to formal education in an important way. Inspection of trends

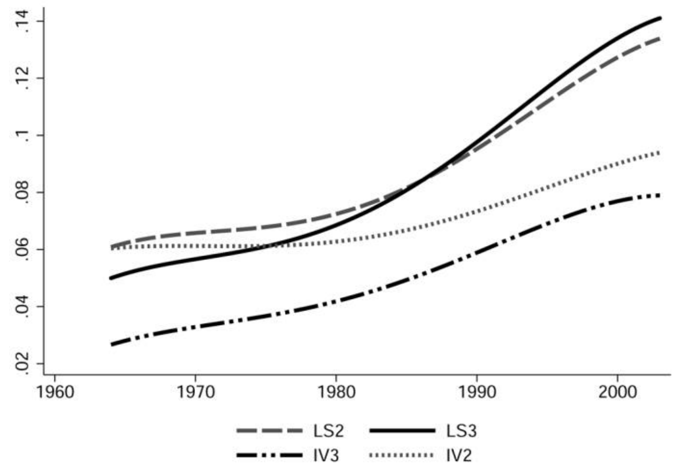

Figure 9.—Rising return to education, 1964–2003. The figure shows the estimated quartic time trend for the return to education. All specifications control for a quartic function in age, a quadratic trend in year of birth, race, survey year, cohort size by education and year, and the unemployment rate at graduation, a year before, and a year after graduation. Specifications LS3 and IV3 also control for a quartic cohort trend in return to education. Data come from the March CPS.

LS3 and IV3 reveals that the rise in the return to education due to marketwide factors has been approximately at the same rate since 1963, with only a mild acceleration around the 1980s. However, the ability bias, the difference between IV3 and LS3, was more or less constant until the 1980s and increased substantially afterward. This finding is consistent with the conclusions of Juhn, Murphy, and Pierce (1993), who relied on increasing residual wage inequality to conclude that the price of unobserved ability is the main reason for the acceleration of the rise in wage inequality.

Columns 5 and 6 in table 6 display the estimates for the last four decades. The estimate of the return to education for the 1964–73 period is 4.76 percent, rising at a rate of about 1.6 percent per decade to 8.40 percent in 1993–2003. At the same time, the point estimates of ability bias are 1.84 percent, 1.95 percent, 3.05 percent, and 4.70 percent for the last four decades, indicating that most of the rise in ability bias occurred in the last two decades.

V. Discussion

The second half of the twentieth century witnessed a substantial rise in the education premium. It has been argued that the earnings gap between workers of different educational backgrounds overstates the causal effect of education. Therefore, it is uncertain whether the higher education premium reflects a higher return to education or a higher market value of unobserved skills. The findings here indicate that both the return to education and the ability bias have increased since the 1960s. The return to formal education is responsible for approximately half of the observed rise in the education premium. Moreover, the acceleration of the rise in the education premium observed after 1980 is almost entirely explained by growing ability bias.

Earlier studies have emphasized the role of changing supply of educated workers in understanding changes in the skill premium (Katz and Murphy 1992; Autor et al. 2008; Goldin and Katz 2008).[27] The estimates here are generally robust to changes in the supply of skills; therefore, the acceleration in the rate of increase after 1980 is not explained by a slowdown in supply factors. Nevertheless, supply-side considerations help to explain the drop in the return to formal education in the late 1970s.

The results are consistent with a steady increase in the demand for skills acquired at school since the 1960s. In contrast, we observe a rise in the ability bias, especially after 1980. A number of studies emphasized a change in the nature or the pace of technological change in the United States since the late 1970s (Greenwood and Yorukoglu 1997; Acemoglu 1998, 2003; Galor and Moav 2000; Krussel et al. 2000; Aghion, Howitt, and Violante 2002). While our results are compatible with the results of these studies, they further underline the particular role of unobserved ability characteristics relative to formal education in explaining the rise in wage inequality. The developments in electronics and information technologies since the 1970s could have generated a higher demand for certain skills that were not captured by formal education at the time. It is conceivable that these skills were better acquired on the job rather than at school because, for instance, they were too specific to an industry or full-time skill accumulation was too costly, at least for older workers. In either case, workers would have preferred to acquire these skills on the job rather than at school. If workers also differed in their ability to learn and adapt to new technologies, the new technology would lead to an increase in the ability bias. This particular story of technology diffusion is also emphasized by Heckman et al. (1998), Ehrlich and Kim (2007), and Guvenen and Kuruscu (2010), who analyze the effects of skill-biased technical change on wage distribution in an overlapping generations framework with endogenous human capital accumulation and heterogeneous agents.

The rise in the ability bias could be due to a change in the composition of abilities required in production. For example, a shift in production from manufacturing to services may raise the relative demand for cognitive ability or communication skills. If the distribution of these skills has a larger variance than the skills required in the manufacturing sector or if they are more strongly correlated with educational attainment,

[27] See Guryan (2008) for a review of Goldin and Katz's book.

then this shift would manifest itself through a larger ability bias due to increased sorting. Bartel and Sicherman (1999) hint for such an effect of technological change on the sorting of workers to industries caused by a rise in the demand for unobserved characteristics and innate ability.

Alternative explanations of rising ability bias differ in their predictions for the persistence of the elevated role of ability. If the skills required by a new technology can eventually be made widely accessible through formal education, then we would expect the increased role of ability to be only temporary.[28] However, if the new technology demands new skills that are specific to industries or firms or that have to be learned by experience, the schools may never provide these skills. The results in figures 8 and 9 do not indicate any decline in the ability bias so far.[29] The larger ability bias, therefore, could be an indication of a permanent change in the way workers acquire human capital, in particular, of a heightened role of cognitive skills in acquiring new skills and thereby maintaining the value of one's time at work. Insofar as such is the case, the education policy should emphasize the development of cognitive skills, which are mostly acquired before formal schooling begins. Future work may help to further identify the specific causes underlying the trends estimated here.

This paper also finds a relatively low return to education compared to some other studies in the literature. The benchmark estimate of the return to a year of education is around 5–6 percent, which is significantly lower than the observed wage difference of about 8–9 percent. The literature has adhered to diminishing returns to education or credit constraints to explain the high IV estimates. The findings presented here fail to support significantly diminishing returns to scale and indicate that the marginal worker has a lower return relative to the average worker.[30]

Carneiro and Heckman (2002) and Carneiro, Heckman, and Vytlacil (forthcoming) argue that the ability bias can, in fact, be negative when ability is multidimensional, as in a Roy model of selection. Willis and Rosen (1979) point out that when workers select into different schooling choices associated with job prospects that reward different types of ability (teachers with a college degree vs. plumbers with a high school di-

[28] Ehrlich and Kim (2007) present an exception to this claim in which the interaction of the heterogeneity in human capital accumulation with endogenous fertility decisions may lead to more persistent effects.

[29] This is in contrast with studies that rely on residual inequality to infer the role of unobserved ability in wage inequality. In contrast to weekly earnings, the increase in the residual inequality based on hourly wages was limited if not stable since the 1990s (Lemieux 2006; Autor et al. 2008).

[30] In a related study, Carneiro and Heckman (2002) question the validity of the common instruments used in the literature. Drawing data from test scores from the National Longitudinal Survey of Youth, the authors show that most of the instruments used in the literature are correlated with ability in ways that lead to higher IV estimates. This could also help reconcile the discrepancy of the estimates.

ploma), the LS estimate of the return could be biased downward if these different types of ability are negatively correlated (i.e., if an average plumber earns more than an average teacher would make if the teacher were to become a plumber). These findings suggest that the different ability types are more likely to be positively correlated.

An advantage of using variation in educational attainment across birth cohorts is that it can be applied to different periods. This feature is used here to understand the changes in the returns to education and unobserved ability. Similarly, future work could apply this estimation strategy to other countries using micro-level data. Although studies have documented the world distribution of Mincer returns (Psacharopoulos and Patrinos [2004], among others), a uniform estimation of ability bias across countries is not readily available. An extension of the current approach to other nations would improve our understanding of the relationship between schooling and cross-country variation in income and growth.

Appendix A

A Cohort-Based Weighting Procedure

This technical appendix demonstrates that when the return to education varies over time, a reweighting procedure can be used to estimate the unweighted average of the return to education with pooled cross-sectional data. Let the marginal productivity of worker i of cohort $j \in 1, 2, \ldots, J$ at time $t \in 1, 2, \ldots, T$ be given by

$$y_{ijt} = \gamma_t \alpha_i + \beta_t s_{ijt} + \nu_{ijt}. \tag{A1}$$

This specification explicitly allows for a changing return to education and ability over time.

Let $\pi = (\pi_{ijt})$ be a vector of weights. Define $n_{jt}(\pi) = \sum_i \pi_{ijt}$, $n_j(\pi) = \sum_t \sum_i \pi_{ijt}$, and $n(\pi) = \sum_j \sum_t \sum_i \pi_{ijt}$. Define the weighted average of a variable x with these weights as

$$\bar{x}(\pi) = \frac{1}{n(\pi)} \sum_j \sum_t \sum_i \pi_{ijt} x_{ijt}$$

and the cohort average as

$$\bar{x}_j(\pi) = \frac{1}{n_j(\pi)} \sum_t \sum_i \pi_{ijt} x_{ijt}.$$

Definition 1. For each j, a weighted Wald estimator based on cohort j is given by

$$\beta_j^w(\pi) = \frac{\bar{y}_j(\pi) - \bar{y}(\pi)}{\bar{s}_j(\pi) - \bar{s}(\pi)}. \tag{A2}$$

Theorem 1 (Angrist 1988). The instrumental variables estimator of β using vectors of **Z** as instruments for s is equivalent to

$$\beta^{\mathrm{IV}}(\pi) = \sum_j \omega_j \beta_j^w(\pi),$$

where

$$\omega_j = \frac{[\bar{s}_j(\pi) - \bar{s}(\pi)]^2}{\sum_j [\bar{s}_j(\pi) - \bar{s}(\pi)]^2}.$$

This theorem essentially states that the IV estimator is a variance-weighted average of cohort-specific Wald estimators. In a more general setting, Angrist shows that the IV estimator is an efficient combination of the Wald estimators.

Let π^o be sampling weights.

Assumption 1. π^o is such that, for each j and t, $n_{jt}(\pi^o) > 0$.

This assumption guarantees that at any time t in the sample, each cohort appears with a positive measure in the data.

Proposition 1. Define $\tilde{\pi}_{ijt} = \pi^o_{ijt} n_j(\pi^o)/n_{jt}(\pi^o)$. Then the IV estimator using vectors of **Z** as instruments for s and $\tilde{\pi}$ as weights consistently estimates $\sum_t \beta_t / T$.

Proof.

Step 1: For any j, $\beta_j^w(\tilde{\pi})$ is a consistent estimator of $\sum_t \beta_t / T$.

With equation (A1), the average marginal productivity for a cohort can be calculated as

$$\begin{aligned}
\bar{y}_j(\tilde{\pi}) &= \frac{1}{n_j(\tilde{\pi})} \sum_t \sum_i \tilde{\pi}_{ijt} y_{ijt} \\
&= \frac{1}{T n_j(\pi^o)} \sum_t \frac{n_j(\pi^o)}{n_{jt}(\pi^o)} \Big(\beta_t \sum_i \pi^o_{ijt} s_{ijt} + \gamma_t \sum_i \pi^o_{ijt} \alpha_{ijt} + \sum_i \pi^o_{ijt} \nu_{ijt} \Big) \\
&= \frac{1}{T} \Big[\sum_t \beta_t \bar{s}_{jt}(\pi^o) + \sum_t \gamma_t \bar{\alpha}_{jt}(\pi^o) + \bar{\nu}_{jt}(\pi^o) \Big].
\end{aligned}$$

Given that $p\lim \bar{s}_{jt}(\pi^o) = E[s|j]$, $p\lim \bar{\alpha}_{jt}(\pi^o) = E[\alpha|j] = E[a]$, and $p\lim \bar{\nu}_{jt}(\pi^o) = 0$, average cohort productivity converges to

$$p\lim \bar{y}_j(\tilde{\pi}) = E[s|j] \frac{1}{T} \sum_t \beta_t + E[\alpha] \frac{1}{T} \sum_t \gamma_t. \tag{A3}$$

Similarly, overall average productivity is

$$\begin{aligned}
p\lim \bar{y}(\tilde{\pi}) &= p\lim \frac{1}{n(\tilde{\pi})} \sum_j \sum_t \sum_i \tilde{\pi}_{ijt} y_{ijt} \\
&= p\lim \frac{1}{T n(\pi^o)} \sum_j n_j(\pi^o) \sum_t \bar{y}_{jt} \\
&= p\lim \frac{1}{T n(\pi^o)} \sum_j n_j(\pi^o) \Big[\sum_t \beta_t \bar{s}_{jt}(\pi^o) + \sum_t \gamma_t \bar{\alpha}_{jt}(\pi^o) \Big] \\
&= E[s] \frac{1}{T} \sum_t \beta_t + E[\alpha] \frac{1}{T} \sum_t \gamma_t.
\end{aligned} \tag{A4}$$

Note that with the new weights,

$$
\begin{aligned}
p\lim \bar{s}_j(\tilde{\pi}) &= p\lim \frac{1}{n_j(\tilde{\pi})} \sum_t \sum_i \frac{n_j(\pi^o)}{n_{jt}(\pi^o)} \pi^o_{ijt} s_{ijt} \\
&= \frac{1}{Tn_j(\pi^o)} \sum_t n_j(\pi) p\lim \bar{s}_{jt}(\pi^o) \\
&= E[s|j] \qquad \text{(A5)}
\end{aligned}
$$

and

$$
p\lim \bar{s}(\tilde{\pi}) = \frac{1}{Tn(\pi^o)} \sum_j n_j(\pi^o) \sum_t p\lim \bar{s}_{jt}(\pi^o) = E[s]. \qquad \text{(A6)}
$$

From equations (A3), (A4), (A5), and (A6), the Wald estimator based on a cohort j converges to

$$
\begin{aligned}
p\lim \beta^w_j(\tilde{\pi}) &= \frac{E[s|j](1/T)\sum_t \beta_t + E[\alpha](1/T)\sum_t \gamma_t - E[s](1/T)\sum_t \beta_t - E[\alpha](1/T)\sum_t \gamma_t}{E[s|j] - E[s]} \\
&= \frac{1}{T}\sum_t \beta_t.
\end{aligned}
$$

Step 2: Since the IV estimator is a weighted average of consistent estimators, it is itself consistent. QED

When the purpose is to estimate the average return for several subperiods, T_1, T_2, and so forth, the estimation can be carried out separately for each subperiod with the weighting procedure described above. Nevertheless, this may be impractical or undesirable when additional control variables with time-consistent coefficients are present in the regression. For instance, estimation of a common, long-term trend in year of birth requires that the return to education

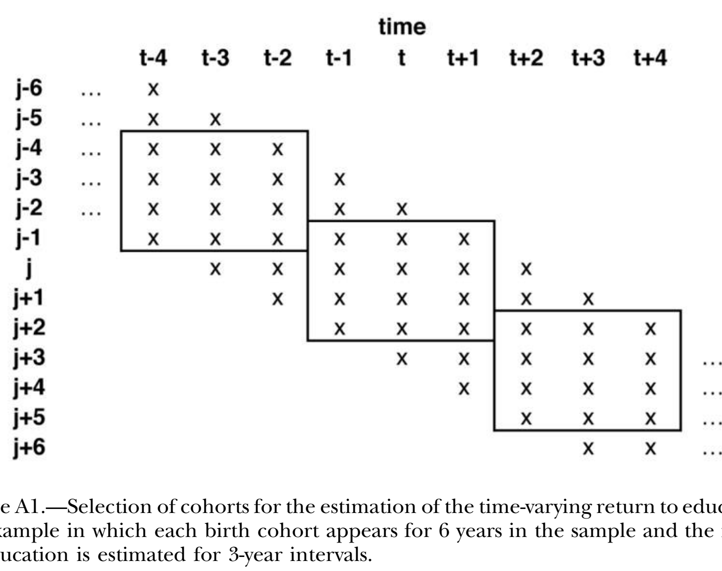

Figure A1.—Selection of cohorts for the estimation of the time-varying return to education: an example in which each birth cohort appears for 6 years in the sample and the return to education is estimated for 3-year intervals.

for all subperiods are estimated at once using all cohorts. In this case, the weighting procedure can be extended by interacting indicator variables for subperiods with indicator variables for those cohorts that appear in each and every year of each subperiod. Then the observations can be reweighted, as described above, within each subperiod. As an example, figure A1 illustrates the selection of cohorts for the estimation of return to education during 3-year intervals when each cohort appears in the sample for 6 years. After the selected cohorts are reweighted, horizontal averages of γ_t and β_t will be the same for each 3-year period. The specification corresponding to the example in the figure would include three dummy variables, one for each block illustrated by a rectangle.

Appendix B

Year of Birth versus Quarter of Birth

This appendix compares the estimates from the benchmark model with estimates that use variation in earnings and education by quarter of birth, as in Angrist and Krueger (1991), for robustness purposes. This allows us to see if the results are simply due to differences in the sample or to the estimation strategy. Table B1 reports the estimates using the quarter of birth as an instrument for educational attainment. Angrist and Krueger argue that the students born later in the year are likely to attend school for an extra year because of a minimum age requirement for leaving school, whereas their wages are similar to those of others, conditional on education. Columns 1–6 replicate the specifications in Angrist and Krueger's paper. These are comparable to columns 5 and 6 in their tables 4, 5, and 6. The instruments are quarter of birth indicators interacted with year of birth indicators. The control variables are age, age squared, and indicators for the year of birth and race. These results are very close to theirs, indicating no major differences in the sample.[31]

Since the estimation strategy here relies on several cross sections and cohorts, I also estimate their model pooling three census years for which the data on the quarter of birth are available. I restrict the sample to cohorts that appear in all three surveys. The results are reported in columns 7 and 8. The point estimate of the return to education is 5.2 percent, relative to 6 percent obtained by LS, with a standard error of 1.21. When birth year cohort dummies are used, as in our benchmark estimates, the estimate of the return to education is 4.9 percent with a standard error of 0.82 in column 10. While the latter estimate is slightly lower, the major difference between the two designs appears in the first step of the estimation, where the F-statistic for the quarter of birth instruments is 2.47 compared to 19.69 for the birth year instruments. The presence of weak instruments could explain why the former estimate is closer to the LS estimate. Staiger and Stock (1997) report corrected IV estimates using quarter of birth that are robust to weak instruments, and these are around 8–10 percent.

Appendix C

Data

Both the decennial census surveys, 1960–2000, and the annual March supplements to the CPS for the years 1964–2003 are extracted from the IPUMS-USA database (Ruggles et al. 2004).

[31] The estimates are not exactly the same as theirs, possibly because of (i) differences in the sample and (ii) absence of controls for the standard metropolitan statistical area.

TABLE B1

Return to Education: Quarter of Birth and Year of Birth as Instruments

	LS (1)	IV (2)	LS (3)	IV (4)	LS (5)	IV (6)	LS (7)	IV (8)	LS (9)	IV (10)
100 × educ	6.94	8.27	6.14	3.47	6.79	5.03	6.00	5.18	6.00	4.89
	(.03)	(1.64)	(.03)	(1.55)	(.05)	(2.93)	(.02)	(1.21)	(.02)	(.82)
First-step Wald statistic		3.44		8.79		.75		2.47		19.69
Controls:										
Year of birth indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	. . .	. . .
Quadratic trend in birth year	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	Yes	Yes
Survey year indicators	. . .	. . .	. . .	. . .	. . .	. . .	Yes	Yes	Yes	Yes
Instruments used	Quarter of birth × year of birth						Quarter of birth × year of birth		Year of birth	
Sample restrictions:										
Census year	1980		1980		1970		1960–80		1960–80	
Birth cohorts in the sample	1930–39		1940–49		1920–29		1920–35		1920–35	

Note.—The table compares year of birth with quarter of birth as instruments for education. Although the two methods yield similar estimates, year of birth is a stronger instrument and produces more precise estimates. All specifications control for race and a quartic trend in age. Specifications 1–8 use quarter of birth interacted with birth year as instruments for educational attainment, while controlling by birth year indicators in the main regression. Specifications 1–6 are comparable to cols. 5 and 6 in tables 4, 5, and 6 in Angrist and Krueger (1991). Specifications 7 and 8 pool three census surveys, 1960, 1970, and 1980, and control for survey year effects. Columns 9 and 10 report our benchmark estimates for surveys 1960, 1970, and 1980 using year of birth as instruments for education. Data are taken from the DCS. Robust standard errors are reported in parentheses.

Education is measured by the highest grade completed prior to 1992. The March CPS changed the coding to a categorical variable, which reported the degree obtained. To maintain consistency, all observations after 1991 were recoded to match the pre-1992 variable. In particular, if the educational attainment was first to fourth grade, the number of years of schooling completed is set to 3. Similarly, for the category of fifth to eighth grades, years of completed education is set to 7. These are the median number of years completed in these categories in the survey years prior to 1992. Grades 9–12 are reported separately and were therefore coded accordingly. Also, 1–3 years of college was assigned 14 years of schooling. The 4+ years of college category and higher educational categories were assigned 16 years of education. The data prior to 1992 do not report degrees beyond college. All observations that reported more than 4 years of college prior to 1992 were recoded to 16 years of schooling.

Total numbers of weeks worked were categorized before 1976. Each category is replaced by the average number of weeks worked in surveys after 1976. The sample is limited to workers who worked 14 weeks or more during a year. Total wage and salary income is top-coded for anonymity reasons in the data, and the top codes changed almost every year. Beginning in 1996, the Census Bureau started reporting the average income of workers above the top code instead of the top code itself. To resolve this discrepancy, top-coded earnings are replaced by an estimated mean under the assumption that the distribution of weekly earnings can be approximated by a lognormal distribution.[32]

The race indicator was imputed to its 1962 definition with three categories (white, black, and other) for consistency over time. Average wage inflation is used to calculate real income in 2003 prices. Using the consumer price index to calculate real values does not change the results.

References

Acemoglu, Daron. 1998. "Why Do Technologies Complement Skills? Direct Technical Change and Wage Inequality." *Q.J.E.* 113:1055–90.

———. 2002. "Technical Change, Inequality, and the Labor Market." *J. Econ. Literature* 40 (1): 7–72.

———. 2003. "Patterns of Skill Premia." *Rev. Econ. Studies* 70 (2): 199–230.

Acemoglu, Daron, and Joshua D. Angrist. 2000. "How Large Are the Social Returns to Education? Evidence from Compulsory Schooling Laws." *NBER Macroeconomics Annual* 2000:9–59.

Aghion, Philippe, Peter Howitt, and Giovanni Violante. 2002. "General Purpose Technology and Within Group Inequality." *J. Econ. Growth* 7 (4): 315–45.

Angrist, Joshua D. 1988. "Grouped-Data Estimation and Testing in Simple Labor Supply Models." Working Paper no. 234, Indus. Relations Sec., Princeton Univ.

———. 1990. "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records." *A.E.R.* 80:313–36.

Angrist, Joshua D., and Guido W. Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treament Intensity." *J. American Statis. Assoc.* 90 (430): 431–42.

Angrist, Joshua D., and Alan B. Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" *Q.J.E.* 106:979–1014.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2008. "Trends in U.S.

[32] A standard remedy for this right censoring is to multiply the top-coded income values by a constant number (1.4, e.g., in Katz and Murphy [1992]). Estimates obtained here suggest that the average ratio is 1.36, but it varies from 1.2 to 1.7.

Wage Inequality: Re-assessing the Revisionists." *Rev. Econ. and Statis.* 90 (2): 300–323.
Autor, David, Alan B. Krueger, and Lawrence F. Katz. 1998. "Computing Inequality: Have Computers Changed the Labor Market?" *Q.J.E.* 113 (4): 1169–1213.
Baker, George, Michael Gibbs, and Bengt Holmstrom. 1994. "The Wage Policy of a Firm." *Q.J.E.* 109 (4): 921–55.
Bartel, Ann P., and Nachum Sicherman. 1999. "Technological Change and Wages: An Interindustry Analysis." *J.P.E.* 107 (2): 285–325.
Beaudry, Paul, and John DiNardo. 1991. "Implicit Contracts and the Movements of Wages over the Business Cycle." *J.P.E.* 99:665–88.
Belzil, Christian, and Jorgen Hansen. 2002. "Unobserved Ability and the Return to Schooling." *Econometrica* 70 (5): 2075–91.
Berger, Mark C. 1985. "The Effect of Cohort Size on Earning Growth: A Reexamination of Evidence." *J.P.E.* 93 (3): 561–73.
Blackburn, McKinley L., and David Neumark. 1993. "Omitted-Ability Bias and the Increase in the Return to Schooling." *J. Labor Econ.* 11 (3): 521–44.
Bound, John, David A. Jaeger, and Regina M. Baker. 1995. "Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable Is Weak." *J. American Statis. Assoc.* 90 (430): 443–50.
Bound, John, and Sarah Turner. 2002. "Going to War and Going to College: Did World War II and the G.I. Bill Increase Attainment for Returning Veterans?" *J. Labor Econ.* 20 (4): 784–815.
———. 2007. "Cohort Crowding: How Resources Affect Collegiate Attainment." *J. Public Econ.* 91:877–99.
Card, David. 1995. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." In *Aspect of Labour Market Behavior: Essays in Honour of John Vanderkamp,* edited by L. N. Christophides, E. K. Grant, and R. Swidinsky, 201–22. Toronto: Univ. Toronto Press.
———. 1999. "Causal Effect of Education on Earnings." In *Handbook of Labor Economics*, vol. 3A, edited by Orley Ashenfelter and David Card. Amsterdam: Elsevier Sci.
———. 2001. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica* 69 (5): 1127–60.
Card, David, and Alan B. Krueger. 1992. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *J.P.E.* 100 (1): 1–40.
Card, David, and Thomas Lemieux. 1996. "Wage Dispersion, Returns to Skill and Black-White Wage Differentials." *J. Econometrics* 74 (2): 319–61.
———. 2000. "Dropout and Enrollment Trends in the Post-war Period: What Went Wrong in the 1970's?" In *Risky Behavior among Youth: An Economic Analysis,* edited by Jonathan Gruber. Chicago: Univ. Chicago Press.
———. 2001. "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis." *Q.J.E.* 116 (2): 705–46.
Carneiro, Pedro, and James J. Heckman. 2002. "The Evidence on Credit Constraint in Post-secondary Schooling." *Econ. J.* 112 (482): 705–34.
Carneiro, Pedro, James J. Heckman, and Edward Vytlacil. Forthcoming. "Estimating Marginal Returns to Education." *A.E.R.*
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." *J. Econometrics* 99 (1): 1–38.
Corman, Hope. 1983. "Postsecondary Education Enrollment Responses by Recent High School Graduates and Older Adults." *J. Human Resources* 18 (2): 247–67.

Dellas, Harris, and Plutarchos Sakellaris. 2003. "On the Cyclicality of Schooling: Theory and Evidence." *Oxford Econ. Papers* 55 (1): 148–72.

Deschenes, Olivier. 2006. "Unobserved Ability, Comparative Advantage, and the Rising Return to Education in the United States, 1979–2000." Manuscript, Univ. California, Santa Barbara.

Ehrlich, Isaac, and Jinyoung Kim. 2007. "The Evolution of Income and Fertility Inequalities over the Course of Economic Development: A Human Capital Perspective." *J. Human Capital* 1 (1): 137–74.

Falaris, Evangelos M., and H. Elizabeth Peters. 1992. "Schooling Choices and Demographic Cycles." *J. Human Resources* 27 (4): 551–74.

Freeman, Richard. 1981. "Career Patterns of College Graduates in a Declining Job Market." Working Paper no. 750, NBER, Cambridge, MA.

Galor, Oded, and Omer Moav. 2000. "Ability-Biased Technological Transition, Wage Inequality, and Economic Growth." *Q.J.E.* 115:469–99.

Goldin, Claudia. 1999. "A Brief History of Education in the United States." Historical Paper no. 119, NBER, Cambridge, MA.

Goldin, Claudia, and Lawrence F. Katz. 2008. *The Race between Education and Technology*. Cambridge, MA: Belknap/Harvard Univ. Press.

Gould, Eric D. 2002. "Rising Wage Inequality, Comparative Advantage, and the Growing Importance of General Skills in the United States." *J. Labor Econ.* 20 (1): 105–47.

Greenwood, Jeremy, and Mehmet Yorukoglu. 1997. "1974." *Carnegie-Rochester Conf. Ser. Public Policy* 46:19–95.

Griliches, Zvi. 1970. "Notes on the Role of Education in Production Functions and Growth Accounting." In *Education, Income and Human Capital*. Studies in Income and Wealth, vol. 35, 71–127. New York: NBER.

———. 1977. "Estimating the Returns to Schooling: Some Econometric Problems." *Econometrica* 45 (1): 1–22.

Guryan, Jonathan. 2008. "The Race between Education and Technology: A Review Article." *J. Human Capital* 3 (2): 177–96.

Gustman, Alan L., and Thomas L. Steinmeier. 1981. "The Impact of Wages and Unemployment on Youth Enrollment and Labor Supply." *Rev. Econ. and Statis.* 63 (4): 553–60.

Guvenen, Fatih, and Burhanettin Kuruscu. 2010. "A Quantitative Analysis of the Evolution of the Wage Distribution: 1970–2000." *NBER Macroeconomics Annual*, forthcoming.

Heckman, James J., Anne Layne-Farrar, and Petra E. Todd. 1996. "Human Capital Pricing Equations with an Application to Estimating the Effect of Schooling Quality on Earnings." *Rev. Econ. and Statis.* 78 (4): 562–610.

Heckman, James J., Lance J. Lochner, and Christopher Taber. 1998. "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." *Rev. Econ. Dynamics* 1: 1–58.

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2008. "Earnings Functions and Rates of Return." *J. Human Capital* 2 (1): 1–31.

Heckman, James J., and Edward Vytlacil. 1998. "Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Average Rate of Return to Schooling When the Return Is Correlated with Schooling." *J. Human Resources* 33 (4): 974–87.

———. 2001. "Identifying the Role of Cognitive Ability in Explaining the Level of and Change in the Return to Schooling." *Rev. Econ. and Statis.* 83 (1): 1–12.

Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce. 1993. "Wage Inequality and the Rise in Returns to Skill." *J.P.E.* 101 (3): 410–42.

Katz, Lawrence F., and David H. Autor. 1999. "Changes in the Wage Structure and Earnings Inequality." In *Handbook of Labor Economics*, vol. 3A, edited by Orley Ashenfelter and David Card, 1463–1558. Amsterdam: Elsevier Sci.

Katz, Lawrence F., and Kevin M. Murphy. 1992. "Changes in Relative Wages 1963–1987: Supply and Demand Factors." *Q.J.E.* 107 (1): 35–78.

Krussel, Per, Lee Ohanian, Victor Rios-Rull, and Giovanni Violante. 2000. "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis." *Econometrica* 68:129–54.

Lang, Kevin, and David Kropp. 1986. "Human Capital vs. Sorting: The Effects of Compulsory Attendance Laws." *Q.J.E.* 101:609–24.

Lemieux, Thomas. 2006. "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *A.E.R.* 96 (3): 461–98.

Manski, Charles F., and John V. Pepper. 2000. "Monotone Instrumental Variables: With an Application to the Returns to Schooling." *Econometrica* 68 (4): 997–1010.

Mincer, Jacob. 1974. *Schooling, Experience and Earnings.* New York: Columbia Univ. Press.

Murnane, Richard J., John B. Willett, and Frank Levy. 1995. "The Growing Importance of Cognitive Skills in Wage Determination." *Rev. Econ. and Statis.* 77 (2): 251–66.

Psacharopoulos, George, and Harry A. Patrinos. 2004. "Returns to Investment in Education: A Further Update." *Educ. Econ.* 12 (2): 111–34.

Raaum, Oddbjørn, and Knut Røed. 2006. "Do Business Cycle Conditions at the Time of Labor Market Entry Affect Future Employment Prospects?" *Rev. Econ. and Statis.* 88 (2): 193–210.

Ruggles, Steven, et al. 2004. *Integrated Public Use Microdata Series: Version 3.0.* Machine-readable database. Minneapolis: Minnesota Population Center. http://www.ipums.org.

Sakellaris, Plutarchos, and Antonio Spilimbergo. 2000. "Business Cycles and Investment in Human Capital: International Evidence on Higher Education." *Carnegie-Rochester Conf. Ser. Public Policy* 52 (1): 221–56.

Staiger, Douglas, and James H. Stock. 1997. "Instrumental Variables Regressions with Weak Instruments." *Econometrica* 65 (3): 557–86.

Stapleton, David C., and Douglas J. Young. 1988. "Educational Attainment and Cohort Size." *J. Labor Econ.* 6 (3): 330–61.

Stock, James H., and Motohiro Yogo. 2005. *Testing for Weak Instruments in Linear IV Regression.* Cambridge: Cambridge Univ. Press.

Taber, Christopher. 2001. "The Rising College Premium in the Eighties: Return to College or Return to Unobserved Ability?" *Rev. Econ. Studies* 68 (3): 665–91.

Welch, Finis. 1979. "Effect of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust." *J.P.E.* 87, no. 5, pt. 2 (October): S65–S97.

Willis, Robert J., and Sherwin Rosen. 1979. "Education and Self-Selection." *J.P.E.* 87, no. 5, pt. 2 (October): S7–S36.

Wooldridge, Jeffrey. 1997. "On Two-Stage Least Squares Estimation of the Average Treatment Effect in a Random Coefficient Model." *Econ. Letters* 56:129–33.

———. 2003. "Further Results on Instrumental Variables Estimation of Average Treatment Effects in the Correlated Random Coefficient Model." *Econ. Letters* 79:185–91.