Returning to school for higher returns
Seonyoung Park *

Department of Economics, University of Minnesota, 4-101 Hanson Hall, 1925 Fourth Street South, Minneapolis, MN 55455, United States

A R T I C L E   I N F O

Article history:
Received 1 February 2011
Received in revised form 1 July 2011
Accepted 1 July 2011

JEL classification:
J31
J24

Keywords:
Returns to education
Nonlinearity
Fixed effects
Risk tolerance

A B S T R A C T

On the basis of those respondents in the National Longitudinal Survey of Youth (NLSY) who change jobs with an intervening period of education reinvestment, the conventional assumption of linearity of log wages in years of schooling is strongly rejected: a typical reinvestment for the 1980 through 1993 period is associated with a rise of about 3.5 percentage points in the estimated return to an additional year of schooling. The estimated marginal rate of return generally rises in the former education level, and reaches the maximum at 15 years of the former level (therefore 16 years of education after reinvestment), where an additional year of investment is associated with a rise in real hourly rate of pay by approximately 20%. Evidence also shows that, while the level of individuals’ risk tolerance affects significantly the probability of returning to school, correcting for sample selectivity makes little difference in the results. Findings in the current paper survive a variety of robustness tests. The current cohort-based evidence is more helpful than existing evidence from cross-sectional data to individuals making schooling decisions.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

As well recognized by many researchers, the marginal rate of private return (henceforth return) to education is an important determinant of an individual’s education attainment and earnings. Repeated efforts have been made to identify the causal effect of education on earnings, with an emphasis on the potential bias associated with Ordinary Least Squares (OLS) estimates that may be subject to ability bias, attenuation inconsistency due to measurement errors in education, and/or, discount-rate bias (Lang, 1993; Card, 1995) . It is not the purpose of this paper to survey these issues, nor do we attempt to participate in any of these discussions with new evidence. Instead, our paper asks what type of information on the return to schooling is more helpful for individuals making decisions on their optimal levels of schooling. Does the return increase with years of investment? Or, is it diminishing with years of education? Or, is the return independent of the educational level, as is conventionally assumed in the vast majority of the literature on the education–earnings relationship? More importantly, for a marginal decision maker who tries to determine whether or not to pursue an additional year of schooling by comparing the marginal rate of return with the marginal discount rate, how much return can he/she expect at each level of education?

These questions are best answered when we analyze cohort data rather than cross-sectional data. As well explained by Heckman, Lochner, and Todd (2003), estimates of the return to different schooling levels based on cross-sectional data may not be representative of the rate of return that governs individuals’ decisions on human capital investment, especially when skill prices change during a period of economic transition or when the quality of schooling changes across different cohorts. For example,
as the relative price of college graduates increases permanently, a cohort of individuals who invest in college education just prior to the increase will enjoy higher returns to college investment than earlier cohorts do. Cross-sectional comparison, however, of college graduates and high school at a point in time would understate the true return to college (for a cohort of individuals who anticipate the rise in the skill premium), as the sample also includes earlier cohorts whose skill premium was relatively low. A similar argument can be made when school quality changes.\footnote{Using the 1964–2000 Current Population Survey (CPS) March Supplements, Heckman et al. (2003, Fig. 7a) find that the cohort-specific return to college tends to be higher than the return obtained from a cross-section of workers since the late 1970s.}

On the basis of those respondents in the National Longitudinal Survey of Youth cohort (NLSY) who return to school after a certain period of job experience and take another job after completion of reinvestment in education, we investigate nonlinearity in the return to education. To the best of our knowledge, this paper is the first cohort-based longitudinal study that deals with the nonlinearity issue. To be specific, first we test if the ability-free return becomes higher, lower, or remains the same at the higher level of education after reinvestment, relative to the lower level prior to reinvestment. The test for nonlinearity is conducted in a very flexible way without imposing any restrictive form on the return function. Rejecting the linearity assumption, we estimate a schedule of the ability-free marginal rate of return individuals can expect as they increase their education level over the course of the lifecycle. We also check robustness in the results by allowing different specifications, by controlling for the ‘Ashenfelter’s dip’ that may present in our sample of education changers, and by correcting for sample selectivity that may arise from using only those who actually decide to return to school.

Major findings are as follows. First, like many existing studies including Angrist and Newey (1991), our ability-free estimate is somewhat larger than the OLS counterpart. Second, the conventional assumption of linearity of log wages in years of schooling is strongly rejected in our cohort sample: a typical reinvestment in schooling for the 1980 through 1993 period (with the former education level being 13.4 years and the amount of reinvestment being 1.4 years) is associated with a rise of about 3.5 percentage points in the estimated return to an additional year of schooling. The change is also statistically significant. Third, the estimated return generally rises in the former education level, and reaches the maximum at 15 years of the former level (therefore 16 years of education after reinvestment), where an additional year of investment is associated with a rise in real hourly rate of pay by approximately 20%.\footnote{Although estimated returns are somewhat reduced beyond that point, we find the increasing portion of the return function more interesting, as a majority of the population completes up to 16 years of education. Our estimates based on the Current Population Survey data reveal that the population proportion of those who completed up to 16 years of schooling is about 92.0%, 91.0%, 90.2% and 92.2% in 1980, 1985, 1990, and 1995, respectively. This is approximately our sample period.} Fourth, while the level of individuals’ risk tolerance affects significantly the probability of returning to school, correcting for sample selectivity makes little difference in the results. Fifth, findings in the current paper survive a variety of additional robustness tests. The results are preserved even when sheepskin effects, year effects, and effects of the ‘Ashenfelter’s dip’, among others, are controlled for. The current cohort-based evidence of the marginal return schedule, free of ability effects, is more helpful than existing evidence from cross-sectional data to individuals making schooling decisions.

The remainder of this paper is organized as follows. Section 2 provides a brief summary of existing studies in relation to the nonlinearity issue. Section 3 introduces our sample and related data issues along with an estimation strategy. Empirical findings are reported in Section 4. Section 5 concludes the paper.

2. A review

It is a well-known fact that, in cross-sectional data, log earnings are approximately linearly related with schooling. For example, Card (1995) writes,

“The approximate linearity of the cross-sectional relation between log earnings and schooling is an important stylized fact.”

Heckman and Polachek (1974), Card and Krueger (1992), and Park (1994) all present supporting evidence of this fact. Consequently, in the literature of the schooling returns, the linearity assumption has been maintained in most studies, and relatively little attention has been paid to relaxing the linearity assumption.

Some studies address the nonlinearity issue based also on cross-sectional data. For example, on the basis of cross-country evidence, Psacharopoulos (1985, 1994) finds diminishing return to education investment with returns being the highest for primary education. Using repeated cross-sectional data from the UK General Household Survey for the 1974–1994 period, Harmon and Walker (1999) report that the marginal returns to schooling beyond age 18 is significantly lower than the marginal returns to schooling up to that age. Using the British National Child Development Survey, which have surveyed all individuals born in Britain between 3 and 9 March 1958 at five different points, Blundell, Dearden, Goodman, and Reed (2000) compare hourly wages (at age 33 observed in the fifth wave) of individuals with Higher Education (HE) qualifications with those who had the prospect of undertaking HE but chose not to. It is found that estimated return to an undergraduate degree is about 17% for men and 37% for women, and that the returns to higher degrees courses are slightly lower than those to undergraduate degrees, but still statistically significant. Using also cross-sectional individual data for 12 countries received from the International Social Survey Programme, Trostel (2005) finds that the marginal rate of return is essentially nil for the first several years of schooling, it then increases rapidly until about year 12, and then it declines. A recent study by Song, Orazem, and Wohlgemuth (2008) focuses on estimated returns to postgraduate education. Using data from the 1993 National Survey of College Graduates, they find that, with nonrandom sorting into graduate school corrected for,
estimated returns to graduate levels are not lower than those to a bachelor's degree. Another strand of research by Hungerford and Solon (1987), Belman and Heywood (1991), and Jaeger and Page (1996) also use cross-sectional data to present evidence of nonlinearity in the return that arises from sheepskin effects, where nonlinearity in the return depends on particular levels of education, not the level of education in general. To summarize, although these cross-sectional studies provide supporting evidence of nonlinearity, they differ in the shape or the peak point of the return function.4

However, as well demonstrated by Card (1999) and many other studies cited therein, OLS estimates of returns to schooling based on cross-sectional data are generally subject to various types of biases: ability bias, attenuation inconsistency due to measurement errors in education, and/or, discount-rate bias. By the same token, with a few exceptions,5 existing studies dealing with the nonlinearity issue (aforementioned) are not free of these issues either, as their OLS estimates are also based on cross-sectional data. In an effort to estimate sheepskin effects in the return, Hungerford and Solon (1987) also mention that their estimates may be biased by omission of ability or other factors correlated with their education variable. Although they conjecture that omission of these factors may not distort their results, it is an empirical matter.

More importantly, as emphasized by Heckman et al. (2003), the rate of return to schooling estimated from cross-sections of workers often differs from the true rate of return faced by cohorts making their schooling decisions. This is particularly so, when skill prices or cohort quality changes.6 Using data from 1964 to 2000 March CPS data, they find that estimated returns from a cross-section of workers are not only biased in levels, but time patterns suggested by these estimates are often inconsistent with those obtained using a cohort-based estimation strategy.7 In addition, patterns of high school completion and college attendance decisions over their sample period are generally found to be consistent with cohort-based estimates of the return to high school and college for the same period.

To fill the gap in the existing literature, we adopt data generated by cohort experience to test formally for non-

4 Psacharopoulos (1985, 1994) shows significant diminishing returns with returns being the highest for primary education, Trostel (2005) reports that returns reach the maximum at around 12 years of education, Blundell et al. (2000) find that returns are higher at the first degree (undergraduate level) compared with lower school or higher degree levels, and Hungerford and Solon (1987) find that estimated returns are generally large and significant at diploma years with the maximum being reached at completion of 16 years. Finally, contrary to Blundell et al. (2000) and Hungerford and Solon (1987), Song et al. (2008) find that estimated returns are not lower at the postgraduate level than at the undergraduate level.

5 Harmon and Walker (1999) use various instruments to find that the results of nonlinearity are sensitive to the instruments used. Blundell et al. (2000) use individuals’ math and reading ability test scores at age seven as proxy variables for their innate ability.

6 Citing Park’s (1994) work, Card (1995) also observes that the slope of the log earnings—education relation varies from year to year in the Current Population Survey data.

7 Heckman, Layne-Farrar, and Todd (1995) also test and reject the conventional assumption of linearity of the earnings—schooling relationship.

8 This is the case when the difference in the education level (reported as of May in each survey) is one year between two adjacent biennial surveys.

9 The main purpose of Angrist and Newey (1991) is to test if the fixed-effects assumption is appropriate in estimating human capital earnings function. In the process, they report that fixed-effects estimates of the return to schooling in the NLSY are roughly twice as large as OLS estimates.
is an endeavor to capture full returns to full investment in education. Finally, we include women in our sample and test if there exist gender differences in the return.

As noted by Griliches (1979) and Card (1999) among others, survey reports of the education variable are often subject to measurement errors. It is well recognized by researchers that measurement errors in reported education tend to attenuate estimated returns to education by a greater extent when we match wage changes with changes in education using longitudinal data. To reduce measurement errors in education, we use the longitudinal nature of the NLSY data to make reported education levels internally consistent. First, we restrict our sample only to the cases when the reported years of completed education (as of May) between two adjacent survey years remain the same or increase by one year. Second, as a way of pinpointing the timing of changes of education, our sample requires that education levels remain the same for at least two consecutive years both right before and right after the end of education change. As an example, if a person reports completed education as 12, 12, 12, 13 and 14 only for five consecutive survey years, the case is discarded, even though numbers are internally consistent, as it is not clear if the person completed her/his education in the last survey year.

Using a sample of only those who actually change education may produce biased results, when the choice of returning to school is not exogenously made. For example, estimated return based on a sample of actual education changers tends to overstate the true return, if the probability of returning to school is greater for those who expect greater return to an additional year of education than for otherwise comparable individuals who do not return to school. We consider this sample selectivity using a standard Heckman (1979) type two-step estimation method, where heterogeneity in individuals’ risk preference is used to identify the selection effect. A meaningful identification, however, requires that individual’s risk attitudes do not affect their wage changes, conditional on years of education reinvestment. The usage of individuals’ risk preference as an excluded instrumental variable for their educational attainment is motivated by Brunello (2002), who sets up a simple static model to show that risk aversion affects in a natural way educational choice by influencing the marginal utility of schooling. In fact, while researchers have typically used school reforms, family backgrounds, and smoking as instruments for schooling, as noted by Brunello (2002), little attention has been paid to risk aversion as an instrument for attained education probably due to difficulty of measuring risk aversion in survey data. Using the Italian Survey on Household Income and Wealth data, Brunello measures the Arrow-Pratt index of absolute risk aversion for each individual and shows that measured risk aversion can be safely excluded from the Mincerian earnings function. Unlike Brunello, however, we use individuals’ risk attitudes at the time of reinvestment, not current, which allows us to effectively control for time-varying as well as time-invariant components of individuals’ risk attitudes. The Appendix A provides a brief discussion of how individuals’ attitudes toward risk satisfy an exclusion restriction in our NLSY sample. Finally, in the empirical execution, we borrow Light and Ahn’s (2010) measure of risk tolerance computed for each respondent in the same NLSY sample that is adopted in the current study. During three out of 22 interviews conducted from 1979 through 2004, respondents were asked whether they would accept two hypothetical, lifetime income gambles of varying riskiness. Light and Ahn used multiple responses to these questions to estimate an Arrow-Pratt index of relative risk tolerance that accounts for both measurement error and aging effects.11

3.2. Econometric model

We begin with a standard Mincer (1974) wage function in estimating the returns to education:

\[ \ln W_{it} = \theta_1 + \theta_2 \text{Age}_{it} + \theta_3 \text{Age}^2_{it} + \gamma \text{Education}_{it} \delta X_{it} + \mu X_{it} + \varepsilon_{it} \]

where \( W_{it} \) is real hourly rate of pay on the main job held by individual \( i \) during the survey week in year \( t \), deflated by the Consumer Price Index (CPI) of that year, \( \text{Age} \) represents the person’s age, \( \text{Education}_{it} \) is years of education completed as of year \( t \), \( X_{it} \) and \( X_i \) are time-varying and time-invariant characteristics, respectively, and \( \varepsilon \) is the error term. Also, \( \gamma \) represents the return to an additional year of schooling for someone whose education level is \( \text{Education}_{it} \). Eq. (1) shares the common specification adopted by many existing studies, such as Ashenfelter and Krueger (1994) and Angrist and Krueger (1991), that estimate pure returns to schooling.

Following Angrist and Newey (1991), the error term is composed as follows:

\[ \varepsilon_{it} = \alpha_i + u_{it} \]

where \( \alpha_i \) represents innate ability of individual \( i \) which is probably correlated with the education level, and the individual-and-time specific error term, \( u_{it} \), is at a minimum assumed to be independent of regressors.

One easy way of controlling for \( \alpha_i \) as well as \( X_i \), and therefore estimating \( \gamma \) consistently, is to use the sample of education changers, observe wages received from the job taken after reinvestment in education, and explain the wage growth rate between the two points by corresponding years of education change. In our empirical implementation, year \( t \) is the last year wages are observed from the previous job, and year \( (t+s) \) is the first year wages are observed from the reemployed job. Subtracting Eq. (1) from the equation observed at time \( t+s \) yields

\[
\ln \frac{W_{i,t+s}}{W_{it}} = \theta_2 s_i + \theta_3 [2(s_i \times \text{Age}_{it}) + \gamma s_i^2] \\
+ \gamma (\text{Education}_{i,t+s} - \text{Education}_{i,t}) + \delta (X_{i,t+s} - X_i) \\
+ (u_{i,t+s} - u_{it}) \\
= \beta Z + \gamma (\text{Education}_{i,t+s} - \text{Education}_{it}) + \Delta u
\]

---

10 Naturally our sample includes only employer changers, while Angrist and Newey’s sample includes both employer stayers and employer-to-employer changers.

11 The data were generously supplied by the authors.
where $\beta Z = \theta_2 s_i + \theta_3 [2(s_i \times \text{Age}_{it}) + s_i^2] + \delta'(X_{it+5} - X_{it})$ and $\Delta u = u_{it+5} - u_{it}$. In Eq. (3), all individual-specific but time-invariant characteristics, observable or not, are differentiated out.

Ordinary Least Squares (OLS) estimation of Eq. (3) would still produce inconsistent estimates of $\gamma$, if education changes are not randomly determined. To be specific, if changers and non-changers are systematically different in some attributes that are correlated with the error term in Eq. (3), estimation of Eq. (3) by OLS may be subject to omitted variable bias. Suppose decision of education change is affected primarily by the person’s attitude toward risk:

$$\text{Education}_{it+5} - \text{Education}_{it} = \lambda_0 + \lambda_1 R_{it} + v_{it}$$ (4)

where $R_{it}$ represents the extent of risk tolerance individual $i$ has at the time of making a decision of whether or not to return to school, year $t$. Then, conditional on education change,

$$E \left( \ln \left( \frac{W_{it+5}}{W_{it}} \right) (\text{Education}_{it+5} - \text{Education}_{it}) > 0 \right) = \beta' Z + \gamma (\text{Education}_{it+5} - \text{Education}_{it}) + E[\Delta u | v_{it} > 0]$$

$$= (\lambda_0 + \lambda_1 R_{it})$$

$$= \beta' X + \gamma (\text{Education}_{it+5} - \text{Education}_{it}) + IMR(R_{it})$$ (5)

where $IMR(R_{it}) = \rho_{\Delta u, v} \sigma_{\Delta u} \frac{(\Phi(\lambda_0 + \lambda_1 R_{it}))/(\Phi(\lambda_0 + \lambda_1 R_{it}))}$, where $\rho_{\Delta u, v}$ represents the correlation of the two error terms in Eqs. (3) and (4), $\sigma_{\Delta u}$ is standard deviation of the error term in Eq. (3), and $\Phi(\cdot)$ and $\phi(\cdot)$ are a density and a cumulative distribution function of a standardized normal random variable, respectively. If education changers tend to enjoy higher wage growth than otherwise comparable non-changers ($\rho_{\Delta u, v} > 0$), estimated return from Eq. (3) is subject to omitted variable bias.

Specification of Eq. (5) is restrictive in the sense that the return to an additional year of education is the same as $\gamma$ regardless of the education level already accumulated. To allow the rate of return to vary with the education level, we construct an unrestricted model:

$$E \left( \ln \left( \frac{W_{it+5}}{W_{it}} \right) (\text{Education}_{it+5} - \text{Education}_{it}) > 0 \right) = \beta' Z + \gamma_{it+5} - \gamma^I \text{Education}_{it+5}$$

$$+ \gamma^I (\text{Education}_{it+5} - \text{Education}_{it}) + IMR(R_{it})$$ (6)

$$E \left( \ln \left( \frac{W_{it+5}}{W_{it}} \right) (\text{Education}_{it+5} - \text{Education}_{it}) > 0 \right) = \beta' Z + \gamma_{it+5} - \gamma^T \text{Education}_{it}$$

$$+ \gamma^T (\text{Education}_{it+5} - \text{Education}_{it}) + IMR(R_{it})$$ (7)

where $\gamma^I$ and $\gamma^T$ represent the marginal returns at the education levels before and after reinvestment, respectively. Eqs. (6) and (7) are identical except for the time index used: in Eq. (6), the educational level and the coefficient of the education change are indexed at $(t+5)$ and $t$, respectively, while Eq. (7) uses opposite time indices. In both equations, significance of the coefficient of the educational level implies non-linearity in the returns to schooling with a positive and negative number showing increasing and decreasing returns to education, respectively. Eqs. (6) and (7) are flexible in the sense that no specific form is specified for the return function. As such, our test results are robust to misspecification of the return function.

Lastly, for the purpose of estimating a more concrete schedule of the marginal rate of return, a set of dummy variables for the former education level are constructed, interacted with education change, and are included in the place of the education change variable in Eq. (5). The coefficient of each interaction term represents the marginal return a potential investor would expect at each education level he/she already accumulated.

4. Empirical findings

4.1. Sample characteristics

Table 1 summarizes education "changes" experienced by individuals in our NLSY random subsample. With the poverty and the military sample excluded, there are originally 6111 individuals aged 14–22 as of 1979. The entire annual surveys from 1979 through 1994 are used in the current analysis. With all the restrictions to reduce measurement errors in education, 1142 individuals are identified to have changed their education levels at certain points in time. Due to the sample restriction that reported education levels be the same at least for two consecutive years right before and right after education change, the first education change observed in our sample period is 1980 and the last one ended in 1993. Among 1142 respondents, approximately 16% experience education change more than one time. Figures in the second column are based on the sample which is used in our regression analysis. In addition to those who do not have valid observations on the variables used in the regression, those "on-the-job" education changers who change their educational level as an employee (Angrist and Newey’s sample) are also dropped from the final sample, which leaves us with 885 person-cases experienced by 767 individuals. However, relative frequency distributions remain very similar between the two columns.

<table>
<thead>
<tr>
<th>Number of changes</th>
<th>Number of individuals (% of the total)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Original sample</td>
</tr>
<tr>
<td>1</td>
<td>958 (0.8389)</td>
</tr>
<tr>
<td>2</td>
<td>169 (0.1480)</td>
</tr>
<tr>
<td>3</td>
<td>13 (0.0114)</td>
</tr>
<tr>
<td>4</td>
<td>2 (0.0018)</td>
</tr>
<tr>
<td>Total</td>
<td>1142</td>
</tr>
</tbody>
</table>

Data source: National Longitudinal Survey of Youth for the 1980–93 period. Figures are based on the random subsample, excluding the poverty and the military sample.
Using the entire sample in the first column of Table 1, Fig. 1 displays frequency of education change by the size of education change and by the former education level. Among the total number of case-person observations (1344), 73% pertain to a one-year change in education, 17% to a two-year change, 7% to a three-year change, and 2% belong to a four-year change in education. The former education level before returning to school varies greatly with the 12th grade being the most frequently observed level. Returning to school to benefit from the “sheepskin effects” is also observed in the data. For example, although one-year upgrading is the most-frequently observed in most cells, comparison of within-cell relative frequencies across cells reveals that one-year, two-year, three-year, and four-year change appear the greatest as 93%, 42%, 13%, and 3%, when the former educational levels are 15, 14, 13, and 12 years, respectively.

Although, for brevity, not reported in a separate table, the sample mean of is 2.84 years, in comparison with 1.41 years of sample mean of education change. The difference between the two sample means is statistically significant. Standard deviations of and education change are 2.12 and 0.82 years, respectively. Lastly, in our regression sample, the sample mean, standard deviation, the minimum, and the maximum value of risk tolerance are 0.89, 1.09, 0.12, and 9.02, respectively.

4.2. Estimated returns to education

Table 2 reports the estimated returns to education. The estimates in column 1 through 6 are obtained by our differenting method and, therefore, free of ability bias. F-tests cannot reject the null hypothesis of no gender difference in regression coefficients in either our basic model (column 1 through 3) or the model with additional control variables (column 4 through 6). Our ability-free estimate in column 6 shows that an additional year of schooling leads to a rise in the wage change by 0.0881** (0.0396) per 0.0014 (0.0001) in standard error. The returns to education are similar for women and men with both coefficients being equal to 0.0860*** (0.0278) for both men and women. The returns to education are similar for women and men with both coefficients being equal to 0.0860*** (0.0278) for both men and women.
in real hourly wage rate by 8.6%. This estimate is very similar to the ability-free estimates found in existing studies, in particular to those in twin-based studies and the one obtained by Angrist and Newey (1991). The similarity between the current study and Angrist and Newey (1991) is notable considering differences in the sample between the two studies: while we consider cases where individuals either take jobs or are in school, not both, respondents in Angrist and Newey’s sample attend school as an employee.

Estimates in the last column are obtained by applying OLS to the level equation based on the very sample used in estimating the differenced equation. In this “balanced” sample, real hourly wages increase by 7.3% following an additional year of schooling, which is somewhat smaller than the ability-free estimate. This finding is quite consistent with recent estimates of the return to education, as well summarized in Card (1995). In particular, Angrist and Newey’s results also show a higher return to education from the fixed-effects model than the corresponding OLS model, although the difference between the fixed-effects estimate and the OLS one is somewhat larger in Angrist and Newey than the current study.

Estimates in Table 2 may be subject to a selection bias by including in the sample only those who actually changed their education level. If the probability of returning to school is greater for those who expect greater returns to an additional year of education than for otherwise comparable individuals who do not come back to school, the estimated return based on actual education changers tend to overstate the true return. To check the potential selection bias associated with using the endogenously selected sample of education changers, Table 3 reports estimates based on Eq. (5). Column 1 imports our ability-free estimates in column 6 of Table 2, and the remaining columns show selectivity-corrected ability-free estimates. Looking at the estimated Probit equation in the last column, it is evident that a greater level of risk tolerance increases the probability of returning to school. In addition, the inverse Mills ratio enters significantly into all wage growth equations from column 2 through 6, implying that the probability of returning to school and the wage growth rate are positively correlated. Including the inverse Mills ratio in the wage growth equation, however, does not reduce estimated return significantly, as is evident in comparison of estimates between column 1 and 2. In column 3, we include a dummy variable for diploma years as additional regressor, which equals one if reinvestment of schooling ends up with 8, 12, or 16 years of education, and found that controlling for sheepskin effects reduces estimated return to education slightly. With year dummies added additionally (column 4), estimated return is reduced slightly to 7.1%. Estimated sheepskin effects are significant, whether or not year dummies are controlled for.

Estimated returns obtained by our differencing method may exaggerate true returns if wages at the timing of decision making \( W_{it} \) in Eq. (5) are already subject to the ‘Ashenfelter’s dip’, which is an empirical regularity that the mean earnings of participants in employment and training programs generally decline during the period just prior to participation (Ashenfelter, 1978). Our basic method of dealing with the ‘Ashenfelter’s dip’ is to use more lagged wages from the pre-investment job \( W_{i(t-1)} \) and subtract them from post-investment wages \( W_{i(t+1)} \) so that wage changes observed over a longer period are less affected by the ‘dip’ impact. Given that the sample size adopted in the current analysis is relatively small, we examine pre-investment wages lagged by one year \( (t = 1) \). We first restrict our sample to those who are in the same job and report wages for at least two consecutive years prior to reinvestment, which reduces the number of our differenced observations in our full sample dramatically to 422. Then, using the 422 observations, we re-estimate Eq. (5) with the most expanded specification in column 4 of Table 3, and report the results in column 5 of Table 3. Finally, using the same ‘balanced sample’, we redo the analysis with the dependent variable replaced by \( \ln(W_{i(t+1)}/W_{i(t-1)}) \), whose results are in column 6. In the balanced sample, estimated return is somewhat reduced from 10.1% to 8.0% by adopting longer differencing, supporting evidence for the ‘Ashenfelter’s dip’. As will be demonstrated in a subsequent section, however, our main results of nonlinearity in returns are quite robust to consideration of the ‘dip’.

4.3. Nonlinearity in the returns to schooling

Table 4 reports estimated Eqs. (6) and (7), where the return to education is estimated in a more flexible way: the return is allowed to vary with the education level and no functional form is pre-specified about non-linearity of the return. Because the two equations produce identical results except for the coefficient of education change, we report the estimated coefficient of education change in Eq. (6) \( \gamma^a \) and that in Eq. (7) \( \gamma^{a+b} \) in the same column along with other estimated coefficients. Results show a strong evidence for increasing returns to education. For example, figures in column 3 show that, all the other things being constant, the estimated return increases from 4.3% at the former education level to 8.6% at the higher education level after reinvestment. The difference (4.3 percentage points), the estimated coefficient of the education level, is statistically significant even at the 1% significance level. The difference remains significant even when we control additionally for sheepskin effects (column 4) and for year effects (column 5). Column 6 uses the same specification as column 5 except that, in column 6, s is replaced by \( s - \Delta Education \). This is the case where Eq. (1) is specified as a quadratic function of potential experience, not age.

---

13 Assuming that \( u_i \)'s are identically distributed, the variance of the differenced error becomes greater if \( \text{Cov}(u_{i1}, u_{i2}) \) decreases in \( i \). In all cases, however, a \( \chi^2 \)-test cannot reject the homoskedasticity hypothesis even at the 10% significance level.

14 See Table 6 of Card (1999) for a good summary of these twin studies.

15 Although not reported for brevity, we also try \( t = 2 \) with an even smaller sample and find little evidence of the ‘Ashenfelter’s dip’ at \( (t = 1) \), implying that the ‘dip’ impact is prominent just prior to the period of investment. This line of tests, however, generally suffers from a small sample problem.

16 An F-test cannot reject the null hypothesis of no gender difference in regression coefficients at the 10% significance level \( (F_{8.659} = 0.58) \).
Table 3
Selectivity-corrected ability-free estimates.

<table>
<thead>
<tr>
<th>Variables</th>
<th>OLS on wage change equation</th>
<th>Selectivity-corrected estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wage change equation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Full sample</td>
<td>Effects of the Ashenfelter's dip</td>
</tr>
<tr>
<td></td>
<td>Probit equation</td>
<td></td>
</tr>
<tr>
<td>( \Delta \text{Education} )</td>
<td>0.0863*** (0.0278)</td>
<td>0.0862*** (0.0278)</td>
</tr>
<tr>
<td>( S )</td>
<td>0.0674 (0.0473)</td>
<td>0.0780* (0.0476)</td>
</tr>
<tr>
<td>( 2 \times \text{Age} \times S + S^2 )</td>
<td>–0.0013* (0.0008)</td>
<td>–0.0014* (0.0008)</td>
</tr>
<tr>
<td>( \Delta \text{Union} )</td>
<td>–0.0333** (0.0146)</td>
<td>–0.0337** (0.0146)</td>
</tr>
<tr>
<td>( IMR )</td>
<td>–0.1632*** (0.0502)</td>
<td>0.1636*** (0.0501)</td>
</tr>
<tr>
<td>Sheepskin effects</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Year dummies</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Risk tolerance</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>( N )</td>
<td>885</td>
<td>885</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.0366</td>
<td>0.0400</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

The sample includes both men and women. In column 1 through 6, the dependent variable is change in logarithm of the real hourly rate of pay between the last survey week point before education change and the first survey week point after. In the last column, the dependent variable is a dummy variable which equals one for education changers and zero for non-changers. Each equation also includes an intercept term. \( s \) represents time distance in years between the two survey week points, and a subscript t represents the last survey week year before returning to school. \( IMR \) stands for inverse Mills ratio. Risk tolerance is represented by estimated Arrow-Pratt index of relative risk tolerance, supplied by Light and Ahn (2010). Sheepskin effects is a dummy variable which equals one if the education level is 8, 12, or 12 years after education change. Numbers in parentheses stand for estimated standard errors.

* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.
Table 4
Testing for nonlinearity in returns to education.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Men</th>
<th>Women</th>
<th>All</th>
<th>All</th>
<th>All</th>
<th>All</th>
<th>Effects of the Ashenfelter’s dip</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ln $W_{t+1}$</td>
<td>In $\frac{W_{t+1}}{W_{t-1}}$</td>
<td>ln $W_{t+1}$</td>
<td>In $\frac{W_{t+1}}{W_{t-1}}$</td>
<td>ln $W_{t+1}$</td>
<td>In $\frac{W_{t+1}}{W_{t-1}}$</td>
<td>ln $W_{t+1}$</td>
</tr>
<tr>
<td>Education level</td>
<td>0.0542*** (0.0177)</td>
<td>0.0314* (0.0167)</td>
<td>0.0429*** (0.0121)</td>
<td>0.0351*** (0.0123)</td>
<td>0.0347*** (0.0126)</td>
<td>0.0339*** (0.0124)</td>
<td>0.0390*** (0.0175)</td>
</tr>
<tr>
<td>$\Delta$Education in Eq. (6)</td>
<td>0.0242 (0.0429)</td>
<td>0.0608 (0.0418)</td>
<td>0.0432 (0.0299)</td>
<td>0.0448 (0.0298)</td>
<td>0.0397 (0.0301)</td>
<td>0.0504 (0.0279)</td>
<td>0.0679 (0.0512)</td>
</tr>
<tr>
<td>$\Delta$Education in Eq. (7)</td>
<td>0.0784* (0.0387)</td>
<td>0.0922* (0.0397)</td>
<td>0.0861** (0.0276)</td>
<td>0.0799*** (0.0275)</td>
<td>0.0744** (0.0280)</td>
<td>0.0843*** (0.0269)</td>
<td>0.1068* (0.0491)</td>
</tr>
<tr>
<td>$S$</td>
<td>0.1451 (0.0752)</td>
<td>0.1085 (0.0650)</td>
<td>0.1282*** (0.0494)</td>
<td>0.1089*** (0.0495)</td>
<td>0.0634 (0.0537)</td>
<td>0.0785 (0.0593)</td>
<td>0.1158 (0.0946)</td>
</tr>
<tr>
<td>$2 \times Age_{-1} \times S + S^2$</td>
<td>−0.0026*** (0.0012)</td>
<td>−0.0018* (0.0010)</td>
<td>−0.0022*** (0.0008)</td>
<td>−0.0018*** (0.0008)</td>
<td>−0.0010 (0.0010)</td>
<td>−0.0013 (0.0010)</td>
<td>−0.0015 (0.0016)</td>
</tr>
<tr>
<td>$\Delta$Union</td>
<td>0.1125 (0.0724)</td>
<td>0.2275*** (0.0683)</td>
<td>0.1658*** (0.0498)</td>
<td>0.1710*** (0.0496)</td>
<td>0.1647*** (0.0499)</td>
<td>0.1645*** (0.0499)</td>
<td>0.1575*** (0.0706)</td>
</tr>
<tr>
<td>IMR</td>
<td>0.4231 (0.2299)</td>
<td>0.3863 (0.4841)</td>
<td>0.3463 (0.1882)</td>
<td>0.3829*** (0.1876)</td>
<td>0.3722** (0.1089)</td>
<td>0.3722** (0.1087)</td>
<td>0.6665 (0.3141)</td>
</tr>
<tr>
<td>Sheepskin effects</td>
<td>X</td>
<td>X</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Year dummies</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>$N$</td>
<td>461</td>
<td>424</td>
<td>885</td>
<td>885</td>
<td>885</td>
<td>885</td>
<td>422</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.0502</td>
<td>0.0687</td>
<td>0.0536</td>
<td>0.0647</td>
<td>0.0825</td>
<td>0.0829</td>
<td>0.0789</td>
</tr>
</tbody>
</table>

The dependent variable is change in logarithm of the real hourly rate of pay between the last survey week point before education change and the first survey week point after. Each equation also includes an intercept term. Coefficients of Education level represent changes in returns to education, those of $\Delta$Education in Eq. (6) represent returns to education at the former level of education, and those of $\Delta$Education in Eq. (7) stand for returns to education after education change. Except for $\Delta$Education, coefficients of all the other variables are identical between the two equations. In column 6, $s$ is replaced by $s - \Delta$Education. Numbers in parentheses stand for estimated standard errors.

* Significant at the 10% level.

** Significant at the 5% level.

*** Significant at the 1% level.
As the education level increases by $\Delta Education$, potential experience increases by $s - \Delta Education$ between $t$ and $t + s$. This exercise, however, does not change any of our previous results. Last two columns test if the current results are robust with respect to consideration of the ‘Ashenfelter’s dip’. For the most expanded specification in column 5, we follow the same procedure as in column 5 and 6 of Table 3. Results show that this exercise makes little change to the observed pattern of increasing returns. In fact, the estimated coefficient of education level becomes greater when we consider the ‘dip’ effects in column 8.

While estimates in Table 4 suggest that estimated returns increase as we invest more in education, the results are not entirely informative in the sense that they hold for a typical or an average case: a typical education changer whose former education level is 13.4 years invests about 1.4 years and enjoys an additional return to education of 3–4 percentage points from a post-reinvestment job. A potential investor in education would require a more complete return schedule, the size of expected return at each education level. For that purpose, we create seven dummy variables for the former education level (less than or equal to 11, 12, 13, 14, 15, 16, and more than or equal to 17), interact them with education change, and include these seven interaction terms in the place of education change in Eq. (5).

While, for brevity, estimated coefficients of other control variables are not reported, they are generally very similar to those in Table 4. The first three columns of estimates in Table 5 correspond to the three sets of estimates from column 3 to 5 in Table 4, and the last two columns examine effects of the ‘Ashenfelter’s dip’ on the current issue. Focusing on the third column where a full set of control variables are included, estimates confirm the existence of nonlinearity in the return. An F-test rejects a null hypothesis of inter-education group homogeneity in the marginal return at the 5% significance level ($F_{6,885} = 2.11$). And, estimated returns continue to rise until the former education level reaches 15 (so that an additional year of investment ends up with 16 years of education), where an additional year of schooling is associated with a rise in the real wage rate by about 20%. After that, the marginal return drops substantially and then rises again to the level of return at 14 years of former education. This pattern is preserved in all three columns.

Additional tests are conducted about robustness of the observed pattern of nonlinearity. First, controlling for effects of the ‘Ashenfelter’s dip’ does not change the observed pattern of nonlinearity, as is evident from the last two columns of Table 5. Increased standard error estimates that result from the reduced sample size make us barely accept the null hypothesis of inter-group homogeneity in the return at the 10% significance level ($F_{6,397} = 1.51$ in the last column).

Second, including group dummies for diploma years separately, one for 12 years and the other for 16 years in the regression equation makes little difference in the observed pattern of nonlinearity. With this new exercise, the seven estimated coefficients in our most preferred specification (the most expanded specification in the third column of Table 5) change to 0.0416, 0.0517, 0.0620, 0.1058, 0.1747, 0.0899, and 0.1672, respectively, with the last four coefficients generally decreasing with the diploma years.

### Table 5

<table>
<thead>
<tr>
<th>Education level</th>
<th>Full sample</th>
<th>( \Delta E_{0.0416} )</th>
<th>( \Delta E_{0.0517} )</th>
<th>( \Delta E_{0.0620} )</th>
<th>( \Delta E_{0.1058} )</th>
<th>( \Delta E_{0.1747} )</th>
<th>( \Delta E_{0.0899} )</th>
<th>( \Delta E_{0.1672} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>0.0618</td>
<td>0.0615</td>
<td>0.0617</td>
<td>0.0622</td>
<td>0.0622</td>
<td>0.0622</td>
<td>0.0622</td>
<td>0.0622</td>
</tr>
<tr>
<td>12</td>
<td>0.1006</td>
<td>0.1006</td>
<td>0.1006</td>
<td>0.1006</td>
<td>0.1006</td>
<td>0.1006</td>
<td>0.1006</td>
<td>0.1006</td>
</tr>
<tr>
<td>13</td>
<td>0.1595</td>
<td>0.1595</td>
<td>0.1595</td>
<td>0.1595</td>
<td>0.1595</td>
<td>0.1595</td>
<td>0.1595</td>
<td>0.1595</td>
</tr>
<tr>
<td>14</td>
<td>0.2284</td>
<td>0.2284</td>
<td>0.2284</td>
<td>0.2284</td>
<td>0.2284</td>
<td>0.2284</td>
<td>0.2284</td>
<td>0.2284</td>
</tr>
<tr>
<td>15</td>
<td>0.2973</td>
<td>0.2973</td>
<td>0.2973</td>
<td>0.2973</td>
<td>0.2973</td>
<td>0.2973</td>
<td>0.2973</td>
<td>0.2973</td>
</tr>
<tr>
<td>16</td>
<td>0.3662</td>
<td>0.3662</td>
<td>0.3662</td>
<td>0.3662</td>
<td>0.3662</td>
<td>0.3662</td>
<td>0.3662</td>
<td>0.3662</td>
</tr>
<tr>
<td>&gt;17</td>
<td>0.4351</td>
<td>0.4351</td>
<td>0.4351</td>
<td>0.4351</td>
<td>0.4351</td>
<td>0.4351</td>
<td>0.4351</td>
<td>0.4351</td>
</tr>
<tr>
<td>( N )</td>
<td>885</td>
<td>885</td>
<td>885</td>
<td>885</td>
<td>885</td>
<td>885</td>
<td>885</td>
<td>885</td>
</tr>
</tbody>
</table>

\( \Delta E \) is the marginal rate of return by education level.
coefficients estimated significantly. Due to relatively large standard error estimates, however, the null hypothesis of inter-group homogeneity in the return cannot be rejected even at the 10% level. A larger sample size would change the result. For example, when we omit the union dummy variable from the regression, the number of differenced observations increases from 885 to 973 (due to missing values of the union variable), and, although not reported for brevity, all our previous findings are preserved in this new sample. In addition, this slightly increased sample makes us reject the null hypothesis of inter-group homogeneity in estimated returns at the 10% level, even though we allow different sheepskin effects for different diploma years ($F_{6,947} = 1.78$). It is also worth noting that estimated coefficient of the dummy for completion of 16 years is 0.136 with associated standard error estimate 0.060, supporting for existence of sheepskin effects.

Finally, we redo most of previous analyses using average hourly earnings, defined by the ratio of total annual labor income from all jobs to total annual hours, instead of the hourly rate of pay from the main job. We also try a sample of 656 individuals who experience education change only once. All the current findings are preserved even with these two exercises.

Estimates in Table 5 suggest that the general shape of the estimated return function appears similar between the current study and existing ones in the sense that estimated returns go up to a certain level of investment and then decline. However, aside from differences between cohort-based evidence and cross-sectional evidence, previously mentioned, and different peak points of the estimated return function, the current study also departs from existing ones by formally testing if observed nonlinearity is generated by sheepskin effects. Existing cross-sectional studies often estimate returns to different degrees, not to different years, which makes it difficult to identify if the observed nonlinearity is due to education or credentials at specific stages of education. As noted by Trostel (2005), similar difficulty is found in estimating returns as a smoothing function of years of education. Hungerford and Solon (1987) are most comparable to the current study in that they also estimate the return by an unrestricted step function. In fact, their large CPS sample, May 1978, enables them to estimate the step at each stage of education from one through eighteen years. Table 2 of their results show that, with an exception of thirteen years of investment (dropout at the first year of college), step-specific returns are large and significant only at diploma years. Unlike Hungerford and Solon, the current study finds that estimated returns increase up to 16 years even after sheepskin effects are controlled for.

In Table 6, we conduct a very preliminary test of whether or not the observed nonlinearity is related with technological progress. In measuring technological change, we follow suggestions of Bartel and Sicherman (1999) to use the NBER total factor productivity growth series described in Bartelsman and Gray (1996) (henceforth TFP), the ratio of R&D funds to net sales reported by the National Science Foundation (R&D – RATIO), and the ratio of scientific and engineering employment to total employment calculated from the 1979 and 1989 CPS by Allen (1996) (SCIENTISTS – RATIO). We compute technological change between $t$ and $t + s$ by the growth rate of total factor productivity between the two time points. For $R&D – RATIO$ and $SCIENTISTS – RATIO$, we simply calculate the difference in the ratio between $t$ and $t + s$. Then, in Eq. (6) or (7), we allow the coefficient of the education level, which reflects change in the return ($Y_{d}^{t} - Y_{d}^{t+s}$), to be a linear function of technological change observed for the same period of individual investment, which is equivalent to including in the equation an interaction of the education level and the technological change variable as an additional regressor. In the expanded equation, therefore, measures of technological change vary across individuals depending on starting and ending time points of education reinvestment. For brevity,

17 Although, unlike Hungerford and Solon, we find substantial returns even for dropouts, the current analysis does find evidence of sheepskin effects, which is different from Layard and Psacharopoulos (1974). The first and the second variables are directly imported from the NBER productivity database and the National Science Foundation, respectively. For the third technology variable, we follow Allen (1996) to compute the series using the Current Population Survey data for an expanded period from 1979 to 1994. While Bartel and Sicherman (1999) suggest three more technology variables in addition to the above three, they are not adopted by the current study, because those variables are not available at least for part of our sample period.
Table 6 reports coefficients of only education-related variables. As the two equations produce virtually identical results regarding the coefficient of the education level, we report estimation results of Eq. (7) only. In this test, we use the full set of control variables, as appeared in the fifth column of Table 4. In this set-up, a coefficient of the education level implies change in the return that is not related with technological change, and a coefficient of the education level interacted with technological change means change in the return associated with technological change. Results show that, no matter what variables are used to represent technological change, the coefficient of the education level remains unchanged, as appeared in the estimate in the fifth column of Table 4. In all cases, the estimated coefficient of the interaction term enters the equation very insignificantly, and it appears even negative when SCIENTISTS – RATIO is used. This implies that the pattern of increasing returns to investment observed in our cohort sample is not explained by concurrent technological change.19

5. Conclusion

This paper is the first cohort-based study that provides a relatively complete schedule of ability-free returns by the education level. On the basis of those respondents in the National Longitudinal Survey of Youth (NLSY) who return to school after a certain period of job experience, this paper finds strong evidence of increasing returns to investment: a typical reinvestment for the 1980 through 1993 period is associated with a rise of about 3.5 percentage points in the estimated return to an additional year of schooling. Estimated returns to an additional year of schooling continue to rise in the former education level, reach the maximum at 15 years of the former level (so that reinvestment ends up with 16 years of education), and decline beyond that point. Cohort-based estimates like current ones are helpful for individuals making schooling decisions.

The current results survive a variety of robustness tests. While the level of individuals’ risk tolerance affects significantly the probability of returning to school, correcting for sample selectivity makes little difference in our results. The current findings also remain valid when we control for year effects, effects of the ‘Ashenfelter’s dip’, and even for sheepskin effects. The observed pattern of increasing returns to education is not explained even by concurrent technological change. These observations suggest that an individual’s lifecycle production is a convex function of education until she/he invests up to 16 years in education.

A number of important limitations need to be considered. The main weakness of this study is the small sample, which limits our analysis in several ways. Not only does it prevent us from reporting a full schedule of the marginal rate of return, but also it limits test power of various hypotheses and even hinders us from testing some hypothesis more effectively. For example, a larger sample would allow us to investigate the issue of the ‘Ashenfelter’s dip’ more thoroughly, by examining more lagged wages.

Another related issue is that the current study exploits information on education investment and wage outcome obtained from an early stage of work career. In our sample, respondents return to school at age 24 on average, and 35 at the oldest. To the extent on-the-job training has differential effects on wages before and after reinvestment, so that wage growth rates are affected by this factor, our estimates will be biased. Finally, it is also acknowledged that the proxy variables for technological change adopted in the current study is far from ideal. In our wage change equation, they vary across individuals depending on when they increase their education attainment and how long they invest. Ability of the measures of technological change to explain nonlinearity in estimated returns may be increased by allowing more variation in these variables, e.g., across industries or regions.

Acknowledgements

We are grateful to two anonymous referees for their extremely helpful comments and suggestions. Valuable comments were also received from seminar participants at the University of Minnesota. All remaining errors are our own.

Appendix A. Discussion of exclusion restrictions of instrumental variables

In an effort to deal with endogeneity of schooling that arises from unobserved ability and measurement errors, researchers have often used instrumental variables that affect schooling but not earnings conditional on schooling. Examples include school reforms and institutional source of variation of schooling (e.g., minimum school leaving age, tuition costs or geographic proximity of schools), family background (e.g., parents’ education attainment, number of siblings, birth order), smoking habits, and the use of sample of twins.20 Recently, Brunello (2002) presents another candidate, individuals’ attitudes toward risk. Brunello sets up a simple static model, where individual risk aversion affects schooling because it affects the marginal utility of income (and consumption), but does not affect the marginal returns to schooling. He also tests empirically if risk attitude affects the log earnings of individuals with the same education attainment by influencing their occupational choice. Using various definitions of job/occupation, he finds that conditional on schooling and other characteristics, individuals’ risk attitudes do not affect their occupational choice.

Our choice of individuals’ risk attitude as an exclusion restriction is primarily motivated by Brunello. Following Brunello, we examine if, conditional on schooling and other characteristics, risk aversion affects individuals’ occupa-

---

19 When year dummies are excluded from the equation, estimated coefficient of the interaction term appears positive in all cases, and even significant at the 5% level when TFP is used to measure the level of technology. However, little change is made to estimated coefficients of the education level in Table 6, even with omission of year dummy variables.

20 See Card (1999) for a detailed review of these instruments, and Fersteter and Winter-Ember (2003) for a review of articles and new evidence that use smoking habits as instruments for education attainment.
tion choice. Unlike Brunello, however, in the current study, educational choice depends on individuals’ risk attitudes at the time of the choice, not on current risk aversion. As the current study focuses on wage changes generated by education investment, we naturally investigate if, with years of education investment and other characteristics controlled for, individuals’ risk attitudes at the time of investment decision make changes to their current occupational choice. However, as noted by Bound, Brown, and Mathiowetz (2001) and many other researchers, survey reports of occupations of respondents are subject to great measurement error. In particular, the occurrence of changes in occupation is exaggerated when estimates of such changes are obtained by comparing the reports of the occupations obtained at two points in time. To reduce that type of measurement errors, we adopt five broader occupation categories than those at the 1 digit level. Then, using our full sample of 885 differenced observations, we run a probit regression of a dummy for occupation change (equals one for changers) against our risk variable along with all the regressors used in our most expanded specification. The p-value of the estimated coefficient of the risk variable is 0.704. We also test if an individual’s occupation choice after education reinvestment (at time t+s) is affected by their risk attitudes at the time of decision making (t). For that purpose, we run a multinomial logit model of the five occupation dummies against the risk variable along with other control variables. A χ^2(4)-test cannot reject the null hypothesis of zero coefficients on the risk variable in all four categories at any significance level (p-value = 0.369).

We also test directly if individuals’ attitudes toward risk at time t are correlated with their wage growth between t and (t + s), conditional on years of reinvestment and other characteristics (so the error term in equation). To conduct the Sargan test of over-identifying restrictions, at least two instrumental variables are needed in our case. We follow Evans and Montgomery (1994) and Ferster and Winter-Ember (2003) to use individuals’ smoking habits as an additional instrument for education attainment. The idea is that individuals’ smoking habits are a good predictor of discount rates, which are not correlated with their (conditional) earnings. Unlike existing studies, however, we consider individuals’ smoking status at the time of decision making, not at an early stage, say age 16. To put it in another way, we consider time-varying as well as time-invariant components of individuals’ smoking habits and risk attitudes when using these variables as instruments for education. On the contrary, as existing studies measure individuals’ smoking habits or risk attitudes at a point in time, effectiveness of these variables as instruments for education hinges on dominance of the time-invariant component of these variables. On the basis of a reduced sample of 773 observations (due to missing values of respondents’ smoking status), the Sargan test cannot reject the null hypothesis that both instruments are uncorrelated with the error term (p-value for χ^2(1) = 0.219).

Finally, using both instruments, we reproduce Tables 3–6 and obtain virtually identical results.22

References


21 These are (1) professional and managers, (2) sales and clerical, (3) craftsmen, foremen, and operatives, (4) laborers, and (5) service and private household.

22 These tables, along with a more complete set of test results of the relevance and validity of these instruments, are electronically available upon request.