Changes across Cohorts in Wage Returns to Schooling and Early Work Experiences

Jared Ashworth, *Pepperdine University*

V. Joseph Hotz, *Duke University, National Bureau of Economic Research (NBER), and Institute of Labor Economics (IZA)*

Arnaud Maurel, *Duke University, NBER, and IZA*

Tyler Ransom, *University of Oklahoma and IZA*

This paper investigates the wage returns to schooling and actual early work experiences and how these returns have changed over the past 20 years. Using the NLSY surveys, we develop and estimate a dynamic model of the joint schooling and work decisions that young men make in early adulthood and quantify how they affect wages using a generalized Mincerian specification. Our results highlight the need to account for dynamic selection and changes in composition when analyzing changes in wage returns. In particular, we find that ignoring the selectivity of accumulated work experiences results in overstatement of the returns to education.

We thank Christian Belzil, Michael Böhm, Richard Blundell, Flavio Cunha, Attila Lindner, Charles Manski, Matt Masten, Michela Tincani, and participants at various seminars and conferences for useful comments and discussions at various stages of this research. We especially wish to thank Vladislav Slanchev for providing us with a version of his statistical software for estimating discrete choice models with a latent

Submitted June 4, 2018; Accepted September 25, 2020; Electronically published July 23, 2021.

Journal of Labor Economics, volume 39, number 4, October 2021.
 https://doi.org/10.1086/711851

I. Introduction

Since the 1970s, there have been dramatic changes in the structure of the US labor market. Foremost among these is a steep increase in the college wage premium during the 1980s, followed by a slower increase thereafter (see, e.g., Katz and Murphy 1992; Card and Lemieux 2001; Carneiro and Lee 2011; Valletta 2019). The characteristics and skill accumulation of American youth also have changed over this same time period. For example, using data from the 1979 and 1997 panels of the National Longitudinal Survey of Youth (NLSY), Altonji, Bharadwaj, and Lange (2012) note an increase in skills over time along with an overall widening of the skill distribution, which appears to be driven by trends in parental education. College attendance has drastically increased, college graduation has been delayed, and the average amount of in-college accumulated work experience has gone up (see, e.g., Bacolod and Hotz 2006; Bound, Lovenheim, and Turner 2012; Scott-Clayton 2012). Accounting for these changes in composition is important to understand how the premium for skill investment has evolved over time.

Our paper addresses three related research questions. First, what are the cross-cohort changes in the wage returns to schooling and early-career work experiences? Second, how much of the cross-cohort change in the college wage premium actually reflects an increase of in-school and, more generally, early work experience? Third, how did the returns to cognitive ability and other noncognitive skills change across cohorts of young men? Answering these questions requires controlling for selection into schooling and work experiences. We do this by specifying and estimating, for two different cohorts, a dynamic model of schooling and work decisions. We estimate this model with data on two longitudinal data sets, the 1979 (NLSY79) and 1997 (NLSY97) panels of the NLSY, with our 1979 panel containing data on young men born in 1959–64 and our 1997 panel containing data on those born in 1980–84.

Our use of longitudinal rather than repeated cross-sectional data allows us to more accurately measure early-career schooling and actual accumulated work experiences and to account for their endogeneity.[1] From each of the NLSY surveys, we construct comparable measures of schooling, employment,

factor error structure that we used in our structural analysis and for his generous assistance in adapting it to our context. Any remaining errors are ours. This research was conducted with restricted access to Bureau of Labor Statistics (BLS) data. The views expressed here do not necessarily reflect the views of the BLS. Contact the corresponding author, V. Joseph Hotz, at hotz@econ.duke.edu. Information concerning access to the data used in this paper is available as supplemental material online.

[1] See also Bacolod and Hotz (2006), Altonji, Bharadwaj, and Lange (2012), Castex and Dechter (2014), Lee, Shin, and Lee (2015), Böhm (2020), and Deming (2017), who also have used NLSY data to make cross-cohort comparisons about the labor market.

and military histories from ages 16 through 35, along with comparable measures of earnings, educational attainment, cognitive skill, local labor market and higher-education conditions, and personal and family background characteristics. From these histories, we are able to construct measures of multiple dimensions of human capital investment, including whether work experience occurred simultaneously with schooling. Of particular relevance for us is the important work of Altonji, Bharadwaj, and Lange (2012), who also use NLSY79 and NLSY97 surveys to document the cross-cohort changes in the unobserved distribution of cognitive ability. We follow Altonji, Bharadwaj, and Lange (2012) to construct comparable measures of correlates of unobserved cognitive ability from the Armed Services Vocational Aptitude Battery (ASVAB) administered to respondents in each panel of the NLSY.[2] However, while their analysis highlights the implications of the changes in skill distribution in terms of wages and employment, in this paper we focus instead on the endogeneity of skill accumulation—that is, schooling and work experiences—and document how the returns to these skills, as well as to cognitive ability, have changed across these cohorts.

Our analysis builds on the extensive literature that estimates the returns to schooling, beginning with the seminal work of Mincer (1974), who introduced what has become known as the Mincer model. This model interprets the coefficient on schooling in a log wage equation that controls for a quadratic in potential experience as a rate of return. Focusing on earnings, Heckman, Lochner, and Todd (2006) show that using flexible polynomials of schooling and potential work experience, as well as allowing for nonlinearities associated with degree completion (also known as "sheepskin effects"), is essential to accurately estimate the returns to schooling. An important contribution of our paper is to show that it is crucial to use actual, rather than potential, work experience when estimating the wage returns to the latter and that accounting for actual work experiences also affects estimated wage returns to schooling.

We deal with selection into schooling and work experiences by specifying and estimating a dynamic model of schooling and work decisions that controls for person-specific unobserved heterogeneity.[3] We follow Cameron and Heckman (1998, 2001) and Heckman, Stixrud, and Urzúa (2006), among

[2] A subset of the measures in the ASVAB are used to construct the Armed Forces Qualifying Test (AFQT) score, which is used by the US military in determining qualifications of young adults for military enlistment.

[3] See also recent work by Belzil and Hansen (2020), who, using data from the NLSY79 and NLSY97, estimate a dynamic model of schooling choices in which they control for dynamic selection on unobservables. Unlike Belzil and Hansen (2020), we account in our paper for dynamic selection into schooling as well as work experiences. Furthermore, while their paper focuses on the determinants of schooling outcomes, our main focus is on the wage returns to schooling and work experiences. Related research by Kejriwal, Li, and Totty (2020) also examines changes

others, and use a factor model to reduce the dimensionality of the unobserved state space.[4] We use initial background conditions, local college conditions, cognitive test scores, and the panel structure of the data to identify the heterogeneity factors. Noteworthy, unlike most of the literature on the wage returns to schooling and work experience, we separately account for work experience that is accumulated before or after graduation. Distinguishing between these two forms of work experience is important, since they may be rewarded differently upon postschooling labor market entry. Furthermore, failure to account for pregraduation work experience may bias estimates of the returns to schooling by incorrectly attributing to schooling the portion of the wage that in fact corresponds to in-school work experience.[5]

Our paper also contributes to the literature on understanding the effect of in-school work on future educational and labor market outcomes (Hotz et al. 2002; Bacolod and Hotz 2006; Scott-Clayton 2012; Baum and Ruhm 2016). Working while in school may cause students to take longer to complete schooling or drop out altogether. However, accumulating work experience during school also may have long-term benefits in the form of higher wages. Key to distinguishing between the costs and benefits of in-school work is accounting for the selection decisions of the individuals who participate. If, for example, high-ability students disproportionately obtain in-school work experience and are much more likely to graduate from high school and/or college, then failure to account for this type of selection will produce misleading policy conclusions about the labor market benefits of in-school work experience. We attempt to account for such selection in our econometric analyses.[6]

Using estimates of our dynamic model, we examine the selection-corrected returns to schooling and work experiences, as well as to unobservable cognitive and other, noncognitive skills, and how they changed across the cohorts we study. Our findings contribute to a small but growing empirical literature that has focused on decomposing the trends in the returns to education

over time in the return to schooling. Using data from the Survey of Income and Program Participation linked with administrative earnings data, they account for multidimensional unobserved heterogeneity using an interactive fixed effects framework. Their approach does not account for actual work experience.

[4] For other examples of factor models that have been used in the context of the returns to schooling, see, among others, Taber (2001), Hotz et al. (2002), Cunha, Karahan, and Soares (2011), and Heckman, Humphries, and Veramendi (2018).

[5] For example, Arcidiacono et al. (2016) find that pre- and postgraduation work experience is rewarded differently for college graduate workers.

[6] Hotz et al. (2002) also control for dynamic selection into in-school work when estimating the returns to in-school work experience. Unlike Hotz et al. (2002), our paper explicitly accounts for the fact that unobserved skills are multidimensional in nature.

(Taber 2001; Fang 2006; Fortin 2006; Lee and Wolpin 2010; Cunha, Karahan, and Soares 2011; Carneiro and Lee 2011).[7]

We find that failure to account for selection into various types of schooling and work experience results in sizable overstatements of the wage returns to degree attainment and a slight understatement of the wage returns to completed years of schooling. In addition, our selection-corrected estimates indicate that the return to an additional year of schooling is 3 percentage points higher among recent cohorts. With respect to degrees, we find that the return to a high school degree was slightly higher for the NLSY97, but we find no meaningful difference across cohorts in the returns to a bachelor's degree.

At the same time, controlling for selection has less of an effect on the returns to actual in-school and postschooling work experiences than on the returns to schooling. The selection-corrected estimated returns to working while in high school are negative for both cohorts, more so for more recent ones, while early cohorts had a 6% return to working while in college but none for more recent ones. With respect to postschooling work experiences, the selection-corrected return to part-time work is negative for both cohorts, more so for the earlier cohorts, while the return to full-time work ranges from 2% to 4%, with more recent cohorts having a higher one.

Finally, based on our selection-correction factor model, we find sizable returns to both cognitive and other, noncognitive skills, with the returns to cognitive skills being lower for more recent cohorts relative to earlier ones, while returns to other, noncognitive skills are considerably higher for the more recent cohorts than the earlier ones.

The remainder of the paper is organized as follows. Section II details the data we use and its construction, and section III presents descriptive statistics for the two cohorts we examine. In sections IV and V, we lay out the specification and estimation of our econometric model. In section VI, we present the results for our various models and their implied returns to schooling and work experiences, along with those for unobserved skills. Finally, section VII summarizes the paper and discusses some implications of our findings.

II. The Data

The data we use to determine the wages, education, and types of work experience across cohorts are derived from two panels of the NLSY, the NLSY79 and NLSY97. These surveys interview American youth beginning in their adolescent years and follow them through adulthood. They contain information on education, employment, background variables, and location

[7] While they focus on a different set of questions, Lee and Wolpin (2010) is particularly relevant for us as, to do so, the authors estimate a dynamic structural (equilibrium) model of schooling and work decisions in which they distinguish between six different types of sector-occupation–specific skills.

(county), among many others. The NLSY79 began in 1979 with a sample of respondents born in 1957–64, when they were aged 14–22. The respondents in the NLSY97 were born in 1980–84 and were first interviewed in 1997, when they were aged 12–17.

From these data, we make several sample selections. First, we restrict our analysis to male respondents.[8] Second, we restrict ourselves to the male respondents in the NLSY79 who were no more than age 20 in 1979 (i.e., were born in 1959–64), in order to minimize recall error at the first interview about their work and schooling experiences during adolescence (no such restrictions were imposed on the NLSY97, given that the oldest respondents were only 17 at the start of the latter survey). Third, we drop respondents in the military and in the economically disadvantaged white NLSY79 oversamples, since the former oversample was not followed after 1984 and the latter oversample was not followed after 1990. Finally, we drop respondents who were screened as "mixed race" in the NLSY97, since this was not an option in the NLSY79. After these restrictions, which are documented in detail in table A1 (tables A1, A2, B1, C1–C4 are available online), we end up with 3,862 male respondents from the NLSY79 and 4,559 from the NLSY97. In all of the analysis presented below, we split our data by these two NLSY surveys. One set of birth cohorts consists of NLSY79 respondents born in years 1959–64, while the other set consists of all birth cohorts in the NLSY97.

In both of the NLSY surveys, individuals are interviewed annually for the first 15 survey rounds and biennially thereafter. At each interview, respondents provide a history of what has transpired in their lives since the previous interview.[9] For example, the survey collects information on all jobs held between the current and previous interview, the wage and hours worked at each of those jobs, and the industry and occupation code of each job. Data related to educational attainment and schooling enrollment/attendance are similarly rich. Linking the survey reports together, it is possible to get measures of employment, schooling enrollment, military service, and hourly wages for those employed on a month-by-month basis. We track activities on a monthly basis so as to be able to distinguish between work experience that occurred during school as opposed to over the summer or between semesters

[8] We focus on men for two main reasons: (i) including women during early adulthood would require us to model their fertility decisions, which is outside the scope of the present analysis, and (ii) much of the literature that has studied human capital formation to which our analysis is comparable has focused on men.

[9] At the first interview, the survey asked extensive questions related to working and schooling history before the survey. Thereafter, for respondents who missed an interview, interviewers attempted to contact the individual during the following cycle and collect data on experiences between the current interview and the most recently completed interview.

as well as work experience that occurred before graduation as opposed to after graduation. In the analysis below, we focus on the activities of respondents in our two cohorts over the ages 16–35, covering the years 1975–99 for the NLSY79 and 1996–2016 for the NLSY97.[10]

With respect to initial conditions, young men in both cohorts are asked detailed questions in their first interview about their family situations. These family background characteristics (parental education, family income, and household structure) are assumed to affect labor market outcomes only through activity choices and, as such, serve as exclusion restrictions in our econometric model. In addition, the NLSY tracks the location of each individual in the surveys. Using the restricted-access Geocode supplement of the NLSY data, we are able to match individuals in the NLSY with county-level data from the Census Bureau, Bureau of Labor Statistics, and Bureau of Economic Analysis. This allows us to analyze the local labor market conditions that each individual faces over time. With additional data from the Integrated Postsecondary Education Data System, we create variables representing the higher-education landscape that these young men faced as teenagers (presence and number of 4-year colleges in the age 16 county of residence and tuition at state flagship university), which serve as further exclusion restrictions (see sec. V.C).

Our analysis is conducted on the following two samples: 3,852 men in the NLSY79 (854,179 person-month observations) and 4,443 men in the NLSY97 (792,652 person-month observations).[11] The additional sample cuts are due to attrition from the survey or missing interview spells of three or more years. A complete summary of sample selection criteria is included in table A1. In appendixes A and B (apps. A–D are available online), we provide full details about the creation of our analysis samples and the construction of the variables used in our analysis from the NLSY79 and NLSY97 data as well as the other data sources.

III. Cohort Differences in Background Characteristics, Skill Attainment, and Skill Wage Premia

In this section, we present some stylized facts across our two cohorts about differences in backgrounds, skill attainment, and wage premia to skills. We present these numbers at age 29—an age by which almost all individuals have completed their educational attainment.[12]

[10] See table A1 for the number of person-month observations for each of our birth cohorts and table B1 for the ages and years covered for each.

[11] Our wage analysis comprises 464,330 person-month observations in the NLSY79 and 422,114 person-month observations in the NLSY97.

[12] Keeping in mind that we are using monthly data, the numbers are calculated in the month before the respondents turn 29.

A. Personal and Family Background

We start by describing the differences across our two cohorts in personal and family background characteristics.

In the first panel of table 1, we show differences in race/ethnicity and nativity. There is no change in the percentage of African Americans, but we do see a very significant increase in the percentage of Hispanics across cohorts (from 7% to 14%). Interestingly, there is no significant change in the percentage of those who were born outside the United States.

In the next panel of table 1, we display differences in mothers' and fathers' education, family income, and status of who is the head of the household.[13] Between the NLSY79 and the NLSY97 cohorts, parental education increased by more than one grade level for mothers and more than four-fifths of a grade level for fathers. With respect to the more recent cohorts, they grew up in households with higher family income ($33,580 vs. $32,860), although this difference is not statistically significant. Finally, the share of young men in our samples who grew up in female-headed households increased by 11 percentage points between the NLSY79 and NLSY97.[14]

There also are differences across the two cohorts in measures of cognitive skills. We focus here on differences in scores on the Armed Forces Qualification Test (AFQT).[15] The third panel of table 1 displays the median and standard deviation of AFQT scores for the two cohorts as well as cross-cohort differences. AFQT scores for the NLSY97 are, in general, higher and more dispersed than those for the NLSY79, with an overall large (but not statistically significant) increase of 0.07 standard deviations in the median score as well as a small (but statistically significant) increase in the standard deviation itself. These results are consistent with the findings of Altonji, Bharadwaj, and Lange (2012), who document a widening of the AFQT distribution between the NLSY79 and NLSY97 cohorts.

[13] These are the family background variables that make up some of our model's exclusion restrictions.

[14] In table C2, we show cross-cohort differences in local labor markets and local college characteristics. The local college characteristics account for the remainder of our exclusion restrictions.

[15] The AFQT is a subset of the ASVAB. Specifically, AFQT scores are a weighted average of four ASVAB subtests: arithmetic reasoning, mathematics knowledge, paragraph comprehension, and word knowledge. In our model, we make use of six ASVAB subtests, the four in the AFQT as well as coding speed and numerical operations. To make both the AFQT and the ASVAB scores comparable across cohorts, we follow Altonji, Bharadwaj, and Lange (2009, 2012) by making use of an equipercentile mapping in ASVAB test scores that corrects for both testing medium (i.e., pencil and paper vs. computer assisted) and age at test (NLSY97 respondents were much younger than NLSY79 respondents when they took the ASVAB).

Table 1
Demographic, Family, and Armed Forces Qualification Test (AFQT) Characteristics

Variable	NLSY79	NLSY97	NLSY97 − NLSY79
Demographics:			
White	.79	.71	−.08***
Black	.15	.16	.01
Hispanic	.07	.14	.07***
Foreign-born	.04	.05	.01
Family characteristics:			
Mother's education	11.76	12.91	1.14***
Father's education	12.17	12.98	.81***
Family income	32.86	33.58	.71
Share lived in female-headed household	.12	.23	.11***
AFQT:			
Median of AFQT score	.37	.44	.07
Standard deviation of AFQT score	.96	.97	.01***
N at age 29	3,464	3,569	

Note.—Education is highest grade completed by the respondent's biological parents. Family income is in thousands of 1982–84 dollars. All demographic and family variables are measured as of the first survey round in both cohorts except female-headed household, which is from age 14 in NLSY97. The AFQT distribution is normalized so that the distribution including all cohorts is mean 0 and variance 1. For median AFQT score, the significance comes from bootstrapped standard errors of the median (500 replications). For standard deviations of AFQT score, the significance comes from two-tailed *F*-tests of the ratio of the variances. Statistics are weighted by National Longitudinal Survey of Youth (NLSY) sampling weights. Sample size for statistical analysis varied for some variables because of missing values (see table A1 for more on sample creation).

*** Significant at the 1% level.

B. Educational Attainment and Work Experiences

We now consider differences across the two cohorts in months of accumulated schooling and work experiences and in educational degree attainment.

Table 2 describes schooling attainment and college completion at age 29 for both cohorts. In the first panel, there is a clear increase across cohorts in educational attainment. While there is little change in the high school dropout rate across cohorts, there is a 3 percentage point increase in those who complete some college and a 4 percentage point increase in those who receive a bachelor's degree. For comparison purposes, in table C1 we report educational attainment from identically aged men in the Current Population Survey (CPS); the CPS shows cross-cohort changes similar to those in the NLSY.[16]

In the second panel of table 2, we find an increase in the number of young men starting college, although there is not a significant change in the college graduation rate among those who start (although there is a nominal increase). Furthermore, we see a significant increase of two-fifths of a year

[16] For a more complete comparison of educational wage premia in the NLSY, CPS, and other major US household surveys, see Ashworth and Ransom (2019).

Table 2
Schooling Attainment and Graduation Probabilities at Age 29

Variable	NLSY79	NLSY97	NLSY97 − NLSY79
Schooling attainment:			
% high school dropouts	.11	.09	−.01**
% high school graduates	.29	.25	−.05***
% some college	.38	.40	.03**
% college graduates	.22	.26	.04***
Graduation probabilities and time to degree:			
Pr(start college)	.60	.66	.06***
Pr(graduate college | start college)	.37	.39	.02
Time to college degree (years)	5.08	5.49	.41***

NOTE.—High school graduates included in this table are those who have either a GED or a diploma but who never attended college. Those with some college attended college but did not graduate with a 4-year degree. College graduates are those who graduated with a 4-year degree. As in Bound, Lovenheim, and Turner (2012), time to college degree is defined as the number of calendar months between high school graduation and 4-year college graduation. Statistics utilize National Longitudinal Survey of Youth (NLSY) sampling weights.
** Significant at the 5% level.
*** Significant at the 1% level.

in the time to a college degree, which is consistent with Bound, Lovenheim, and Turner (2012).

We also examine differences across the two cohorts in months of accumulated schooling and work experience. Table 3 reports average levels of schooling and work experience (in months) by age 29 (beginning at age 16). Consistent with table 2 and Bound, Lovenheim, and Turner (2012), we find that students in the NLSY97 spent longer in school by almost a full year. Despite this, those in the NLSY97 also accumulated slightly more total work experience by age 29 as those in the NLSY79 (almost 2 months more). That

Table 3
Changes in School and Work Experience

Variable	NLSY79	NLSY97	NLSY97 − NLSY79
Overall:			
Total months of schooling	40.53	52.24	11.71***
Total months of work experience	116.85	118.61	1.76**
By type:			
Months of school only	19.52	20.64	1.12**
Months of work in high school	9.41	11.84	2.43***
Months of work in college	11.60	19.76	8.16***
Months of part-time work	13.39	14.30	.91***
Months of full-time work	82.45	72.71	−9.74***

NOTE.—Monthly activities are displayed in table 5 and fully described in app. A. Months of an activity as of age a is the sum of incidence in that activity from age 16 to current age, a. Thus, the average individual in the NLSY79 had a total of 40.5 months of school after turning 16. Statistics are weighted by National Longitudinal Survey of Youth (NLSY) sampling weights.
** Significant at the 5% level.
*** Significant at the 1% level.

said, there were differences in the types of work experience the two cohorts accumulated by this age. In particular, there was an increase across cohorts in the accumulated level of in-high-school work experience (about 2.5 months) and a much larger increase in in-college work experience (more than 8 months). Furthermore, while the overall level of out-of-school part-time work was basically the same, the overall level of out-of-school full-time work sharply declined (by more than 9 months).

These differences across cohorts in the types of accumulated work experiences that young men experienced motivate our differential treatment of in-school and out-of-school work experience.

C. Wage Premia

Finally, we examine how wage premia have varied across our two cohorts by documenting how the association between wages at age 29 and amounts of schooling or work experience has changed across cohorts. Herein, we refer to differences in wages across school and work experience levels as "wage premia," although we hasten to add that these measures are not to be interpreted as causal effects. Below, in sections IV and V, we develop a model to estimate the causal effects of schooling and work experience on wages.

The first panel of table 4 reports the wage premia associated with various experiences for those working full time at age 29. Each row shows the mean change in the full-time log wage with an additional year of each type of experience. The wage premia are highest for working in college, in the range of 7% to 9%.[17] On the other hand, out-of-school part-time work experience is associated with lower wages, in the range of −9% to −13% for an additional year of experience.[18] For full-time work experience, the wage premia are small and not statistically different from zero. For each of the work experience wage premia, we see a decrease across cohorts. The model we present in the next section will shed light on whether these patterns in wage premia similarly hold for wage returns.

The other panels of table 4 allow us to assess how the observed wage premia associated with educational attainment have changed across these cohorts. The second panel shows average log wages associated with the four different educational attainments described in table 2 and reveals a decrease in inflation-adjusted wage levels across cohorts of between 3 and 10 log points for each education level. The third panel shows the wage premia for each

[17] To further investigate whether the timing of in-college work experience matters, we separate in-college work experience into two types: experience attained in the freshman and sophomore year and experience attained in later years of college. Table 4 shows that earlier in-college work experience has a larger premium in the NLSY79 but that the two have similar premia in the NLSY97.

[18] As we will show later in the paper, this negative association partly reflects negative selection into part-time work.

Table 4
Changes in Wage Premia for Experience and Educational Attainment at Age 29 for Full-Time Workers

Variable	NLSY79	NLSY97	NLSY97 − NLSY79
Average log wage premia for one more year of experience:			
Year of school	.066	.056	−.010*
Work in high school	.055	.031	−.024
Work in college	.089	.067	−.022***
Early college work	.137	.065	−.072**
Late college work	.073	.068	−.005
Work part time	−.091	−.128	−.037***
Work full time	.005	.002	−.003
Average log wages by highest educational attainment:			
High school dropouts	1.81	1.75	−.05
High school graduates	1.95	1.92	−.03
Some college	2.09	1.99	−.10***
College graduates	2.35	2.30	−.05*
Average log wage premia for highest educational attainment:			
High school wage premium	.14	.16	.02
Some college wage premium	.14	.07	−.07***
College wage premium	.41	.38	−.03***

NOTE.—The sample is conditional on working full time. Estimates for work experience are coefficients from separate bivariate regressions of log wage on each cumulative experience term. The exception is for the breakout of work in college. Early college work refers to work done as a freshman or sophomore (no more than 16 months of college, i.e., four semesters), while late college work refers to work done as a junior or senior (more than 16 months of college). The premia for these two experiences are from a joint regression. High school graduates included in this table are those who never attended college. Those with some college attended college but did not graduate with a 4-year degree. College graduates are those who graduated with a 4-year degree. High school wage premium refers to the log wage difference between high school graduates and high school dropouts. Some college wage premium refers to the log wage difference between those with some college and high school graduates. College wage premium refers to the log wage difference between college graduates and high school graduates. Statistics are weighted by National Longitudinal Survey of Youth (NLSY) sampling weights.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

degree. Most notable is the significant decrease in the college wage premium, which is 3 log points lower in the more recent cohort.[19]

As noted above, our discussion thus far has ignored the possibility that selective differences in educational attainment and accumulated work experiences

[19] While our finding of a decreasing college wage premium between the NLSY79 and the NLSY97 is at odds with some previous research (Castex and Dechter 2014; Deming 2017; Böhm 2020), it is consistent with some recent studies of changes in wages over time and is robust to a number of different specifications. Ashworth and Ransom (2019) perform a full comparison of the college wage premium using five different US surveys and find that, compared with other US surveys, the NLSY shows a much lower college wage premium for the NLSY97 cohorts and a much lower advanced degree premium for the NLSY79 cohorts born in 1960–64.

may affect the suggested impacts of the latter on wages among young men and how they changed across cohorts. In the next section, we introduce the model that we use to account for selection into the various types of experience and, in our final results, present and discuss selection-corrected wage returns. The differences we have documented in schooling and work experiences, as well as in personal and family background characteristics, are the prime motivation for our model in which we estimate the evolution of wage returns to skills by accounting for these changes in composition.

IV. The Model

In this section, we develop a dynamic model of schooling and work decisions. We use it to form an econometric model that accounts for the endogeneity of accumulated schooling and work experiences in the estimation of wage returns across our two cohorts.

A. Activity Choices

We assume that, at each age a—which is measured in months in our case—individual i, who is a member of birth cohort c, chooses activity j from a set of possible activities, which may vary with age and/or the occurrence(s) of one or more previous events. For simplicity, we suppress notation indexing the individual's cohort. We estimate the model separately for both the NLSY79 and NLSY97 cohorts, so all the parameters should be understood as cohort specific. Let R_{ia} denote the choice set for individual i at age a, where we assume that there are K possible choice sets (i.e., $R_{ia} = r \in 1, \ldots, K$). Then, conditional on facing choice set $R_{ia} = r$, individual i chooses from among J^r activities, where we define

$$d^r_{iaj} = \begin{cases} 1 & \text{if } i \text{ chooses activity } j \text{ from choice set } r \text{ at age } a, \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

and $\Sigma^{J^r}_{j=1} d^r_{iaj} = 1$, for all i, a, and r. In practice, we consider $K = 3$ choice sets, which are composed of the potential activities for those who (i) have not graduated from high school ($R_{ia} = 1$), (ii) have graduated from high school but have not graduated from college ($R_{ia} = 2$), and (iii) have graduated from college ($R_{ia} = 3$). The three choice sets and the activities associated with each are given in table 5, and the definitions of these activities can be found in appendix A.

B. School and Work Experiences

We are interested in estimating the effects of accumulated experiences on various outcomes. In particular, we are interested in accumulated years of school attendance as well as years of work experiences. We also use our

Table 5
Definitions of Activities by Educational Choice Sets

Activity (j^r)	Description
$R_{ia} = 1$ (pre-high-school graduate):	
1	School only, no high school diploma or GED
2	Work in school, no high school diploma or GED
3	Work part time (no school), no high school diploma or GED
4	Work full time (no school), no high school diploma or GED
5	Military, no high school diploma or GED
6	Other, no high school diploma or GED
7	Graduate from high school at age a (attainment activity)
$R_{ia} = 2$ (high school graduate):	
1	School only, has high school diploma or GED
2	Work in school, has high school diploma or GED
3	Work part time (no school), has high school diploma or GED
4	Work full time (no school), has high school diploma or GED
5	Military, has high school diploma or GED
6	Other, has high school diploma or GED
7	Graduate with bachelor's degree at age a (attainment activity)
$R_{ia} = 3$ (college graduate):	
1	School only, has bachelor's degree
2	Work in school, has bachelor's degree
3	Work part time (no school), has bachelor's degree
4	Work full time (no school), has bachelor's degree
5	Military, has bachelor's degree
6	Other, has bachelor's degree

NOTE.—The creation of our school and work activity variables is fully described in app. A.

model to estimate the effect of educational attainment, such as high school and college graduation, on these outcomes. In the following, we will refer to these work experiences, schooling activities, and graduation outcomes collectively as "experiences."

The vector of types of experience is given by

$$\mathbf{x}^r_{ia} \equiv (x_{1ia}, \mathbf{x}^r_{2ia}, x_{3ia}, x_{4ia}, x_{5ia}, x_{6ia}, I_{ia}(R_{ia} > 1), I_{ia}(R_{ia} = 3))', \quad (2)$$

where the experience variables are as follows: x_{1ia}, the number of years of schooling attendance as of age a; $\mathbf{x}^r_{2ia}$, the number of years of in-school work experience (given the relevant choice set r); x_{3ia}, the total number of years of part-time (nonschool) work as of age a; x_{4ia}, the total number of years of full-time (nonschool) work as of age a; x_{5ia}, the number of years in the military as of age a; x_{6ia}, the number of years spent in other activities as of age a;[20]

[20] This residual category includes home production as well as unemployment.

$I_{ia}(R_{ia} > 1)$, an indicator equal to 1 if individual i has received a high school degree as of age a; and $I_{ia}(R_{ia} = 3)$, an indicator equal to 1 if individual i has received a bachelor's degree as of age a.

The experience variables $\mathbf{x}^r_{ia}$ are accumulated from the starting age, $a_0 = 192$ in months (i.e., age 16), to age $a - 1$.[21] The first element of $\mathbf{x}^r_{ia}$, accumulated years of attendance in school-related activities, is defined as

$$x_{1ia} = \frac{1}{12}\sum_{\ell=a_0}^{a-1}(d_{i\ell 1} + d_{i\ell 2}), \tag{3}$$

which corresponds to the years-of-schooling variable used in the wage returns literature originating with Mincer (1974).[22] The in-school work experience vectors, $\mathbf{x}^r_{jia}, j = 2$, are defined as follows. In-school work experience before high school graduation is a scalar equal to the number of years spent working in high school since a_0, $x_{2,\mathrm{HS},ia}$. For individuals who graduate from high school, the vector for in-school work while in college (or graduate school) contains two elements: the number of years working while in high school, $x_{2,\mathrm{HS},ia}$, and the number of years spent working while in college or graduate school, $x_{2,\mathrm{COL},ia}$. That is,

$$\mathbf{x}^r_{2ia} \equiv \begin{cases} x_{2,\mathrm{HS},ia} & \text{if } R_{ia} = 1, \\ (x_{2,\mathrm{HS},ia}, x_{2,\mathrm{COL},ia}) & \text{if } R_{ia} > 1, \end{cases} \tag{4}$$

where

$$x_{2,\mathrm{HS},ia} = \frac{1}{12}\sum_{\ell=a_0}^{a-1} d_{i\ell 2},$$

$$x_{2,\mathrm{COL},ia} = \frac{1}{12}\sum_{\ell=a_{\mathrm{HS}_i}}^{a-1} d_{i\ell 2} \quad \text{if } R_{ia} > 1,$$

and where a_{HS_i} is the age of graduation from high school. Finally, the remaining experience variables in $\mathbf{x}^r_{ia}$ are defined as

$$x_{jia} = \frac{1}{12}\sum_{\ell=a_0}^{a-1} d_{i\ell j}, j = 3, \ldots, 6. \tag{5}$$

C. Wages

Let W_{iaj} denote the potential hourly wage rate that individual i would realize at age a if he were to choose activity $j, j = 2, 3, 4$. We assume that W_{iaj} is determined by the individual's accumulated human capital, or skills, H_{ia},

[21] For now, we suppress the r superscript from the activity indicators $d^r_{i\ell j}$.

[22] In the specification of the activity-specific value functions below, we define x_{1ia} slightly differently, as only the sum of $d_{i\ell 1}$.

as of the beginning of age a, measured in efficiency units; the occupation-specific skill price P_{iaj} per efficiency unit that varies across the local labor market in which i resides at age a;[23] and idiosyncratic shocks, denoted by $e^{\varepsilon_{iaj}}$, that are unanticipated by the individual,

$$W_{iaj} = P_{iaj} H_{ia} e^{\varepsilon_{iaj}}, \tag{6}$$

so that the log of wages, denoted by w_{iaj}, is given by the following linear function:

$$\begin{aligned} w_{iaj} &= p_{iaj} + h_{ia} + \varepsilon_{iaj} \\ &= w^e_{iaj} + \varepsilon_{iaj}, \end{aligned} \tag{7}$$

where $p_{iaj} \equiv \ln P_{iaj}$, $h_{ia} \equiv \ln H_{ia}$, and $w^e_{iaj} \equiv p_{iaj} + h_{ia}$ is i's expected log wage at age a (i.e., the wage that i expects to get if he chooses activity j). We assume that p_{iaj} is the following function of the conditions of the local labor market in which i resides at age a, $\mathbf{m}_{ia}$:

$$p_{iaj} = \beta_{0j} + \boldsymbol{\beta}_{\mathbf{m}} \mathbf{m}_{ia}. \tag{8}$$

We further assume that the (log of the) individual's stock of human capital, h_{ia}, is determined by some observed personal characteristics, for example, one's birth year, race, and so on, denoted by the vector $\mathbf{z}_i$; the individual's accumulated schooling and work experience and degree completion, $\mathbf{x}^r_{ia}$; and the individual's unobserved characteristics, $\boldsymbol{\xi}_i$, which are broken out into elements pertaining to the individual's cognitive (ξ_{1i}) and other (non-cognitive) abilities (ξ_{2i}):

$$h_{ia} = \boldsymbol{\beta}_{\mathbf{z}} \mathbf{z}_i + \boldsymbol{\beta}_{\mathbf{x}} g(\mathbf{x}^r_{ia}) + \beta_{\xi 1j} \xi_{1i} + \beta_{\xi 2j} \xi_{2i}. \tag{9}$$

It follows that

$$\begin{aligned} w_{iaj} &= w^e_{iaj} + \varepsilon_{iaj} \\ &= \beta_{0j} + \boldsymbol{\beta}_{\mathbf{m}} \mathbf{m}_{ia} + \boldsymbol{\beta}_{\mathbf{z}} \mathbf{z}_i + \boldsymbol{\beta}_{\mathbf{x}} g(\mathbf{x}^r_{ia}) + \beta_{\xi 1j} \xi_{1i} + \beta_{\xi 2j} \xi_{2i} + \varepsilon_{iaj}, \end{aligned} \tag{10}$$

where $g(\cdot)$ contains (i) a cubic polynomial in all types of accumulated experience,[24] (ii) pairwise interactions between school experience and each of the work experience variables (work in school, part-time work, and full-time work), and (iii) indicators for having graduated high school and for having graduated college (see also Heckman, Lochner, and Todd 2006).

[23] See Moretti (2011) for a survey of models of local labor markets.

[24] See also Belzil and Hansen (2002), who estimate the returns to schooling using an extended Mincerian specification in which they relax the assumption that wages are linear in the number of years of schooling.

One of our primary interests is obtaining consistent estimates of the parameters in equation (10). As we make clear below, the central obstacle is that the elements of $\mathbf{x}^r_{ia}$ are endogenous unless one conditions on the unobserved factors, $\boldsymbol{\xi}_i$. We now develop the nature of that linkage through the sequences of activity choices individual i makes over his life cycle.

D. Activity-Specific Value Functions

Let the value function for individual i who is of age a and who engages in activity j (from choice set r) be denoted by V^r_{iaj}. These value functions depend on the elements of the individual's information set at age a, namely, personal characteristics, $\mathbf{z}_i$; family background characteristics, $\mathbf{f}_i$; local college characteristics at age 16, $\mathbf{c}_{i,16}$; local labor market characteristics at age a, $\mathbf{m}_{ia}$; accumulated school and work experiences at that age, $\mathbf{x}^r_{ia}$; and the individual's unobserved characteristics, $\boldsymbol{\xi}_i$.[25] For computational simplicity, we approximate the V^r_{iaj}'s as a sum of a linear function of these characteristics and interactions between $\mathbf{x}^r_{ia}$ and $\mathbf{z}_i$:

$$\begin{aligned} V^r_{iaj} &= \boldsymbol{\alpha}^r_{zj}\mathbf{z}_i + \boldsymbol{\alpha}^r_{fj}\mathbf{f}_i + \boldsymbol{\alpha}^r_{cj}\mathbf{c}_{i,16} + \boldsymbol{\alpha}^r_{mj}\mathbf{m}_{ia} + \boldsymbol{\alpha}^r_{xj}b(\mathbf{x}^r_{ia}, \mathbf{z}_i) \\ &\quad + \alpha^r_{\xi 1j}\xi_{1i} + \alpha^r_{\xi 2j}\xi_{2i} + \omega_{iaj} \qquad (11) \\ &= v^r_{iaj} + \omega_{iaj}, \end{aligned}$$

where $b(\cdot)$ contains (i) a set of up to nine bin indicators for each type of accumulated experience and (ii) linear interactions between race/ethnicity and each type of accumulated experience.[26] Finally, ω_{iaj} captures the idiosyncratic factors that affect the individual's value from choosing activity j at age a.

It follows that at each age a, individual i chooses the activity j^{r*}_{ia} from among the activities in the current choice set that yields the highest value:

$$j^{r*}_{ia} = \underset{j}{\operatorname{argmax}}\, V^r_{iaj}. \qquad (12)$$

E. Unobserved Skills

Our model incorporates two unobserved random factors representing the unobserved cognitive and other, noncognitive abilities of individuals.

[25] See table A2 for a detailed description of these elements.

[26] As an example of the bin indicators, we include a set of nine bins for the number of months of full-time work experience outside school. The cut points for each of the bins occur at the following values: 12, 24, 36, 48, 60, 72, 84, and 96 months. While the choice of cut points for each experience is different, the cut points are constant across NLSY cohorts. Allowing the different types of experience to vary in this way allows us to estimate highly nonlinear effects of experience on the decision to invest in different types of human capital. This nonlinear relationship is necessary in order to match the observed data. All experience terms have nine bins except for military, which has five.

To measure unobserved cognitive ability (ξ_{1i}), we use six subject tests from the ASVAB.[27] We chose to include these subjects because (i) each appears in both the NLSY79 and the NLSY97 and (ii) they are measure constructs typically thought to be associated with individuals' cognitive ability or skills. For each subject test s, the z-scored test score y for individual i is expressed as a linear function of personal characteristics $\mathbf{z}_i$ and the cognitive ability ξ_{1i}, namely,

$$y_{is} = \gamma_{0s} + \gamma_{zs}\mathbf{z}_i + \gamma_{\xi 1s}\xi_{1i} + \eta_{is}, \tag{13}$$

where η_{is} captures idiosyncratic variation in test scores not related to the cognitive ability or other test score determinants.[28]

There is little overlap in the measures of noncognitive traits across the two NLSY surveys.[29] Because of this data limitation, we are unable to include comparable measures of noncognitive ability for both of our NLSY cohorts. For this reason, we rely on the panel nature of the data—along with exclusion restrictions to be discussed in the next section—to identify the residual ability factor ξ_{2i}. Thus, this second ability factor can be interpreted as noncognitive in the sense that it captures all unobserved permanent person-specific determinants of the agent's wage and decision process that are orthogonal to the cognitive factor.

V. Inference

In this section, we further characterize our econometric model and the strategy for estimating its parameters. In particular, we summarize the specification of the error structure of our model and the estimation procedure we employ. For now, we continue to not notationally distinguish between the NLSY79 and NLSY97, although we allow all of the parameters of our model to be cohort specific, and we explicitly examine the cross-cohort differences in the estimated marginal returns to schooling and work experiences. Finally, we also discuss the identification of the model.

[27] The six subject tests we use are: arithmetic reasoning, coding speed, mathematics knowledge, numerical operations, paragraph comprehension, and word knowledge. The frequently used AFQT score is a composite of all of these subjects except for coding speed and mathematics knowledge. Our six subject tests are the same as used by Heckman, Humphries, and Veramendi (2018).

[28] The mean and standard deviation used to compute the z-scores are taken across both cohorts.

[29] The NLSY79 contains the Rotter locus of control score and Rosenberg self-esteem scale for all individuals. These have been used in other studies as noncognitive measures (Heckman, Stixrud, and Urzúa 2006; Cunha, Karahan, and Soares 2011). The NLSY97 does not collect information on any of these tests but instead collects information on risky behavior, such as school suspensions, sexual promiscuity, and substance abuse. See, e.g., Aucejo and James (2019), who use school suspensions, fights, precocious sex, grade retention, substance abuse, and eighth-grade GPA as noncognitive measures.

A. Error Structure

We assume that $\boldsymbol{\xi}_i$ is a person-specific vector of factors that is stochastically independent of the distributions of the observables, $\mathbf{z}_i$, $\mathbf{f}_i$, $\mathbf{c}_{i,16}$, and $\mathbf{m}_{ia}$, and of the unobservables, $\boldsymbol{\omega}_{ia}$, $\boldsymbol{\varepsilon}_{ia}$, and $\boldsymbol{\eta}_i$, for all a and i.[30] At the same time, because the choice of past activities determines the accumulated experience in $\mathbf{x}^r_{ia}$, it is not the case that the elements of this vector are independent of $\boldsymbol{\xi}_i$.

We further normalize, for both cohorts, the distribution of the unobserved factors $\boldsymbol{\xi}_i$ to be normally distributed with mean 0 and identity covariance matrix. With respect to $\boldsymbol{\omega}_{ia}$, $\boldsymbol{\varepsilon}_{ia}$, and $\boldsymbol{\eta}_i$, respectively, we assume that they are mutually independent, are independently distributed both across ages and at each age a, and have mean 0 and constant variances.[31] That the vector of activity shocks $\boldsymbol{\omega}_{ia}$ is uncorrelated with $\boldsymbol{\varepsilon}_{ia}$ is the result of assuming that decisions about activities are made at each age a before the actual realizations of wages are known by individual i.

B. Likelihood Function and Estimation Method

We assume that the idiosyncratic errors in the activity payoff functions, ω_{iaj}, have a type I extreme value distribution so that the choice probability for any individual i at age a to choose activity j in choice set r, conditional on the unobserved factors $\boldsymbol{\xi}_i$, has the logistic form

$$P^r_{iaj} = \frac{\exp\left(v^r_{iaj}\right)}{\sum_{k=1,\dots,J^r} \exp\left(v^r_{iak}\right)}, \tag{14}$$

where, as defined in the first line of equation (11), v^r_{iaj} is the component of the value function associated with activity ℓ that is deterministic from individual i's viewpoint. Recall that v^r_{iaj} depends on the unobserved factors $\boldsymbol{\xi}_i$ and on personal characteristics $\mathbf{z}_i$, family background characteristics $\mathbf{f}_i$, local college characteristics at age 16 $\mathbf{c}_{i,16}$, local labor market characteristics $\mathbf{m}_{ia}$ as of age a, and accumulated school and work experiences $\mathbf{x}^r_{ia}$ as of that age:

$$v^r_{iaj} = \boldsymbol{\alpha}^r_{\mathrm{z}j}\mathbf{z}_i + \boldsymbol{\alpha}^r_{\mathrm{f}j}\mathbf{f}_i + \boldsymbol{\alpha}^r_{\mathrm{c}j}\mathbf{c}_{i,16} + \boldsymbol{\alpha}^r_{\mathrm{m}j}\mathbf{m}_{ia} + \boldsymbol{\alpha}^r_{\mathrm{x}j}b(\mathbf{x}^r_{ia}, \mathbf{z}_i) + \alpha^r_{\xi 1j}\xi_{1i} + \alpha^r_{\xi 2j}\xi_{2i}. \tag{15}$$

Additionally, we assume that the idiosyncratic errors entering the wage function in equation (10) are normally distributed with mean 0 and variance $\sigma^2_{w_j}$. Thus, the corresponding contribution to the likelihood, conditional on $\boldsymbol{\xi}_i = \boldsymbol{\xi}$, is given by

[30] The assumption that individual effects $\boldsymbol{\xi}_i$ are independent of the observable characteristics and of the idiosyncratic shocks is very common in dynamic discrete choice models. See, among others, Taber (2001), Belzil and Hansen (2002), Hotz et al. (2002), and Heckman, Stixrud, and Urzúa (2006).

[31] In practice, some of these shocks may exhibit some degree of persistence over time. In our model, this feature would be at least partly accounted for by the time-invariant unobserved factor ξ_{2i}, which would then be interpreted as a mixture of noncognitive skills and the persistent component of the shocks.

$$\ell_{w_{iaj}} = \frac{1}{\sigma_{w_j}} \phi\left(\frac{w_{iaj} - \beta_{0j} - \boldsymbol{\beta}_{\mathrm{m}}\mathbf{m}_{ia} - \boldsymbol{\beta}_{\mathrm{z}}\mathbf{z}_i - \boldsymbol{\beta}_{\mathrm{x}} g(\mathbf{x}_{ia}^r) - \beta_{\xi j1}\xi_1 - \beta_{\xi j2}\xi_2}{\sigma_{w_j}}\right), \quad j = 2, 3, 4,$$

where $\phi(\cdot)$ is the standard normal probability density function (pdf).[32]

We also assume that the idiosyncratic errors entering the ASVAB test score function in equation (13) are normally distributed with mean 0 and variance $\sigma_{y_s}^2$. Thus, the likelihood contribution, conditional on $\xi_{i1} = \xi_1$, is given by

$$\ell_{y_{is}} = \frac{1}{\sigma_{y_s}} \phi\left(\frac{y_{is} - \gamma_{0s} - \boldsymbol{\gamma}_{\mathrm{z}s}\mathbf{z}_i - \gamma_{\xi s1}\xi_1}{\sigma_{y_s}}\right). \tag{16}$$

It follows that the (unconditional) log likelihood function is given by

$$\log \mathcal{L}(\boldsymbol{\theta}) = \sum_i \log \int \mathcal{L}_i(\theta|\boldsymbol{\xi}) f_{\xi}(\boldsymbol{\xi})\, d\boldsymbol{\xi}, \tag{17}$$

where, conditional on $\boldsymbol{\xi}_i = \boldsymbol{\xi}$, the individual contribution to the likelihood is given by

$$\mathcal{L}_i(\theta|\boldsymbol{\xi}) = \prod_s \ell_{y_{is}} \prod_a \prod_r \left[\prod_{j=1,5,6,7} (P_{iaj}^r)^{d_{iaj}^r} \prod_{k=2,3,4} [P_{iak}^r \ell_{w_{iak}}]^{d_{iak}^r} \right]^{I(R_{ia}=r)}, \tag{18}$$

where $\boldsymbol{\theta} \equiv (\boldsymbol{\alpha}', \boldsymbol{\beta}', \boldsymbol{\gamma}')'$, $I(A)$ is the indicator function that is equal to 1 if A is true and 0 otherwise, and $f_{\xi}(\cdot)$ is the pdf of $\boldsymbol{\xi}$. In the analysis that follows, we employ the assumption that $\boldsymbol{\xi}$ is distributed multivariate normal and estimate the model via maximum likelihood.[33]

C. Identification

In this section, we discuss the identification of key features of the model. Note that we cannot readily identify the effects of endogenously determined schooling and work experiences on wages or subsequent school and work decisions by relying on standard instrumental variables techniques, as finding valid instruments for these sequences of past choices over individuals'

[32] Recall that choice set–specific intercepts are included in $\mathbf{x}_{ia}^r$ through degree attainment dummies.

[33] In practice, we use quadrature to approximate the integral of the likelihood function. Specifically, we use Gaussian quadrature with seven points of support for each dimension of the integral. As starting values for the parameters, we use perturbed point estimates from the specification of the model without unobserved heterogeneity. Finally, standard errors are computed using the estimated cluster-robust asymptotic covariance matrix, which accounts for within-person serial correlation of the error terms.

careers is very challenging, if not impossible.[34] Herein, we deal with dynamic selection into schooling and work experiences by explicitly modeling the underlying choice process, controlling for person-specific unobserved factors as in Cameron and Heckman (1998, 2001) and Heckman, Stixrud, and Urzúa (2006). In what follows, we discuss how identification is achieved within this econometric framework.

First, one can use the results of Hu and Shum (2012) to show nonparametric identification of the conditional choice probabilities, P^r_{iaj}. This identification result relies on the first-order Markov structure and the resulting dynamic exclusion restrictions implied by our dynamic discrete choice model.[35] Under the assumption that the idiosyncratic preference shocks are distributed following a type I extreme value assumption, the conditional value functions are then identified (up to a reference alternative) by inverting the conditional choice probabilities, P^r_{iaj}.

We now turn to the unobserved individual factors, (ξ_1, ξ_2), and the outcome equations. Aside from the aforementioned dynamic exclusion restrictions, we also impose two types of exclusion restrictions that play an important role in identifying the covariate effects in the outcome equations as well as the distribution and the returns to these unobserved factors (i.e., the factor-loading parameters). First, we impose the restriction that the noncognitive factor, ξ_2, does not enter the ASVAB test score equations. This results in a system of six continuous and selection-free measurements that are dedicated to the first factor ξ_1. From this set of measurements, the factor loadings associated with ξ_1 are identified from the covariances of the ASVAB test scores. Having identified the factor loadings, the distributions of ξ_1 and of the idiosyncratic performance shocks are identified in a second step using deconvolution arguments (Kotlarski 1967).

Note, however, that we cannot directly use the same arguments for the second unobserved factor ξ_2, as we do not have access to a set of selection-free continuous measurements dedicated to that factor. In our model, the continuous outcomes (wages) along with the discrete choices of activities play the

[34] A number of papers in the returns to schooling literature follow Card (1995) and use presence of a college (or geographical distance to college) in the local labor market at age 14 as an instrument for college attendance (see, among others, Kane and Rouse 1995; Kling 2001; Currie and Moretti 2003). Kane and Rouse (1995) also use tuition at local public 4-year colleges at age 17. See Card (2001) for a survey. Unlike these papers, our goal is to estimate the wage returns to schooling, along with the different types of work experiences. As such, our approach does not lend itself to a standard instrumental variables strategy. Importantly, though, we build on this literature and use density of local colleges as well as flagship tuition as exclusion restrictions in our model.

[35] In our model, choices and outcomes today depend only on the past sequence of choices through the accumulated experiences at the beginning of the period, once we condition on unobserved heterogeneity.

role of noisy measurements of the underlying factors. Two main aspects of the data and the model are then central to the identification argument. First, the panel dimension of the data—in particular, the autocorrelation of wages and choices (conditional on observed covariates)—along with the correlation between these two sets of variables and the ASVAB measurements play an important role in identifying the returns to unobserved factors (ξ_1, ξ_2) in the outcome and choice equations. Second, as we discuss below, exclusion restrictions in the form of variables affecting individual decisions but excluded from the potential wages are key to addressing the underlying selection issue. Having identified the distribution of ξ_1 in the previous step, these exclusion restrictions make it possible in turn to identify the distribution of the unobserved factor ξ_2 using standard deconvolution arguments applied to the distribution of potential wages.

In practice, we exclude the vectors of family background characteristics, $\mathbf{f}_i$, and local college characteristics at age 16, $\mathbf{c}_{i,16}$, from the wage equations (for similar restrictions regarding family background characteristics, see Willis and Rosen [1979], Taber [2001], Hotz et al. [2002], and Heckman, Stixrud, and Urzúa [2006] as well as Card [1995] and Kane and Rouse [1995], who use exclusion restrictions based on the existence of a local college and tuition at local colleges, respectively). In addition, while we allow current period local labor market conditions, $\mathbf{m}_{ia}$, to directly impact wages, past local labor market variables do not enter the wage equation. The assumption that these characteristics affect wages only indirectly through past activity choices that determine the accumulated experience variables that enter the wage equations is central in identifying the distribution of potential wages and the wage equation parameters from the realized wages of the selected group of labor market participants.

VI. Results

In this section, we present the results of our estimation. We first focus on how the specification of the log wage function impacts the measured returns to schooling and work experiences. In particular, we highlight the importance of generalizing the classic Mincer model by controlling for observable characteristics and selection on unobservable factors. Second, we discuss how the returns to schooling and work experiences, as well as the returns to unobserved ability as measured by our factor-loading estimates, have changed across cohorts. Note that our final and preferred specification—the full heterogeneity specification—consists of our wage equations, ability equations and activity-choice equations. The estimates and standard errors for the full set of parameters for this specification are provided in appendix D. We do not discuss the results of our activity-choice equations, as this part of our preferred specification is included solely for the purposes of dealing with the selection of wages and the endogeneity of experience terms in the wage equation.[36] Rather, in this section we focus on the results on returns to schooling, work-related experiences, and unobserved skills in our wage equations.

A. Specifications of Wage Equations

Our empirical framework allows us to estimate wage returns to various types of school and work experiences by accounting for the endogeneity of schooling and work choices. As described above, our most comprehensive (and preferred) specification of the wage equation includes nonlinear functions of school and work experience variables, indicators for graduation attainment and type of work, personal background characteristics, local labor market conditions, and measures for unobserved cognitive and noncognitive abilities. We compare this specification with other models, specifically, an augmented version of the classic Mincerian (1974) model where we control for high school and college graduation dummies and type of work dummies in addition to Mincer's quadratic in potential work experience and an augmented version of the flexible specification introduced in Heckman, Lochner, and Todd (2006). While our version of the latter specification (referred to as "augmented HLT" hereafter) is parametric, it remains very flexible and includes controls for race, ethnicity, high school and college graduation, cubic polynomials in school and potential work experience, and interaction between schooling and potential experience.

The classic Mincerian model restricts log wages to be a linear function of the number of years of schooling and a quadratic function of the number of years of potential experience (defined as age − years of schooling − 6). Focusing on earnings, Heckman, Lochner, and Todd (2006) consider a more

[36] As noted in sec. V.C, we imposed a number of exclusion restrictions to identify our model. To assess their importance, we used the following metrics of the statistical and economic significance of these restrictions. With respect to statistical significance, we first examined what fraction of each exclusion restriction has a t-statistic larger than 1.96 in magnitude in each activity value function in equation (11). Across all activity choices, we found that just under one-third of both the background characteristics, $\mathbf{f}_i$, and the local college characteristics, $\mathbf{c}_{i,16}$, were individually statistically significant. The local labor market conditions, $\mathbf{m}_{ia}$, were statistically significant in just under half of all of the choice equations. Second, we conducted a likelihood ratio test of the joint significance of the college characteristics, $\mathbf{c}_i$, in our preferred model. This test yielded p-values close to zero for both cohorts, thus rejecting the hypothesis that the college characteristics are not important. To assess the economic significance of imposing these restrictions, we compared the impact of a 1 standard deviation increase in each exclusion restriction with a 1 standard deviation increase in the unobserved cognitive factor. To do this, we calculated the ratio of marginal effects of each of these variables in each of the activity value functions, each multiplied by its respective standard deviation (note that the unobserved cognitive factor has a standard deviation of 1) and determined the percentage of these ratios that exceeded 1. By this criterion, we found that 40% of the $\mathbf{f}_i$ variables are economically significant (ratios exceeded 1.0) vs. 30% of the $\mathbf{m}_{ia}$ variables and 20% of the $\mathbf{c}_i$ variables.

flexible specification that uses indicators for each year of schooling and each year of potential experience and allow returns to potential experience to vary by levels of schooling: high school dropout, high school graduate, some college, and college graduate. They find that the internal rate of return to schooling changes drastically with the introduction of nonlinearities in schooling as well as nonseparability between schooling and work experiences.

Our preferred specification differs from Heckman, Lochner, and Todd (2006) in three notable ways.[37] First, we include controls for personal background characteristics, in particular nativity (native-born or foreign-born), birth year, and local labor market conditions (employment rate and income per capita).[38] The second difference relates to work experience. This is one of our main contributions, as we use actual work experience accumulated at each age a instead of potential work experience, distinguishing between in-high-school, in-college, part-time, full-time, and military work experiences. Third (and importantly), we control for selection into schooling and work experience levels on the basis of unobservable characteristics. We do so by allowing the cognitive skill factor, ξ_1, and the other, noncognitive skill factor, ξ_2, to enter the wage equation.

We estimate the model for all individuals i in our data set at each age a for which we observe them, up to and including age 35. We report the marginal effects associated with these different specifications and different variables of interest in tables 6 and 7.[39] For the accumulated experience variables, $\mathbf{x}^r_{ia}$ (i.e., schooling, work, military, etc.), that enter the model in a nonlinear fashion, we evaluate the marginal effects using the average experience vector at age 29 ($\overline{\mathbf{x}}^r_{29}$) but using parameters that are estimated from the entire age range.[40] We also report marginal effects at age 32 in tables C3 and C4.[41] Finally, our generalized Mincerian specification allows the marginal effects to vary over the life cycle through changes in the amount of accumulated experiences.

[37] As noted above, our analyses and those in Heckman, Lochner, and Todd (2006) do differ in two other notable ways. First, we focus on wages, while Heckman, Lochner, and Todd (2006) focused on earnings. Second, Heckman, Lochner, and Todd (2006) focused on the estimation of internal rates of return to schooling, while we focus on estimating marginal rates of return to schooling as well as to actual work experiences.

[38] Note that we do not directly control for the ASVAB test scores, as these are used as noisy measurements for the cognitive factor, ξ_1, which also enters the wage equation.

[39] The full estimation results are reported in app. D.

[40] We use this age because (i) it is an age by which most people have completed schooling and (ii) it is the last age for which we have a full-sized cross section in our panel.

[41] Consistent with sec. III, when we say age 29 (32), we are actually referring to the month before their twenty-ninth (thirty-second) birthday.

Table 6
Measures of Wage Returns to Schooling across Specifications at Age 29

Specification	NLSY79	NLSY97	NLSY97 − NLSY79
A. Return to year of schooling:			
i. Raw	.077***	.072***	−.005
ii. Augmented Mincer	.036***	.043***	.006
iii. Augmented HLT	.054***	.047***	−.006
iv. + Background	.043***	.043***	.000
v. + Actual experience	.006	.006	−.001
vi. + Unobserved	.014**	.046***	.032***
B. Return to graduation from high school (sheepskin):			
i. Raw	.191***	.197***	.007
ii. Augmented Mincer	.101***	.074***	−.027
iii. Augmented HLT	.102***	.073***	−.029
iv. + Background	.104***	.067***	−.037**
v. + Actual experience	.073***	.049***	−.023
vi. + Unobserved	.033**	.049***	.016
C. Return to graduation from college (sheepskin):			
i. Raw	.401***	.417***	.016
ii. Augmented Mincer	.299***	.294***	−.005
iii. Augmented HLT	.261***	.274***	.013
iv. + Background	.238***	.257***	.019
v. + Actual experience	.204***	.227***	.023
vi. + Unobserved	.187***	.187***	.001

NOTE.—Panel A is the wage return at age 29 of one extra year of schooling, panel B is the wage premium (sheepskin effect) of earning a high school diploma relative to not earning a diploma, and panel C is the wage premium (sheepskin effect) of earning a bachelor's degree relative to a high school diploma. Row i indicates raw premium, controlling only for type-of-work dummies (in school, part time, full time). Row ii adds to row i a quadratic in potential experience (age − years of schooling − 6), a linear term for years of schooling, and degree dummies. Row iii increases flexibility similar to Heckman, Lochner, and Todd (2006) and adds a cubic in schooling, a linear interaction between schooling experience and potential experience, and race/ethnicity indicators. Additionally, idiosyncratic error variance is allowed to be heteroskedastic by type of work. Row iv adds personal background characteristics and local labor market conditions. Row v replaces potential experience in row iv with actual work experience type (in school, part time, full time), military experience, and other experience. It also includes a linear interaction between schooling and actual work experiences except for military and other. Row vi adds person-specific random factors to account for dynamic selection (see eq. [10]). All standard errors are clustered at the individual level and are on the order of 0.005–0.020. NLSY = National Longitudinal Survey of Youth.
** Significant at the 5% level.
*** Significant at the 1% level.

B. Returns to Schooling

Table 6 presents estimates of the returns to schooling for our various specifications. Panel A displays the return to an additional year of schooling, while panels B and C present estimates of sheepskin effects for graduating from high school and college, respectively.[42] We report six different

[42] Because we include dummy variables for high school and college degrees in our wage equations, our estimates of the return to schooling measure the return to any additional year of schooling, including to years in which a degree is completed, but they do not capture any nonlinearities associated with degree completion.

Table 7
Measures of Wage Returns of Work Experiences at Age 29 for Selection- and Nonselection-Correction Specifications

Variable	NLSY79	NLSY97	NLSY97 − NLSY79
A. Full model without controlling for selection:			
Years of school	.006	.006	−.001
	(.008)	(.007)	(.011)
Work in high school	.029***	−.005	−.034***
	(.010)	(.008)	(.013)
Work in college	.065***	.044***	−.021
	(.012)	(.008)	(.014)
Work part time only	−.052***	−.049***	.003
	(.007)	(.006)	(.009)
Work full time only	.023***	.041***	.018***
	(.001)	(.002)	(.002)
Four years of college (no work)	.229***	.249***	.020
	(.022)	(.020)	(.029)
Four years of college (+ work)	.292***	.317***	.025
	(.025)	(.021)	(.033)
B. Full model controlling for selection:			
Years of school	.014***	.046***	.032***
	(.005)	(.004)	(.007)
Work in high school	−.001	−.024***	−.024**
	(.008)	(.005)	(.010)
Work in college	.066***	.0001	−.066***
	(.010)	(.0043)	(.011)
Work part time only	−.020***	−.008**	.012**
	(.005)	(.003)	(.006)
Work full time only	.022***	.039***	.017***
	(.001)	(.001)	(.002)
Four years of college (no work)	.242***	.372***	.129***
	(.015)	(.010)	(.018)
Four years of college (+ work)	.306***	.372***	.066***
	(.018)	(.011)	(.021)

NOTE.—Panel A refers to wage equation marginal effects without correcting for selection on unobservables. This is specification v ("+ Actual experience") in table 6. Panel B refers to wage equation marginal effects correcting for selection on unobservables. This is specification vi ("+ Unobserved") in table 6. Marginal effects are evaluated at the cohort-specific sample averages at age 29 for one additional year of each component of experience. Four years of college (no work) is calculated as the sum of the marginal effect of years of school (times 4) plus the return to graduation from college (sheepskin) from the relevant specification in table 6. Four years of college (+ work) is calculated as the sum of 4 years of college (no work) plus the marginal effect of work in college (times the average years spent working in college from table 3). NLSY = National Longitudinal Survey of Youth.
** Significant at the 5% level.
*** Significant at the 1% level.

specifications on separate rows within each panel, beginning with raw premia and ending with our preferred specification, which accounts for selection on observable and unobservable characteristics.

We start by comparing results for the augmented Mincerian and augmented HLT specifications, which are reported in rows ii and iii, respectively. There is virtually no difference in the estimated returns to high school graduation (panel B) across these two specifications, while the estimated returns to college graduation (panel C) for the augmented HLT specification are about 2–3 percentage points lower than for the augmented Mincer specification. In contrast, the estimated returns to an additional year of schooling (panel A) are slightly larger in the augmented HLT specification compared with the augmented Mincer, with the return to an extra year of schooling based on the former specification being about 2 points higher for the NLSY79 but nearly identical for the NLSY97.

In row iv of the panels in table 6, we extend the above specification to include controls for local labor market conditions (displayed in table C2), birth year, and nativity. Adding these variables slightly reduces the estimated returns to a year of schooling by 1.1 points for the NLSY79 cohort and 0.4 points for the NLSY97. This specification also results in smaller returns to college degrees, by about 2 points each. And while there is no impact on the high school sheepskin effect for the NLSY79, adding these controls does reduce it further for the NLSY97.

In row v of the panels in table 6, we present estimates for the wage equation specification in which we replace potential work experience with actual work experience. Note that these estimates do not account for the potential endogeneity of work experience. Relative to the estimates in the preceding rows of the panels, the estimates of returns to an extra year of schooling, high school, and college graduation are all substantially lower. Taken together, these findings suggest that a sizable part of the estimated returns to schooling and sheepskin effects in the previous rows actually may be attributable to returns to the work experiences individuals acquire during their transition from school to work. We examine the role of school-related work experiences in section VI.C below.

The estimated returns to schooling and degrees for the last and preferred specification we consider, which accounts for selection on unobservable characteristics, are found in row vi of the panels in table 6. This specification accounts for selection by jointly estimating the wage equation with our choice model and ability measurement equations, as described in section V.B. Compared with the estimates of our model that do not control for unobserved selection in row v, accounting for selection reduces the returns to college degrees for both cohorts (panel C) reduces the returns to high school for NLSY79 only (panel B) but increases the returns to each additional year of schooling (panel A). Importantly, the returns to schooling and degrees in row vi are much lower than the unadjusted ones in row i of each panel.

Finally, we compare how our estimates of the returns to schooling when one controls for selection in row vi have changed across these two cohorts. These changes are recorded in the last column of table 6 for each panel. We find that the estimated returns to an additional year of schooling (panel A) and the return to a high school degree (panel B) have both increased across the two NLSY cohorts, although the latter is not statistically significant at standard levels. Our estimation results also indicate that the return to college degree (panel C) has been essentially stable across these two cohorts. Finally, an important takeaway from this table is that the cross-cohort changes in the returns to education in row vi are quite different than the corresponding changes for the estimated returns produced by the other specifications, suggesting that the selection processes that govern educational and early work experiences have changed over the past 20 years.

Overall, we find that accounting for the accumulated actual work experiences of young men and their endogeneity not only affects one's conclusions about the magnitudes of returns to years of schooling and to degrees but also alters the conclusions one draws about how these returns have changed across cohorts.

C. Returns to Work Experiences

We next consider the returns to various types of work experiences and how they have changed across cohorts. Estimates for the returns to work experiences are presented in table 7. Panel A displays results for the wage equation specification that corresponds to controlling for actual work experience and was used to produce row v in table 6, while panel B is based on the selection-corrected wage equation used to produce the returns to education estimates in rows vi of table 6. The first marginal effect of both panel A and panel B of table 7 (year of school) is the same as rows v and vi of panel A of table 6, respectively. The second and third marginal effects of both panels display the additional returns to working while in high school and college, respectively. The next two display the estimated returns to part- and full-time out-of-school work experience. Finally, the remaining two rows report the total return to college graduation (assuming 4 years to degree) under two scenarios. The first one is the pure return to college, which is equal to four times the return to any schooling plus the college sheepskin effect. The second one accounts for the additional wage return to in-college work and is computed by adding to the previous return the product of the average number of years worked while in college and the return to an additional year of working while in college. As before, the estimated returns to the various types of work experiences are measured at age 29, with additional results at age 32 included in table C4.

We begin with the returns to working while in school. Consider, first, the returns to working while in college. For this type of work experience, we find sizable returns, with 6% for the NLSY79 and 4% for the NLSY97. Both are higher than those to any other form of work experience we consider or, for that matter, when we do not account for unobserved heterogeneity, the return to an extra year of pure schooling. However, when we account for unobserved heterogeneity, the return vanishes in the NLSY97 but stays the same in the NLSY79. This is notable considering the findings of the previous table: controlling for unobservable heterogeneity resulted in a substantial increase in the return to a year of school in the NLSY97. Thus, for the NLSY97, much of the perceived return to in-college work experience is actually the result of selection (on unobservables) in acquiring those work experiences. Regardless, with and without controls for unobserved heterogeneity, we see a decrease across cohorts in the return to working in college.

With respect to the returns to working while in high school, table 7 shows that they are lower in every instance than the corresponding returns to working while in college. Without controlling for unobserved heterogeneity, the estimated returns to working while in high school initially are 3% in the NLSY79 and negligibly small in the NLSY97. After controlling for unobserved heterogeneity, these returns become negligibly small for the NLSY79 cohorts and negative (−2.4%) for the NLSY97 cohorts. Our finding of a negligible or negative return to working while in high school is consistent with the findings of Hotz et al. (2002), who also estimate the wage returns to early work experiences using data from the NLSY79. Also, we once again see a decrease across cohorts in the returns to in-school work, this time for work in high school.

With respect to non-school-related work experiences, we estimate an increase in the return to an additional year of full-time experience, from 2% in the NLSY79 to 4% in the NLSY97. This return is robust to the inclusion of unobserved heterogeneity. In contrast, the estimated return to part-time, non-school-related experience is quite sensitive to controls for unobserved heterogeneity. The returns are about −5% without considering unobserved heterogeneity but become −2% in the NLSY79 and −1% in the NLSY97 thereafter. In short, it appears that those individuals who tend to accumulate part-time, non-school-related work experience are negatively selected on unobservables so that failure to control for unobserved heterogeneity greatly exaggerates the detrimental consequences of early part-time work on subsequent wages of young men.

Finally, as mentioned above, the last two rows of each panel report the total return to 4-year college, with and without accounting for in-college work experience. In both cases, the estimated returns show only a modest and nonsignificant increase across cohorts when ignoring unobservable selection. However, results from our preferred specification that controls for unobserved heterogeneity point to significant and quantitatively sizable increases across cohorts in the returns to 4-year college. Namely, we find that this return is 12.9 (6.6) percentage points higher in the NLSY97 when one ignores (accounts for) in-college work experience. Furthermore, these results imply that the cohort improvement in the return to college plus work in college is solely the result of the across-cohort increases in the return to the additional 4 years of attending college and in acquiring a college degree, given the negligible return to work while in college found for the NLSY97 cohorts.

Taken together, our results indicate that the returns to work experiences, especially those for in-school and part-time out-of-school work experiences,

Table 8
Full-Time Wage Factor-Loading Estimates

Variable	NLSY79	NLSY97	NLSY97 − NLSY79
Cognitive	.148***	.111***	−.037***
	(.003)	(.002)	(.003)
Noncognitive	.091***	.161***	.070***
	(.003)	(.002)	(.004)

NOTE.—Factor-loading estimates are from the specification found in the "+ Unobserved" row in table 6. NLSY = National Longitudinal Survey of Youth.
*** Significant at the 1% level.

differ substantially depending on whether one controls for unobserved heterogeneity, which has significant impacts on the implied cross-cohort changes in the returns to work experiences.

D. Returns to Unobserved Skills

Finally, we examine the contribution of the unobserved factors to the wages of young men. Table 8 contains estimates of the cognitive and noncognitive factor loadings for the full-time wage equation for each of the three cohorts. Recall that the distribution of the factors is multivariate normal with mean 0 and identity covariance matrix. It follows that these estimates can be interpreted as the change in log wages due to a 1 standard deviation increase in the corresponding unobserved factor, holding fixed all observable characteristics and the other dimension of unobserved ability.

We find that the wage return to cognitive ability (or cognitive skills) of young men decreased across cohorts from 15% to 11% for a 1 standard deviation increase in cognitive skills. On the other hand, assuming stability in the distribution of noncognitive skills across cohorts, the return to these skills increased from 9% to 16% for a 1 standard deviation increase in the cognitive factor.[43] Interestingly, our results are consistent with Castex and Dechter (2014) and Deming (2017), who also examine the wage returns to skills between the NLSY79 and NLSY97 cohorts and find that the returns to cognitive skills (as measured by AFQT) have diminished across the two. Additionally, Deming (2017) also finds an increasing return to noncognitive skills across both cohorts.

VII. Conclusion

This paper examines the returns to both schooling and various forms of work experience for men from two birth cohorts, using longitudinal data from the 1979 and 1997 panels of the NLSY. To deal with the endogenous nature of accumulated work experience and schooling and its potential

[43] If the distribution of noncognitive skills changed across cohorts, then an increase in our factor loadings will instead reflect an increase in the variance of noncognitive skills.

impact on estimating the wage returns to these different types of experience, we develop and estimate a dynamic model of the schooling and work decisions that individuals make in their early adulthood and how they affect subsequent wages for each of these cohorts. Building on previous work by Heckman, Lochner, and Todd (2006), our empirical framework generalizes the classic Mincerian model of returns to human capital in four main ways: (i) it allows for a more flexible function of schooling and work experiences rather than the original linear-quadratic specification; (ii) it incorporates additional controls for an individual's background as well as degree sheepskin effects; (iii) it accounts for individual-specific multidimensional unobservable heterogeneity to correct for the endogeneity of past human capital investment decisions; and, importantly, (iv) it moves away from the concept of potential experience by differentiating among and controlling for various forms of work experience that were actually attained by the individual.

Based on the estimates from this model, we produce several key findings. First, the failure of previous estimates to account for the influences of accumulated actual work experience and its endogenous determination results in sizable overstatements of the wage returns to degree attainment and, for the 1979 cohort, of the wage returns to schooling. Second, we find that the returns to various types of school and work experiences significantly differ between cohorts. For example, we find that the returns to an extra year of schooling increased across cohorts, while the returns to an additional year of in-school work decreased. The latter finding could partly reflect the changing nature of high school and college employment, with students oftentimes holding low-skill jobs that do not significantly improve their future employment prospects, as noted by Baum and Ruhm (2016). Although the return to a college degree has remained stable, the overall return to 4 years of college has increased. Third, consistent with Deming (2017), we find that the return to unobservable cognitive skills has declined, while the return to other, noncognitive skills may have increased.

Overall, our analysis highlights the need to account for dynamic selection and changes in composition of skills when analyzing secular changes in the wage returns to skills. An interesting future research avenue would be to build on our analysis and estimate a dynamic generalized Roy model to quantify the relative importance of cross-cohort changes in wage returns to skills and nonwage components—in particular, increasing costs of college education—in explaining changes in the acquisition of schooling and early work experiences.

References

Altonji, Joseph G., Prashant Bharadwaj, and Fabian Lange. 2009. Constructing AFQT scores that are comparable across the NLSY79 and the NLSY97. Technical report, Yale University.

———. 2012. Changes in the characteristics of American youth: Implications for adult outcomes. *Journal of Labor Economics* 30, no. 4:783–828.

Arcidiacono, Peter, Esteban Aucejo, Arnaud Maurel, and Tyler Ransom. 2016. College attrition and the dynamics of information revelation. NBER Working Paper no. 22325, National Bureau of Economic Research, Cambridge, MA.

Ashworth, Jared, and Tyler Ransom. 2019. Has the college wage premium continued to rise? Evidence from multiple U.S. surveys. *Economics of Education Review* 69:149–54.

Aucejo, Esteban, and Jonathan James. 2019. Catching up to girls: Understanding the gender imbalance in educational attainment within race. *Journal of Applied Econometrics* 34, no. 4:502–25.

Bacolod, Marigee, and V. Joseph Hotz. 2006. Cohort changes in the transition from school to work: Evidence from three NLS surveys. *Economics of Education Review* 25, no. 4:351–73.

Baum, Charles L., and Christopher J. Ruhm. 2016. The changing benefits of early work experience. *Southern Economic Journal* 83, no. 2:343–63.

Belzil, Christian, and Jorgen Hansen. 2002. Unobserved ability and the return to schooling. *Econometrica* 70, no. 5:2075–91.

———. 2020. The evolution of the US family income–schooling relationship and educational selectivity. *Journal of Applied Econometrics* 35, no. 7:841–59.

Böhm, Michael J. 2020. The price of polarization: Estimating task prices under routine-biased technical change. *Quantitative Economics* 11, no. 2: 761–99.

Bound, John, Michael F. Lovenheim, and Sarah Turner. 2012. Increasing time to baccalaureate degree in the United States. *Education Finance and Policy* 7, no. 4:375–424.

Cameron, Stephen V., and James J. Heckman. 1998. Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males. *Journal of Political Economy* 106, no. 2:262–333.

———. 2001. The dynamics of educational attainment for black, Hispanic, and white males. *Journal of Political Economy* 109, no. 3:455–99.

Card, David. 1995. Using geographic variation in college proximity to estimate the return to schooling. In *Aspects of labor market behaviour: Essays in honour of John Vanderkamp*, ed. Louis N. Christofides, E. Kenneth Grant, and Robert Swidinsky. Toronto: University of Toronto Press.

———. 2001. Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica* 69, no. 5:1127–60.

Card, David, and Thomas Lemieux. 2001. Can falling supply explain the rising return to college for younger men? *Quarterly Journal of Economics* 116, no. 2:705–46.

Carneiro, Pedro, and Sokbae Lee. 2011. Trends in quality-adjusted skill premia in the United States, 1960–2000. *American Economic Review* 101, no. 6:2309–49.

Castex, Gonzalo, and Evgenia Kogan Dechter. 2014. The changing roles of education and ability in wage determination. *Journal of Labor Economics* 32, no. 4:685–710.

Cunha, Flavio, Fatih Karahan, and Ilton Soares. 2011. Returns to skills and the college premium. *Journal of Money, Credit and Banking* 43, no. 5:39–86.

Currie, Janet, and Enrico Moretti. 2003. Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *Quarterly Journal of Economics* 118, no. 4:1495–532.

Deming, David J. 2017. The growing importance of social skills in the labor market. *Quarterly Journal of Economics* 132, no. 4:1593–640.

Fang, Hanming. 2006. Disentangling the college wage premium: Estimating a model with endogenous education choices. *International Economic Review* 47, no. 4:1151–85.

Fortin, Nicole M. 2006. Higher-education policies and the college wage premium: Cross-state evidence from the 1990s. *American Economic Review* 96, no. 4:959–87.

Heckman, James J., John Eric Humphries, and Gregory Veramendi. 2018. Returns to education: The causal effects of education on earnings, health and smoking. *Journal of Political Economy* 126, no. S1:S197–S246.

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2006. Earnings functions, rates of return and treatment effects: The Mincer equation and beyond. In *Handbook of the economics of education*, vol. 1, ed. Eric A. Hanushek and Finis Welch, 307–458. Amsterdam: Elsevier.

Heckman, James J., Jora Stixrud, and Sergio Urzúa. 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics* 24, no. 3:411–82.

Hotz, V. Joseph, Lixin Colin Xu, Marta Tienda, and Avner Ahituv. 2002. Are there returns to the wages of young men from working while in school? *Review of Economics and Statistics* 84, no. 2:221–36.

Hu, Yingyao, and Matthew Shum. 2012. Nonparametric identification of dynamic models with unobserved state variables. *Journal of Econometrics* 171:32–44.

Kane, Thomas J., and Cecilia Elena Rouse. 1995. Labor-market returns to two- and four-year college. *American Economic Review* 85, no. 3:600–614.

Katz, Lawrence F., and Kevin M. Murphy. 1992. Changes in relative wages, 1963–1987: Supply and demand factors. *Quarterly Journal of Economics* 107, no. 1:35–78.

Kejriwal, Mohitosh, Xiaoxiao Li, and Evan Totty. 2020. Multidimensional skills and the returns to schooling: Evidence from an interactive fixed-effects approach and a linked survey-administrative data set. *Journal of Applied Econometrics* 35, no. 5:548–66.

Kling, Jeffrey R. 2001. Interpreting instrumental variables estimates of the returns to schooling. *Journal of Business and Economic Statistics* 19, no. 3:358–64.

Kotlarski, Ignacy. 1967. On characterizing the gamma and the normal distribution. *Pacific Journal of Mathematics* 20, no. 1:69–76.

Lee, Donghoon, and Kenneth I. Wolpin. 2010. Accounting for wage and employment changes in the U.S. from 1968–2000: A dynamic model of labor market equilibrium. *Journal of Econometrics* 156, no. 1:68–85.

Lee, Sang Yoon, Yongseok Shin, and Donghoon Lee. 2015. The option value of human capital: Higher education and wage inequality. NBER Working Paper no. 21724, National Bureau of Economic Research, Cambridge, MA.

Mincer, Jacob. 1974. *Schooling, experience and earnings*. New York: Columbia University Press for the National Bureau of Economic Research.

Moretti, Enrico. 2011. Local labor markets. In *Handbook of labor economics*, vol. 4B, ed. Orley Ashenfelter and David Card, 1237–313. Amsterdam: Elsevier.

Scott-Clayton, Judith. 2012. What explains trends in labor supply among U.S. undergraduates? *National Tax Journal* 65, no. 1:181–210.

Taber, Christopher R. 2001. The rising college premium in the eighties: Return to college or return to unobserved ability? *Review of Economic Studies* 68, no. 3:665–91.

Valletta, Robert G. 2019. Recent flattening in the higher education wage premium: Polarization, skill downgrading, or both? In *Education, skills, and technical change: Implications for future U.S. GDP growth*, ed. Charles Hulten and Valerie Ramey, 313–242. Chicago: University of Chicago Press.

Willis, Robert J., and Sherwin Rosen. 1979. Education and self-selection. *Journal of Political Economy* 87, no. 5:S7–S36.