Ability Bias, Human Capital, or Sheepskin Effects? A Factor Model-Based Decomposition of the Returns to Schooling *

Mohitosh Kejriwal *Purdue University*

Xiaoxiao Li *Villanova University*

Evan Totty *U.S. Census Bureau*

March 20, 2025

Abstract

The relative importance of human capital and signaling frameworks in explaining the returns to schooling has been a subject of extensive research and debate due to their vastly different implications for education policy and society at large. Empirical resolution of the issue is complicated by the fact that both frameworks are consistent with a positive earnings-schooling nexus coupled with the challenge of controlling for ability bias which itself can generate such a nexus. This paper offers a new approach to decomposing the apparent returns to schooling into its human capital, sheepskin (i.e., degree attainment), and ability bias components. We adopt a flexible factor model framework that allows treating ability as an unknown multidimensional object that can be estimated using panel data on earnings and schooling. Our approach obviates the need to rely on inadequate ability proxies and does not depend on the sequence in which covariates are added to the model. Our results using data from the NLSY79 show that about 56% of the apparent returns to schooling can be attributed to ability bias while the sheepskin and human capital components contribute about 38% and 6%, respectively. The sheepskin contribution is found to be driven almost entirely by the college degree while the high school degree plays only a very minor role.

Keywords: human capital, degree signaling, interactive fixed effects, bias decomposition
JEL Classification: C33, C38, I26

*The opinions in this paper are those of the authors and may not represent the views of the U.S. Census Bureau or other organizations. Contact information: Mohitosh Kejriwal (mkejriwa@purdue.edu), Xiaoxiao Li (xiaoxiao.s.li@villanova.edu), Evan Totty (evan.scott.totty@gmail.com).

1 Introduction

Consider the following quotes from two influential surveys of the literature on the returns to education:

"Our review of the available empirical evidence on Job Market Signaling leads us to conclude that there is little in the data that supports Job Market Signaling as an explanation for the observed returns to education." (Lange and Topel, 2006)

"My claim is that education is mostly signaling. Given all the evidence, a 20/80 human capital/signaling split seems reasonable." (Caplan, 2019)

These quotes serve to exemplify the lack of consensus in the literature on what explains the observed returns to education. The primary competing frameworks that have been proposed to capture the mechanisms through which education increases earnings are the human capital model and the signaling model. The human capital model (Becker, 1962) asserts that education raises earnings by imparting skills that enhance productivity. In contrast, the signaling model (Spence, 1973) posits that education increases earnings by providing information on pre-existing skills. The two frameworks have vastly different implications for education policy and society at large. According to the human capital view, education provides positive spillover effects to society so that policy interventions directed at raising schooling to the level that balances its marginal social benefit with marginal social cost would be desirable. On the other hand, adherence to the pure signaling view suggests that all returns are private and education is socially inefficient such that government efforts (e.g., tuition subsidies) to increase educational attainment are generally wasteful. Avenues for improving social welfare are those that mitigate the costs of conveying information regarding individual skills. For detailed discussions of the human capital and signaling frameworks, see, inter alia, Lovenheim and Turner (2017) and Gunderson and Oreopolous (2020).

Quantifying the relative importance of human capital versus signaling in practice is difficult.

In both models, individuals choose different levels of education based on their skill levels and other background characteristics that are plausibly of independent value in the labor market. It is therefore onerous to find variation in receipt of a given credential that is unrelated to variation in human capital, which is essential to determining the signaling value of the credential. Moreover, a major challenge in this literature is the difficulty of controlling for ability bias which itself is capable of generating a positive earnings-schooling nexus that is consistent with both the human capital and signaling models. In a comprehensive survey of the literature, Caplan (2019) documents the substantial role of ability bias and the methods used to control for the same in determining the relative contributions of human capital and signaling. Two commonly adopted approaches to address ability bias are the use of instrumental variables that exploit variation in schooling potentially unrelated to ability and the use of observable proxies for ability. While the instrumental variable approach has been criticized on the grounds of susceptibility to problems of weak instruments and potential failure of instrument exogeneity (Bound et al., 1995; Bound and Jaeger, 2000; Buckles and Hungerman, 2013), the proxy approach has been shown to be inadequate in capturing the latent skills (Das and Polachek, 2019; Kejriwal et al., 2024).

This paper offers a new approach to decomposing the apparent returns to schooling into its human capital, signaling, and ability bias components. The signaling component is measured by the effect of degree completion (i.e., a high school or college degree) on returns, commonly referred to as the "sheepskin effect." To model the ability component, we adopt a flexible factor model framework, developed by Pesaran (2006) and Bai (2009), that treats ability as a latent multidimensional vector with associated prices that are potentially time-varying. The factor loadings and the common factors, which represent the unobserved skills and their prices, respectively, are estimated using panel data on earnings and schooling. By exploiting the within individual variation in earnings and schooling, our approach obviates the need to rely on ability proxies or instruments. Our decomposition analysis allows us to express the least squares estimator of the returns obtained from a model that only includes the human capital and experience variables as the

sum of three components obtained from estimating the most general model that also includes the degree completion variables and the factor structure: (i) the factor model estimator of the returns which represents the human capital component; (ii) the signaling component obtained from the estimated effect of the degree completion variables; (iii) the ability bias component obtained from the estimated factor structure.

A notable feature of our approach is that it does not depend on the sequence in which covariates are added to the model. This is an important issue in practice since sequential covariate addition to a base empirical specification can yield considerably different results regarding the importance of covariates, if the covariates are intercorrelated (Gelbach, 2016). In our context, the base specification is one that only involves the human capital and experience variables, while the additional covariates are those pertaining to degree completion which determine the contribution of the signaling component and the estimated factor structure that determines the contribution of the ability component.

Our empirical analysis is based on an unbalanced panel data set obtained from the National Longitudinal Survey of Youth (NLSY79). Our results show that about 56% of the apparent returns to schooling can be attributed to ability bias while the diploma signaling and human capital components contribute abound 38% and 6%, respectively. The signaling contribution is found to be driven almost entirely by the college degree while the high school degree plays only a very minor role. We also find that ignoring the presence of ability bias yields a human capital/signaling split of about 40/60, thereby overstating the contribution of both components (but particularly the human capital component) and underscoring the empirical importance of appropriately accounting for ability bias.

A potential limitation of the panel data approach is that identification requires a sample of individuals who are simultaneously working and obtaining education. This may include, for example, students who work while obtaining a bachelor's or individuals who return to school later in life, whether to finish an uncompleted degree or for additional degrees (e.g., MBA). Moreover,

earnings before or during schooling could be part-time or seasonal work and not fully reflect an individual's earning potential (Card, 1995; Lazear, 1977). This raises the issue of sample selection since such a sample could be different from the traditional idea of a student who completes degrees consecutively and does not work while in school. We believe that these concerns are alleviated by the facts that (i) we restrict our sample of analysis to individuals who reported working at least 30 weeks and 800 hours in the previous year to capture an individual's full earning potential, following Koop and Tobias (2004); (ii) we find similar summary statistics for samples that are less restrictive than our sample of analysis; (iii) the population of students that work while in school is large (Bound et al., 2012; Carnevale et al., 2015).

Our study builds upon previous work in Kejriwal et al. (2020) and Kejriwal et al. (2024) that also employ a factor model framework in conjunction with panel data to estimate the returns to schooling. While all three studies find that the factor model estimates of the returns are substantially smaller than the standard least squares estimates, the present study makes a distinctive contribution relative to these earlier works. Kejriwal et al. (2020) use a unique linked-survey administrative data set to estimate models that allow for individual level heterogeneity in returns and show that ability bias accounts for a larger fraction of the aggregate least squares bias compared to heterogeneity. Kejriwal et al. (2024) utilize the factor model framework as a test bed for evaluating the efficacy of ability proxies in estimating the returns to schooling and document their failure in explaining the estimated ability bias. Unlike the present study, neither Kejriwal et al. (2020) nor Kejriwal et al. (2024) address the impact of signaling on returns or assess its contribution relative to the human capital and ability bias components.

The rest of the paper is organized as follows. Section 2 presents a review of the related literature. Section 3 lays out the empirical framework. Section 4 discusses the data used in the empirical analysis. Section 5 presents the empirical results. Section 6 concludes. Additional empirical results are contained in Supplementary Appendices A-C (not for publication).

2 Issues in the Existing Literature

Quantifying the relative importance of human capital versus signaling remains a critical and ongoing debate in economics (Lange and Topel, 2006; Lovenheim and Turner, 2017; Bradley and Green, 2020). The pure human capital model posits that more education enhances individual productivity, leading to higher wages and, at an aggregate level, greater economic growth (Becker, 1962; Krueger and Lindahl, 2001). In contrast, signaling purism argues that education primarily acts as productivity signals, thereby improving employment prospects and wage outcomes without necessarily increasing productivity (Spence, 1973; Weiss, 1995). How education influences productivity and wages is one of the oldest and most fundamental questions in economics. However, the empirical literature has yet to reach a consensus on the relative contributions of human capital and signaling to the private returns on education. This challenge arises because both models predict a positive relationship between education and wages, making it difficult to empirically disentangle their effects (Deming, 2022).

One prediction of the signaling theory suggests that individuals who earn diplomas are likely to receive higher wages than peers with equivalent years of education but without diplomas. Unlike other educational metrics such as years of schooling, course credits, or GPA, a diploma is simply a document and does not intrinsically improve productivity. As such, any earnings premium linked to holding a diploma primarily reflects its signaling value (Clark and Martorell, 2014). This has prompted numerous empirical studies to investigate the so-called "sheepskin effects" by identifying the wage disparities between those who have obtained a degree or diploma and those who have not, conditional on years of schooling. And the existence of such education-based signaling is well-supported by an extensive body of research (e.g., Hungerford and Solon 1987; Jaeger and Page 1996; Park 1999; Frazis 2002). Caplan (2019) conducts a comprehensive review of the literature and concludes in favor of significant sheepskin effects, with high school diplomas associated with wage increases of approximately 3-27% and even greater effects for college diplo-

mas, ranging from 24-58% (please refer to Caplan 2019, Tables E.1 and E.2 for further details), and suggests a human capital/signaling split of roughly 60/40 for high school and 40/60 for college diplomas.[1]

The interpretation of sheepskin effects as signaling value has faced significant critiques. Chiswick (1973) argues that if graduates are typically more efficient learners who receive a larger productivity boost from schooling than dropouts do, and thus choose to complete their degree programs, then the wage differences observed between graduates and dropouts could be misleading. Such disparities might more accurately reflect a self-selection process among students rather than the signaling value of the diploma (Lange and Topel, 2006). Relatedly, degree completion might be associated with other desirable labor market traits, such as perseverance, which can yield higher wages. Thus, estimators that fail to adjust for these observable and unobservable productivity differences are subject to omitted variable bias. Various studies that incorporate ability measures such as ASVAB or AFQT scores, IQ scores, standardized test results, high school rank, and GPA generally continue to report significant sheepskin effects (e.g., Frazis 1993; Kane and Rouse 1995; Arkes 1999; Light and Strayer 2004). However, as demonstrated in Kejriwal et al. (2024), the ability proxies commonly used in the literature do not adequately capture latent skills. Critics argue that the apparent sheepskin effects could diminish with more and better measures of ability. Supporting this argument, Clark and Martorell (2014) utilizes a regression discontinuity design to analyze earnings differences between high school students who just passed or failed an exit exam, and their causal evidence suggests no signaling value of a high school diploma.

Given the challenges faced by the sheepskin effects approach, the conversation around human capital and signaling continues to evolve, exploring other methodologies. A notable shift towards the employment-learning framework offers an alternative perspective to examine this

[1]The extent of the contribution from sheepskin effects can differ based on the choice of model specifications and variables. For example, Jaeger and Page (1996) analyze the relative contribution of sheepskin effects by comparing coefficients from models that either include or exclude explicit degree indicators. His findings suggest that sheepskin effects represent 26% of the total return on completing 16 years of education and 13.56% on 12 years of education. See Jaeger and Page (1996), Figure 2 for the detailed decompositions.

debate (Farber and Gibbons, 1996; Altonji and Pierret, 2001). This approach is based on the intuition that as firms learn about a worker's productivity over time, the signaling component of the wage premium should diminish, meaning that sheepskin effects become less pronounced as one's career progresses. Utilizing AFQT as a cognitive ability measure that is unobserved by firms, Lange (2007) finds that employers learn quickly, halving their initial errors within three years, and estimates the upper bound of signaling's contribution to the returns to education of about 25% under strong assumptions required for the bounding exercise (see Lange 2007, section IV for detailed explanation). Extending this framework, Pasche (2009) incorporates non-cognitive ability measures using Rotter Locus of Control and Rosenberg Self-Esteem Scales, and estimates lifetime human capital/signaling splits ranging from 39/61 to 95/5, depending on the parameter of interest rate and the speed of employer learning. Most recently, Aryal et al. (2022) integrate transparent and hidden IVs within the employer learning model, exploiting variations in Norway's compulsory schooling laws. Their analysis reveals a private return to education of about 7.9 percent, attributing 70 percent to human capital accumulation and the remaining 30 percent to signaling effects. One of the major critiques of the employer-learning literature is its neglect of unmeasurable abilities, which are crucial yet challenging to quantify.[2] Furthermore, as noted above, Kejriwal et al. (2024) employ the factor model approach and provide a formal evaluation of the commonly used proxies for ability including AFQT scores, Rotter Locus of Control Scale, and Rosenberg Self-Esteem Scale. Their finding reveals the inadequacy of these ability proxies in capturing the underlying latent skills. Taken together, these limitations have significant implications for accurately estimating the speed of employer learning and, consequently, for effectively quantifying the contribution of signaling effects to the returns on education.

Further efforts have been made to delineate the comparative impacts of human capital and signaling theories on the private returns to education. For instance, Bedard (2001) addresses the debate by analyzing how the distribution of high school graduates varies across regions with

[2]For additional critiques of the employer-learning approach, see Caplan 2019, Chapter 4.

differing university access. Contrary to the predictions of a pure human capital model, which suggests that increased university access would not impact dropout rates, her findings reveal that regions with universities actually experience higher dropout rates. This evidence supports the signaling view, indicating that the presence of universities may influence educational decisions in ways not accounted for by human capital models alone. Another approach, rooted in the cross-country comparison and economic growth literature, examines the existence and size of signaling through the lens of disparities between national and private returns to education (see Krueger and Lindahl, 2001 for an introduction). This method is based on the premise that while the human capital model predicts education boosts individual productivity and national income, the signaling model suggests education only enhances perceived productivity, thus affecting individual but not national income. Therefore, any discrepancy between social and private returns highlights the importance of the signaling component. Caplan (2019) surveys this strand of the literature and suggests a 20/80 human capital/signaling split. Fang (2006) develops a general equilibrium structural model of endogenous education choices and wage determination and evaluates the roles of these two competing theories in explaining the college wage premium. His analysis indicates that college education enhances individuals' abilities by approximately 40%, and the signaling channel contributes to at most one-third of the observed college wage premium. Arteaga (2018) leverages a natural experiment involving a curriculum reform at Universidad de Los Andes, which reduced the coursework required for economics and business degrees, to separate the effects of human capital accumulation from signaling. Her results reveal that post-reform, wages of graduates in economics and business fell by 16% and 13%, respectively, which suggests a substantial role for human capital in determining wages and challenges a pure signaling model, as it indicates that the actual content of education, and not just the degree itself, impacts earnings.[3]

The debate on the relative importance of human capital versus signaling in explaining returns

[3]For a summary of other perspectives, such as through the lens of malemployment, see Lange and Topel (2006) and Caplan (2019).

to schooling continues as the available empirical evidence is insufficient to reach a definitive conclusion (Lange and Topel, 2006; Bradley and Green, 2020). In this paper, we advance the ongoing debate by re-evaluating the importance of sheepskin effects using a factor model framework that effectively accounts for previously omitted productivity differences (Kejriwal et al., 2020). The factor model approach, following Pesaran (2006) and Bai (2009), utilizes an unobserved factor structure to model skills and their corresponding prices, where factor loadings represent multidimensional latent skills and common factors denote their associated prices. We apply this framework to the NLSY79, a commonly adopted panel dataset, and exploit within-individual variations in earnings and education to consistently estimate the returns to schooling without resorting to external instruments or ability proxies. Utilizing the methodology of Gelbach (2016), we perform a rigorous bias decomposition of the apparent returns to schooling and quantify the relative contributions of human capital and signaling. This decomposition analysis, not previously undertaken in the sheepskin effect literature, sheds new light on the debate. Specifically, we seek to unravel two critical questions: Firstly, what are the relative contributions of human capital, the sheepskin effect, and ability bias to the native returns to schooling? Secondly, how does ability bias affect the balance between human capital and the sheepskin effect in explaining the returns to education?

Our findings reveal an interesting picture of the returns to education: Ability bias emerges as the dominant factor, accounting for the largest share of the apparent returns to schooling, while the sheepskin effect, predominantly from college degrees, follows. Remarkably, the human capital component appears to contribute minimally, if at all, with an observed distribution of 6% for Human Capital, 38% for Signaling, and 56% for Ability Bias. Furthermore, without adjusting for productivity factors, the split between Human Capital and Signaling is more evenly distributed at 40/60. However, when productivity factors are controlled, the split shifts dramatically to 13/87 in favor of signaling, corroborating Caplan (2019). This stark difference underscores the critical role that ability bias plays in influencing the returns to education and the relative importance of

human capital and sheepskin signaling effects.

3 Empirical Framework

This section details the factor model (FM henceforth) framework that forms the basis of our empirical analysis, provides a description of the various econometric techniques employed, and delineates how our framework addresses the research questions posed in the paper. Section 3.1 discusses the two FM estimators, namely, the interactive fixed effects and common correlated effects estimators. Section 3.2 presents a decomposition, originally developed by Gelbach (2016) for the linear regression model, of the OLS returns into its human capital, sheepskin, and ability bias components. Section 3.3 spells out the research questions of interest and discusses how our framework facilitates their analysis. Much of the material presented in section 3.1 is borrowed from Kejriwal et al. (2024) who employ the FM framework to assess the usefulness of ability proxies in estimating the returns to schooling.

3.1 The Factor Model And Related Estimators

We consider an unbalanced panel data factor model specified as

$$y_{it} = c_i + s_{it}\beta_1 + s_{it}^2\beta_2 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + z_{it}'\delta + v_{it} \tag{1}$$

$$v_{it} = \lambda_i' f_t + u_{it}, \;\; i = 1,...,N; \;\; t = t_i \in \mathscr{J}_i \equiv \{t_i(1), t_i(2), ..., t_i(T_i)\} \tag{2}$$

where y_{it} and s_{it} represent, respectively, the (log of) hourly wage and the years of schooling completed for person i at period t, e_{it} is a measure of labor market experience, and z_{it} is a vector of degree receipt (sheepskin) dummies for person i at period t. Specifically, $z_{it} = (HS_{it}, COLL_{it})'$, where HS_{it} $(COLL_{it})$ takes the value zero before an individual obtains a high school (college) degree and the value one thereafter.

The set $\mathscr{J}_i$ defined in (2) includes the time indices of the non-missing observations for

person i. For each i, there are T_i observations available at times $\{t_i(1), t_i(2), ..., t_i(T_i)\}$ where T_i can be different across i. Let T denote the length of the complete time period, i.e., $T = \max_{i \leq N}\{t_i(T_i)\}$.

The error term v_{it} has an interactive fixed effects (IFE) structure and is composed of a common component $(\lambda_i' f_t)$ and an idiosyncratic component (u_{it}). The $(r \times 1)$ vector λ_i represents a set of unmeasured skills (factor loadings), such as innate abilities, while f_t is a $(r \times 1)$ vector of unobserved, possibly time-varying, prices (or common factors) of the unmeasured skills.[4] Both the loadings and factors are allowed to be correlated with schooling so that the same skills and prices that influence earnings may also influence schooling. The number of common components r is assumed unknown. Note that while the returns to the skill components $(\lambda_i' f_t)$ are identified, the skills and their prices are not separately identified.[5] The estimated factors and loadings thus only estimate a rotation of the underlying true parameters and so cannot be given a direct economic interpretation. We follow Bai (2009) and Pesaran (2006) in referring to skills/abilities as factor loadings and the common time shocks or prices as factors. Heckman et al. (2006) and others in the returns to schooling literature also estimate what is referred to as a factor model, although the methods are not the same (see Kejriwal et al., 2024, for a comparison). Importantly, they refer to the skills/abilities as the factors rather than the factor loadings.

A set of person fixed effects c_i is included to control for time-invariant person characteristics such as gender, race, mother's and father's education, and the number of siblings. We also consider a variant of (1) that includes a set of demographics-by-year fixed effects which allows us to investigate the extent to which the factor structure in (2) can be interpreted in terms of time varying returns to observable time-invariant characteristics.

The quadratic schooling specification has been routinely adopted in empirical studies; see, e.g., Lemieux (2006). The concavity of log earnings as a function of years of schooling arises in a

[4]While we refer to the factor loadings as skills/abilities, there are other time-invariant determinants with possibly time-varying prices, such as motivation and persistence, that can be captured by the factor loadings as well.

[5]For an arbitrary $(r \times r)$ invertible matrix A, we have $F\Lambda' = FAA^{-1}\Lambda' = F^*\Lambda^{*\prime}$, so that a model with $F = (f_1, ..., f_T)'$ and $\Lambda = (\lambda_1, ..., \lambda_N)'$ is observationally equivalent to a model with $F^* = (f_1^*, ..., f_T^*)'$ and $\Lambda^* = (\lambda_1^*, ..., \lambda_N^*)'$ where $F^* = FA$ and $\Lambda^* = \Lambda A^{-1\prime}$.

simple human capital investment model in which individuals have different preferences (discount rates) but all face the same concave production function (the return to a year of schooling declines as years of schooling increase). Mincer (1996) shows that in a model where individuals have heterogenous preferences and earnings opportunities, average log earnings may either be a convex or a concave function of years of schooling. The importance of allowing for interaction between schooling and experience is underscored in Heckman et al. (2006) who show that log earnings-experience profiles are not parallel across schooling levels. We also estimate the benchmark canonical Mincerian regression, with linear (as opposed to quadratic) schooling and quadratic experience as well as specifications with quartic age controls (Murphy and Welch, 1990). Our parameter of interest is the marginal returns to schooling (MRTS) evaluated at the overall mean (across individuals and time):

$$MRTS = \beta_1 + 2\bar{s}\beta_2 + \bar{e}\gamma_3, \tag{3}$$
$$\bar{s} = N^{-1}\sum_{i=1}^{N} T_i^{-1}\sum_{t\in\mathscr{J}_i} s_{it}, \quad \bar{e} = N^{-1}\sum_{i=1}^{N} T_i^{-1}\sum_{t\in\mathscr{J}_i} e_{it}$$

It is important to emphasize that unlike Heckman et al. (2006), our paper does not attempt to distinguish between the role of cognitive and non-cognitive skills in explaining the behavior of earnings. Rather, we are interested in estimating the average rate of growth of earnings with schooling given by (3) employing the interactive fixed effects structure as a device to control for the different components of ability that may affect earnings and are potentially correlated with schooling.

3.1.1 The Interactive Fixed Effects (IFE) Estimator

The IFE approach to estimating the factor model (1)-(2) was originally proposed by Bai (2009) for the balanced panel framework. In this approach, the regression coefficients and the factor structure are jointly estimated using an iterative principal components algorithm. Identification and

consistent estimation of the regression coefficients are facilitated by the following assumptions: (i) a common factor structure in the errors v_{it} where each factor makes a nontrivial contribution to the variance of v_{it} [assumption of "strong factors"; see, e.g., Chudik et al. (2011)]; (ii) mild conditions on the idiosyncratic components u_{it} that allow weak correlation and heteroskedasticity in both dimensions; (iii) finite fourth moments for the regressors, factors and loadings; (iv) large sample size in both dimensions. Note, however, that the IFE estimator remains consistent with a small time dimension as long as the errors are serially uncorrelated and there is no time series heteroskedasticity. Cross-section correlation and heteroskedasticity in the errors can still be allowed for without affecting consistency [see Bai (2003) and Bai (2009) for details].

The extension to unbalanced panels was considered in Bai et al. (2015). They allow for various patterns of missing data including block missing where a group of individuals can join or drop out of the sample at a given time period, regular missing where the missing event occurs at the same time frequency for all the individuals, and random missing where some of the data are randomly missing without any obvious pattern. Bai et al. (2015) develop an estimation procedure based on adapting the EM algorithm. Due to space limitations, the details of the procedure are relegated to Appendix A.1. To evaluate the finite sample performance of their estimator, Bai et al. (2015) conduct Monte Carlo simulations under block, regular and random missing patterns in the data generating process. Their proposed estimator is shown to perform well in terms of both bias and variance regardless of whether the common factors are smooth functions of time or stochastic and non-smooth.

3.1.2 The Common Correlated Effects (CCE) Estimator

An alternative approach to estimation proposed by Pesaran (2006) treats the factors as nuisance parameters rather than parameters of interest. In this approach, the common factors f_t are proxied for using the cross-sectional averages of the dependent and independent variables. The regression in (1) is augmented using these cross-sectional averages (with their coefficients allowed to be

individual-specific) and estimated using OLS. In particular, we include the cross-sectional averages of y_{it}, s_{it}, s_{it}^2 and $s_{it}e_{it}$ to proxy for the factors.[6] This estimator, which does not require knowledge of the number of factors, is referred to as the common correlated effects pooled (CCEP) estimator. Pesaran (2006)'s estimator was subsequently extended to unbalanced panels by Zhou and Zhang (2016) who establish consistency and asymptotic normality assuming a large cross-section dimension but a fixed time dimension. The theory is derived assuming that the data are missing at random while their simulation evidence shows that the approach works well more generally (e.g., when the data are only missing at the end).

In addition to the one-step estimator developed in Zhou and Zhang (2016), we also consider a version of the two-step estimator proposed by Pesaran (2006) in the balanced panel case that involves combining the CCE and principal component approaches. The first step entails using the one-step estimator to obtain the estimates of the common factors which are then employed in the second step as regressors instead of the cross-sectional averages to estimate the regression coefficients. The number of factors is set to six in accordance with a rank condition that is required to validate the CCE approach (see discussion below). Appendix A.2 provides details of the one-step and two-step estimators.

The assumptions required for identification in the CCE approach are similar to those previously outlined for the IFE approach except that the CCE approach does not require all the factors to be strong (Chudik et al., 2011) or the time dimension to be large (Zhou and Zhang, 2016; Westerlund et al., 2019). However, both CCE estimators can be sensitive to a particular rank condition which requires that the total number of factors does not exceed the total number of observed variables. Westerlund and Urbain (2013) showed that without this condition, the validity of the CCE approach hinges on the assumption that the factor loadings of the dependent variable and the regressors are uncorrelated. The relevance of this rank condition can explain potential

[6]The issue of including nonlinear functions of regressors (squares, interactions, etc.) in the CCE approach was recently studied by De Vos and Westerlund (2019). They show that the CCE procedure remains valid once we include cross-sectional averages of the nonlinear functions as additional regressors.

discrepancies between the CCE and IFE estimators in practice since the latter does not require this condition to be satisfied.[7]

3.2 Bias Decomposition

Our analysis adopts a bias decomposition proposed by Gelbach (2016) in the context of the linear regression model. The idea is to express the estimated aggregate OLS bias defined as the difference between the estimate from the base specification (one that includes only the variables of interest) and the full specification (that also includes all other covariates) as the sum of the contributions from each of the covariates. The practical advantage of employing this decomposition is that it allows us obtain the estimated bias contribution of each covariate that is invariant to the sequence in which covariates are added to the model. As we show in our empirical analysis (section 5), the bias estimates can differ considerably depending on the order in which the sheepskin and ability components are appended to the base specification that only includes the human capital and experience variables.

For expositional simplicity, we consider a simple linear-in-schooling model that drops the experience terms and defer treatment of the general model to Appendix A.3. The model is specified as

$$y_{it} = c_i + \beta s_{it} + z_{it}'\delta + v_{it} \tag{4}$$

where the error term v_{it} is specified as in (2). As a matter of notation, for any variable w_{it}, define its time demeaned version as $\widetilde{w}_{it} = w_{it} - \bar{w}_i,\ \bar{w}_i = T_i^{-1}\sum_{t\in\mathscr{J}_i} w_{it}$. Further, let

$$\Sigma_{ss} = \sum_{i=1}^{N}\sum_{t\in\mathscr{J}_i} \widetilde{s}_{it}\widetilde{s}_{it}',\ \Sigma_{sz} = \sum_{i=1}^{N}\sum_{t\in\mathscr{J}_i} \widetilde{s}_{it}\widetilde{z}_{it}',\ \Sigma_{sy} = \sum_{i=1}^{N}\sum_{t\in\mathscr{J}_i} \widetilde{s}_{it}\widetilde{y}_{it}$$

[7] Our empirical analysis includes the cross-sectional averages of six variables: $y_{it}, s_{it}, s_{it}^2, s_{it}e_{it}, z_{h,it}, z_{c,it}$. Thus, the rank condition implies $r \leq 6$. The cross-sectional averages of the age controls (our measure of experience) are not included given that they are equivalent to the inclusion of a deterministic time trend (see Appendix B of KLT for details).

We first estimate the following model by OLS:

$$y_{it} = c_i + \beta s_{it} + v_{it} \tag{5}$$

The resulting estimate of β can be expressed as

$$\hat{\beta}_{OLS} = \Sigma_{ss}^{-1}\Sigma_{sy}$$

Next, estimate the following model using one of the FM methods discussed in Section 3.1:

$$y_{it} = c_i + \beta s_{it} + z_{it}'\delta + \lambda_i' f_t + u_{it} \tag{6}$$

Denote the FM estimates of β and δ by $\hat{\beta}_{FM}$ and $\hat{\delta}_{FM}$, respectively, and the estimates of λ_i and f_t by $\hat{\lambda}_i$ and $\hat{f}_t$, respectively. Then, we can write

$$\begin{aligned}\hat{\beta}_{OLS} &= \hat{\beta}_{FM} + \left\{\Sigma_{ss}^{-1}\Sigma_{sz}\hat{\delta}_{FM}\right\} + \left\{\Sigma_{ss}^{-1}\sum_{i=1}^{N}\sum_{t\in\mathscr{J}_i}\widetilde{s}_{it}\hat{\lambda}_i'\hat{f}_t\right\} \\ &= C_1 + C_2 + C_3 \end{aligned} \tag{7}$$

In (7), *the overall returns to schooling*, $\hat{\beta}_{OLS}$, is decomposed into three components: $C_1(=\hat{\beta}_{FM})$ is the human capital component, C_2 is the signaling component and C_3 is the ability bias component. Each of these components is estimable from the data thereby allowing us to quantify their relative contributions to the returns. The decomposition also enables estimating the high school and college sub-components of C_2 separately which is useful in evaluating the individual contributions of high-school and college completion.

3.3 Research Questions

We use our empirical framework to address the following three research questions:

1. What are the relative contributions of the human capital, signaling (sheepskin) and ability bias components to the returns to schooling?

2. What is the role of ability bias in determining the relative contributions of the human capital and signaling components?

3. What is the role of sheepskin effects in determining the relative contributions of the human capital and ability bias components?

The first question can be analyzed directly using (7) as discussed above. For the second, we need to estimate by OLS the following model that excludes the factor structure:

$$y_{it} = c_i + \beta s_{it} + z_{it}'\delta + v_{it} \tag{8}$$

Denote the OLS estimates of β and δ from (8) by $\widetilde{\beta}_{OLS}$ and $\widetilde{\delta}_{OLS}$, respectively. We can write

$$\hat{\beta}_{OLS} = \widetilde{\beta}_{OLS} + \Sigma_{ss}^{-1}\Sigma_{sz}\widetilde{\delta}_{OLS} = D_1 + D_2 \tag{9}$$

A comparison of the decompositions in (7) and (9) reveals the importance of ability bias. Note that in the presence of ability bias, D_1 and D_2 are generally both biased estimates of the true contributions of the human capital and signaling components. For example, if ability bias is an important determinant of returns but is not controlled for so that one relies on the decomposition (9), the contributions of both the human capital and signaling components may be overstated, i.e., $D_1 > C_1,\ D_2 > C_2$. It is also possible that $D_2 > D_1$ (more able individuals are more likely to complete the degree) but $C_1 > C_2$, i.e., the ranking of the human capital and signaling components

in terms of their estimated contributions to returns varies depending on whether ability bias is accounted for. Of course, the direction and magnitude of the biases in D_1 and D_2 depend on the data at hand and will be discussed in our empirical analysis.

Finally, to address the third research question, we first estimate a model that excludes the sheepskin dummies using one of the FM methods:

$$y_{it} = c_i + \beta s_{it} + \lambda_i' f_t + u_{it} \tag{10}$$

Denote the FM estimates of β and δ from (10) by $\widetilde{\beta}_{FM}$ and $\widetilde{\delta}_{FM}$, respectively. We can express the OLS estimate $\hat{\beta}_{OLS}$ from (5) as

$$\hat{\beta}_{OLS} = \widetilde{\beta}_{FM} + \left\{ \Sigma_{ss}^{-1} \sum_{i=1}^{N} \sum_{t \in \mathscr{J}_i} \widetilde{s}_{it} \widetilde{\lambda}_i' \widetilde{f}_t \right\} = G_1 + G_3 \tag{11}$$

where $\{\widetilde{\lambda}_i\}$ and $\{\widetilde{f}_t\}$ denote the estimates of the factor loadings and common factors, respectively, from (10). Now, a comparison of the decompositions in (7) and (11) clarifies the empirical relevance of sheepskin effects. If sheepskin effects are important, G_1 and G_3 can be substantially biased and therefore quite different from C_1 and C_3, respectively.

4 Data

We use the National Longitudinal Survey of Youth (NLSY79) for our empirical analysis. The information collected by the NLSY79 includes schooling choices, labor market outcomes, family background and individual characteristics, for a nationally representative sample of 12,686 young men and women who were 14-22 years old in 1979. These individuals were interviewed yearly through 1994 and every other year thereafter. We construct an unbalanced panel dataset with hourly earnings, years of schooling, degrees information, and other covariates from 1981 to 2016, with the minimum number of years restricted to be 15, and the maximum number of years as 25.

To remain consistent with the literature, we restrict the sample of analysis to (i) white males who are at least 16 years old in a given year to analyze a population that historically is strongly attached to the labor market, and least likely to experience wage discrimination (Ginther, 2000), and (ii) individuals who reported working at least 30 weeks and 800 hours in the previous year to capture an individual's full earning potential (Koop and Tobias, 2004). We carefully remove observations with missing information on degrees and other covariates, outlier observations in the top and bottom 1% of the wage distribution, and individuals with negative or abnormally large schooling changes. Our final sample contains $1,176$ individuals for a total of $23,779$ person-year observations, where observations within an individual are not required to be consecutive.

Table 2 presents descriptive statistics for not only the main sample of analysis but also how the summary statistics change as we sequentially add the selection criteria to obtain our final sample.[8] Column (1) shows a baseline sample of 9,224 individuals who are at least 16 years old in a given year with no missing covariates. Focusing our attention on white males decreases the sample by more than two-thirds to 2,954 individuals in column (2). It is worth noting that, hourly wage earning is 24% higher on average for white males compared to the baseline sample. Column (3) imposes wages and hours worked restrictions, but only shrinks the sample by 111 individuals, presumably due to the aforementioned reason that white males above age 16 are strongly attached to the labor market. After trimming the wage outliers in Column (4), we lose another person, and the number of observations decreases to 37,452. The average wage converted to 1999 dollars drops to about 16 dollars an hour, and the standard deviation is greatly reduced to a reasonable level. Column (5) displays summary statistics for a sample of $2,420$ individuals for a total of $33,248$ person-year observations, after we carefully exclude individuals with clearly misreported schooling changes. We further restrict the minimum number of years to be 15 in column (6) to mitigate the concern of small T_i in an unbalanced panel, and as a result, the sample shrinks to

[8]The NLSY sampling weights were not used following Heckman et al. (2006); Koop and Tobias (2004); among others. We also did not drop the oversample observations.

1,190 individuals for a total of 24,046 person-year observations.[9] Lastly, column (7) presents our final sample of analysis, confirming that there is no missing information concerning the timing of degree acquisition, including both high school and college degrees. Consequently, our final sample is reduced by 14 individuals, resulting in a total of 1,176 individuals and 33,248 person-year observations. The summary descriptions for the final sample in column (7) closely align with those in column (6), the latter of which was employed in Kejriwal et al. (2024). The average wage is approximately $16.66 per hour in 1999 dollars, with an average educational attainment (highest grade completed) of 13.28 years, slightly more than a high school diploma. In this final sample, 94% of the person-year observations have a high school diploma, and 24% have a college degree.

Delving deeper into the variations of schooling and degree variables provides insightful details. Table B.1 shows detailed information on the variation in years of schooling for the final sample. Panel A presents the number of schooling changes per person. We follow Koop and Tobias (2004) and keep all individuals including 675 people with fixed schooling. About 43% individuals in our final sample experience changes of schooling, mostly once or twice during the sample period. Panel B tabulates the schooling changes by education level after the change, and most of the variation in schooling in our final sample comes from individuals completing grades years 12-16 (i.e., completing high school through 4 years of college). Table B.2 presents a detailed cross-tabulation between years of schooling and degree attainment. Similar to Flores-Lagunes and Light (2010), we also observe significant variations in years of schooling within each degree category. For instance, while over 70% of individuals earned a high school degree with 12 years of education, variations exist: some completed it in less than 12 years, while others took more—notably, 11.7% with 13 years and 12.7% with 14 years of schooling. Moreover, among individuals with identical years of schooling, there is some variation in degree status, albeit to a lesser extent. For instance, while the majority with 16 years of education possess a college degree, a few only hold a high school diploma. These variations are crucial for separately identifying the

[9]Figure B.1 shows the distribution of number of years for our final analysis sample of 1,176 individuals.

schooling and degree effects (Jaeger and Page, 1996).

Although there are several studies that use panel data and within individual variation to study returns to schooling (Angrist and Newey, 1991; Ashworth et al., 2021; Kejriwal et al., 2020, 2024; Koop and Tobias, 2004; Westerlund and Petrova, 2018), one important concern for this study is how sample selection affects the magnitude of the signaling effects of diplomas.[10] The literature on statistical discrimination suggests that sheepskin effects—where diplomas act as signals of productivity—may not be uniform across demographic groups. For instance, results from Belman and Heywood (1991) indicate that minority men and women receive higher returns from completing 16 or more years of education compared to their white counterparts. Similarly, the sheepskin effects could vary if diplomas provide different productivity signals for those individuals who are employed while pursuing their education. We believe this concern is mitigated by the facts that: (1) Other research has shown that the student population that works during school is large and growing (Bacolod and Hotz, 2006; Bound et al., 2012; Carnevale et al., 2015; Hotz et al., 2002), and is thus an important population itself; (2) Most observable covariates in Table 2 remain fairly stable across columns (2)-(7) and are not sensitive to additional restrictions imposed (e.g., minimum weeks and hours worked, minimum number of years to be 15); (3) OLS estimates for high school and college degree effects align closely with studies in the literature that do not incorporate panel data; (4) If we consider a reasonable possibility that the signaling effects of diplomas may be weaker for working students compared to students without accumulated work experience, then the actual diploma signaling effects could be stronger than our estimates suggest.

5 Empirical Results

This section presents the results of our empirical analysis. Section 5.1 details the set of estimated specifications. Section 5.2 presents estimates of the $MRTS$ from the different specifications.

[10] See Kejriwal et al. (2020, 2024) for detailed discussions on sample selection concerns related to part-time or seasonal work, and measurement errors in schooling variations.

Section 5.3 reports the bias decomposition results that facilitate estimation of the relative contributions of the different components (human capital, signaling, ability bias) to the apparent returns to schooling. Our empirical analysis is aimed at addressing the research questions discussed in section 3.3.

5.1 Estimated Specifications

We estimate eight groups of specifications that vary according to whether the model is linear or quadratic in schooling as well as whether degree indicators, interactive fixed effects, and/or demographics are included. A summary of the specifications is presented in Table 2. Panel A contains linear schooling specifications. The specifications in Group 1 exclude the demographic controls and interactive fixed effects: specification (a) incorporates only the schooling variable alongside basic controls, namely, person fixed effects and quadratic age controls; specification (b) adds indicators for the completion of high school and college degrees. The specifications in Group 2 are the counterparts of those in Group 1 that include demographic controls. Group 3 includes specifications that mirror those in Group 1 while allowing for interactive fixed effects. Finally, the specifications in Group 4 include both demographic controls and interactive fixed effects.

The specifications in Panel B are identical to those in Panel A, except that they involve a quadratic function of schooling. We also estimate another set of specifications paralleling those detailed in Table 2 which utilize quartic instead of quadratic age controls. For brevity, we only report the parameter estimates with quadratic age controls and refer to Appendix C for the set of results with quartic age controls which are qualitatively similar.

5.2 Parameter Estimates

Tables 3 and 4 present estimates of the MRTS (in percentage form) evaluated at the overall mean (across individuals and time) as well as the coefficients associated with the degree completion indicators, for the linear and quadratic schooling specifications, respectively. Panel A reports the

results without demographic controls, while Panel B displays the estimates with these controls which include parental education levels and the number of siblings. In addition to the parameter estimates, we also present two measures of model adequacy: the adjusted R^2 for each specification and the CD statistic based on the OLS residuals proposed by Pesaran (2015) for testing the null hypothesis of no cross-section dependence.

Several aspects of these findings are noteworthy. First, as shown in column (1) of Tables 3 and 4, the OLS estimates of the MRTS without degree controls are between 7%-8.5%, consistent with the typical range (6%-10%) reported in the returns to education literature [see, e.g., the survey by Gunderson and Oreopolous, 2020]. Second, once degree completion is accounted for [column (2)], the MRTS estimates decrease to approximately 3%-4%. When examining the degree coefficients, the OLS return on a high school degree is positive but its magnitude varies considerably depending on the model specification. For instance, in the linear schooling model (Table 3), the return on a high school degree is approximately 6%, with marginal statistical significance at the 10% level. However, under a quadratic schooling model (Table 4), the estimate drops to about 3% and becomes statistically insignificant.[11] The OLS return to a college degree is substantially higher, at around 23%-24%, and statistically significant at the 1% level. This notable college premium remains robust across the various model specifications, demonstrating a strong and stable increase in earnings that a college degree provides, conditional on the same years of schooling.[12] It is useful to note that the estimated sheepskin effect for a college degree closely aligns with those reported in the literature that utilizes cross-sectional data, as detailed in Caplan (2018), Tables E.1 and E.2. However, the effect for a high school degree falls at the lower end of the range reported in the

[11] In Appendix C, we examine the robustness of our findings using quartic (as opposed to quadratic) age controls (see, e.g., Cho and Phillips, 2018). The only notable difference is that the OLS return to a high school degree is about 3% and 1% for the linear and quadratic schooling models, respectively - both statistically insignificant, as detailed in column (2) of Tables C.1 and C.2.

[12] Due to limitations in data variability, we do not distinguish between advanced degrees and college degrees; therefore, individuals with advanced degrees are also categorized under college degrees in our analysis. This classification may artificially elevate the estimated returns to a college degree, although only a small percentage (6.7%) of the individuals hold graduate degrees. Excluding those individuals with graduate degrees from our analysis does not alter the results, which are available upon request.

literature, where the year twelve premium varies from 2.6% in Flores-Lagunes and Light (2010) to 27% in Riddell (2008), as summarized in the same tables by Caplan (2018). The evidence thus suggests that the signaling effect of a high school diploma may be less pronounced for working students than those without accumulated work experience.

Third, without controlling for degree completion, the FM estimates of the MRTS decrease to about 3%-5% [columns (3), (5), and (7)], regardless of the particular FM estimator adopted. Thus, after accounting for ability bias via an interactive fixed effects structure markedly reduces the overall returns to schooling, more than halving the OLS estimates in column (1), consistent with previous findings in Kejriwal et al. (2020, 2024). Fourth, and most importantly, including indicators for degree completion [Columns (4), (6), (8)] shrinks the FM estimates to nearly zero with statistical insignificance across all specifications, with this finding being insensitive to the choice of the FM estimator used to control for the factor structure. When revisiting the degree coefficients, the FM return on a high school degree is mostly statistically indistinguishable from zero, aligning with Clark and Martorell (2014)'s finding that a high school diploma offers little signaling value. However, the magnitude of this effect varies depending on the estimator and model specification used. For example, in linear schooling models (Table 3), the results consistently demonstrate both statistically and economically insignificant effects of a high school degree across the various estimators and specifications. When using quadratic schooling models, the IFE estimator without demographic controls [Table 4, column (4)] yields a relatively large effect (around 7.5%, statistically significant at the 10% level) compared to other estimators and specifications (about 2%-3%, statistically insignificant). The FM returns on a college degree are reduced compared to the OLS estimates in column (2) but remain substantial and robust, with values ranging from 14%-17% that consistently maintain statistical significance at the 1% level. This enduring premium is consistent with prior research, which documents a substantive sheepskin effect for college degrees even after adjusting for ability proxies, as evidenced, inter alia, in Frazis (1993); Kane and Rouse (1995); Arkes (1999); Light and Strayer (2004); Tobias and Li (2004).

Fourth, both the adjusted R^2 and CD statistics offer strong evidence in favor of the factor structure: for both linear and quadratic schooling models, the adjusted R^2 notably improves with the inclusion of the factor structure while the CD test statistic comprehensively rejects the null of no cross-section dependence. Finally, it is worth noting that the results with and without demographic controls are qualitatively similar (except for the return to a high school degree using the IFE estimator without demographic controls as noted above) and indicate that the factor structure cannot be interpreted in terms of time-varying returns to observable time-invariant characteristics.

5.3 Bias Decomposition

Table 5 details the bias decomposition results without the factor structure, establishing a baseline that separates the overall return to schooling into the human capital component (D1) and the diploma signaling component (D2), as outlined in equation (9), both of which are subject to bias when ability controls are absent. This baseline is essential for subsequent comparisons with the scenario where ability controls are included to assess the role of ability bias in determining the relative contributions of the human capital and signaling components (research question 2 in section 3.3). Columns (1) and (3) present linear and quadratic schooling models with quadratic age controls, respectively, while columns (2) and (4) add quartic age controls to these schooling specifications. Panel A reports the findings without demographic controls while Panel B includes them.

The results reveal several key insights. First, without incorporating the factor structure, both the human capital (D1) and the diploma signaling components (D2) positively and significantly contribute to the overall returns to schooling. This finding highlights the inherent benefits of educational attainment in terms of both productivity enhancement and perceived market value when ability controls are not included. Moreover, the overall contribution from diplomas is predominantly attributed to the college degree. In contrast, the high school degree's contribution to schooling returns is neither statistically nor economically significant. It is noteworthy that the

results vary only modestly across model specifications. Specifically, when schooling enters the model linearly, the contribution from the sheepskin channel appears slightly more pronounced at about 60% relative to a human capital contribution of around 40%, as observed in columns (1) and (2). However, when schooling enters quadratically, the contributions of human capital and degree completion tend to equalize. The data suggest a nearly even 50/50 split between human capital and diploma signaling, based on the results in columns (3) and (4). These splits remain consistent whether demographic controls are included or not and are unaffected by the modeling choice for the age controls.

Table 6 displays the bias decomposition results incorporating the factor structure, where the overall returns to schooling are split into the human capital component (C1), the diploma signaling component (C2), and the ability bias component (C3) [see equation (7)]. The structure of Table 6 mirrors that of Table 5, wherein each model setup - whether schooling is linear or quadratic, paired with quadratic or quartic age controls - features results for the three different estimators, namely, IFE, CCE, and CCE-2.

Correcting for ability bias yields rather different results regarding the contributions of human capital and signaling to the returns on education. First and foremost, ability bias (C3) emerges as the predominant factor among the three in explaining the apparent returns to schooling. The contribution of the human capital component (C1) is minimal, delineating the limited direct impact of schooling on productivity enhancement after controlling for ability bias. This finding suggests that much of the human capital effects documented in previous studies may be overstated and attributed to self-selection. In contrast, the contribution from degree attainment remains substantial, and is primarily associated with college degrees. In other words, contrary to the skepticism often expressed in the literature, the diploma signaling component, while somewhat diminished, remains important after accounting for the possibility that more capable individuals are more likely to complete the degrees. To provide some context, for specifications without demographic controls (Table 6, Panel A), ability bias make up the largest portion of the overall returns to

schooling, followed by the sheepskin component, which together drive most of the apparent returns to schooling. On average, across all columns, the distribution between human capital, diploma signaling, and ability is represented as a 6/38/56 split.

It is useful to compare our results in Table 6 with those from a sequential covariate addition approach in order to highlight the practical relevance of our bias decomposition analysis. There are two possible sequences in which covariates can be appended to the base specification that only includes the human capital and experience variables (and demographic controls if applicable). One sequence, which we label the HC-SS-AB approach, first adds the degree indicators followed by the factor structure. The second, labeled the HC-AB-SS approach, first adds the factor structure followed by the degree indicators. For the HC-SS-AB approach, we can write

$$\begin{aligned} MRTS_{ols}^{base} - MRTS_{fm}^{full} &= \left\{MRTS_{ols}^{base} - MRTS_{ols}^{ss}\right\} + \left\{MRTS_{ols}^{ss} - MRTS_{fm}^{full}\right\} \\ &= C_2^* + C_3^*, \end{aligned} \tag{12}$$

where $MRTS_{ols}^{base}$ is the OLS estimate from the base specification, $MRTS_{ols}^{ss}$ is the OLS estimate from the specification that adds the degree indicators to the base, and $MRTS_{fm}^{full}$ is the FM estimate (based on one of the three estimators) from the full specification that includes both the degree indicators and the factor structure. These estimates can be obtained, respectively, from column (1), column (2), and columns (4), (6), (8) of Tables 3 and 4. The quantities C_2^* and C_3^* are the estimated contributions of the signaling and ability bias components, respectively, obtained by employing the HC-SS-AB approach. We can easily compute C_2^* and C_3^* from Tables 3 and 4. For instance, consider the quadratic schooling specification that employs quadratic age controls but excludes demographic controls. Then $C_2^*/MRTS_{ols}^{base} = 0.0424/0.0834 \simeq 51\%$. If the factor structure is controlled for using the IFE estimator, $C_3^*/MRTS_{ols}^{base} = 0.0294/0.0834 \simeq 35\%$. In comparison, $C_2/MRTS_{ols}^{base} \simeq 39\%$, $C_3/MRTS_{ols}^{base} \simeq 47\%$. Thus, the HC-SS-AB approach tends to exaggerate the contribution of the signaling component while understating the contribution of the

ability bias component. An analogous pattern holds for the other specifications and FM estimators. A similar analysis for the HC-AB-SS approach shows that the relative importance of the signaling and ability bias components depends on the specific FM estimator employed. In particular, the IFE approach overstates (understates) the ability bias (signaling) contribution while the opposite is true for the CCE and CCE-2 approaches. Overall, these results clearly illustrate the fragility of the sequential covariate addition approach in the present context and underscore the robustness afforded by our adopted approach. This completes our analysis of the first research question concerning the relative importance of human capital, signaling, and ability bias in determining the overall returns to schooling.

Next, we analyze the second research question concerning the role of ability bias. Comparisons between models with and without ability controls reveal that the human capital component is substantively reduced after adjusting for ability bias (i.e., C1 is much smaller than D1). On the other hand, the sheepskin component also diminishes, though to a lesser extent (i.e., C2 is slightly smaller than D2). Consequently, the distribution between human capital and signaling in explaining the returns to schooling shifts towards a 13/87 split in favor of signaling. These findings highlight the crucial role of accounting for ability bias in accurately assessing the relative contributions of human capital and the sheepskin effect to the returns on education.

Finally, to address the third research question pertaining to the role of sheepskin effects, we compare the results between models that include and exclude the indicators for degree completion that characterize the signaling component of returns. Specifically, we can write

$$\begin{aligned} MRTS_{ols}^{base} &= MRTS_{FM}^{nd} + \left\{ MRTS_{ols}^{base} - MRTS_{FM}^{nd} \right\} \\ &= G_1 + G_3, \end{aligned} \tag{13}$$

where $MRTS_{ols}^{base}$ is as defined in (12), $MRTS_{FM}^{nd}$ is the FM estimate when the degree indicators are excluded, and G_1, G_3 are defined in (11). The quantities G_1 and G_3 can be computed from

columns (1), (3), (5), (7) of Tables 3 and 4. A comparison of G_1 and G_3 with C_1 and C_3 (reported in Table 6) indicates that omitting sheepskin effects considerably overstates the human capital component (C_1 is much smaller than G_1) and understates the ability bias component (C_3 is much larger than G_3) regardless of the adopted specification and the FM estimator employed. For example, when considering the IFE estimator in conjunction with a specification that is quadratic in schooling as well as age and excludes demographic controls [Table 6, Panel A, column (7)], $C_1/(C_1+C_3) \simeq 23\%$ and $C_3(C_1+C_3) \simeq 77\%$. In contrast, from columns (1) and (3) of Panel A in Table 3, $G_1/(G_1+G_3) = .0391/0.0834 \simeq 47\%$ and $G_2/(G_1+G_3) = 0.0443/0.0834 \simeq 53\%$. This analysis therefore serves to showcase the empirical importance of controlling for sheepskin effects when estimating the relative contributions of human capital and ability bias to the overall schooling returns.

6 Conclusion

The financial return to education can be ascribed to two distinct sources: first, the prestige and opportunities associated with having a formal degree, and second, the tangible skills and knowledge acquired during the educational process. This study contributes significantly to the ongoing debate about the relative importance of human capital and signaling in explaining the returns to education. Our factor model-based analysis of the NLSY79 panel variation challenges traditional views and suggests that the observed returns to schooling are predominantly driven by ability bias, accounting for about 56% of these returns, while actual benefits from degree attainment and human capital accumulation are less pronounced than previously thought. Moreover, the sheepskin effect is predominantly influenced by college degrees, whereas high school degrees contribute minimally to this aspect. Importantly, without adjusting for ability, the distribution between Human Capital and Diploma Signaling is more evenly split at 40/60. However, once multidimensional ability is accounted for, this distribution shifts dramatically to 13/87, heavily favoring signaling.

These findings underscore the profound impact of ability bias, which significantly diminishes the perceived direct contributions of schooling itself, adding crucial insights to this debate.

Our methodology, which avoids reliance on inadequate ability proxies and does not depend on the sequence of covariate inclusion, offers a more reliable estimation of the true returns to education. By demonstrating the predominance of ability bias in the apparent returns to schooling, this research invites policymakers and scholars to reconsider the economic value attributed to educational programs and credentials as well as the mechanisms through which education impacts economic outcomes. This reevaluation is crucial for designing education policies that genuinely enhance productivity and economic opportunity.

References

Altonji, J. G. and C. R. Pierret (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics 116*(1), 313–350.

Angrist, J. D. and W. K. Newey (1991). Over-identification tests in earnings functions with fixed effects. *Journal of Business & Economic Statistics 9*(3), 317–323.

Arkes, J. (1999). What do educational credentials signal and why do employers value credentials? *Economics of Education Review 18*(1), 133–141.

Arteaga, C. (2018). The effect of human capital on earnings: Evidence from a reform at colombia's top university. *Journal of Public Economics 157*, 212–225.

Aryal, G., M. Bhuller, and F. Lange (2022). Signaling and employer learning with instruments. *American Economic Review 112*(5), 1669–1702.

Ashworth, J., V. J. Hotz, A. Maurel, and T. Ransom (2021). Changes across cohorts in wage returns to schooling and early work experiences. *Journal of labor economics 39*(4), 931–964.

Bacolod, M. and V. J. Hotz (2006). Cohort changes in the transition from school to work: Evidence from three nls surveys. *Economics of Education Review 25*(4), 351–373.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica 71*(1), 135–171.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica 77*(4), 1229–1279.

Bai, J., Y. Liao, and J. Yang (2015). Unbalanced panel data models with interactive effects. In *The Oxford Handbook of Panel Data*. Oxford University Press.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of political economy 70*(5, Part 2), 9–49.

Bedard, K. (2001). Human capital versus signaling models: university access and high school dropouts. *Journal of political economy 109*(4), 749–775.

Belman, D. and J. S. Heywood (1991). Sheepskin effects in the returns to education: An examination of women and minorities. *The Review of Economics and Statistics*, 720–724.

Bound, J. and D. A. Jaeger (2000). Do compulsory school attendance laws alone explain the association between quarter of birth and earnings? In *Research in labor economics*, pp. 83–108. Emerald Group Publishing Limited.

Bound, J., D. A. Jaeger, and R. M. Baker (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American statistical association 90*(430), 443–450.

Bound, J., M. F. Lovenheim, and S. Turner (2012). Increasing time to baccalaureate degree in the united states. *Education Finance and Policy 7*(4), 375–424.

Bradley, S. and C. Green (2020). The economics of education: A comprehensive overview.

Buckles, K. S. and D. M. Hungerman (2013). Season of birth and later outcomes: Old questions, new answers. *Review of Economics and Statistics 95*(3), 711–724.

Caplan, B. (2018). *The case against education: Why the education system is a waste of time and money*. Princeton University Press.

Caplan, B. (2019). The case against education: Why the education system is a waste of time and money.

Card, D. (1995). Earnings, schooling, and ability revisited. In S. Polschek (Ed.), *Research in Labor Economics*, pp. 23–48. JAI Press.

Carnevale, A. P., N. Smith, M. Melton, E. Price, et al. (2015). Learning while earning: The new normal. *Georgetown University Center on Education and the Workforce*.

Chiswick, B. (1973). Schooling, screening, and income. *Does college matter*, 151–158.

Cho, J. S. and P. C. Phillips (2018). Sequentially testing polynomial model hypotheses using power transforms of regressors. *Journal of Applied Econometrics 33*(1), 141–159.

Chudik, A., M. H. Pesaran, and E. Tosetti (2011). Weak and strong cross-section dependence and estimation of large panels. *The Econometrics Journal 14*(1), C45–C90.

Clark, D. and P. Martorell (2014). The signaling value of a high school diploma. *Journal of Political Economy 122*(2), 282–318.

Das, T. and S. W. Polachek (2019). *Microfoundations of earnings differences*. Springer.

De Vos, I. and J. Westerlund (2019). On CCE estimation of factor-augmented models when regressors are not linear in the factors. *Economics Letters 178*, 5–7.

Deming, D. J. (2022). Four facts about human capital. *Journal of Economic Perspectives 36*(3), 75–102.

Fang, H. (2006). Disentangling the college wage premium: Estimating a model with endogenous education choices. *International Economic Review 47*(4), 1151–1185.

Farber, H. S. and R. Gibbons (1996). Learning and wage dynamics. *The Quarterly Journal of Economics 111*(4), 1007–1047.

Flores-Lagunes, A. and A. Light (2010). Interpreting degree effects in the returns to education. *Journal of Human Resources 45*(2), 439–467.

Frazis, H. (1993). Selection bias and the degree effect. *Journal of Human Resources*, 538–554.

Frazis, H. (2002). Human capital, signaling, and the pattern of returns to education. *Oxford Economic Papers 54*(2), 298–320.

Gelbach, J. B. (2016). When do covariates matter? and which ones, and how much? *Journal of Labor Economics 34*(2), 509–543.

Ginther, D. K. (2000). Alternative estimates of the effect of schooling on earnings. *Review of Economics and Statistics 82*(1), 103–116.

Gunderson, M. and P. Oreopolous (2020). Returns to education in developed countries. In *The economics of education*, pp. 39–51. Elsevier.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2006). Earnings functions, rates of return and treatment effects: The mincer equation and beyond. *Handbook of the Economics of Education 1*, 307–458.

Heckman, J. J., J. Stixrud, and S. Urzua (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics 24*(3), 411–482.

Hotz, V. J., L. C. Xu, M. Tienda, and A. Ahituv (2002). Are there returns to the wages of young men from working while in school? *Review of Economics and statistics 84*(2), 221–236.

Hungerford, T. and G. Solon (1987). Sheepskin effects in the returns to education. *The review of economics and statistics*, 175–177.

Jaeger, D. A. and M. E. Page (1996). Degrees matter: New evidence on sheepskin effects in the returns to education. *The review of economics and statistics*, 733–740.

Kane, T. J. and C. E. Rouse (1995). Labor-market returns to two-and four-year college. *The American Economic Review 85*(3), 600–614.

Kejriwal, M., X. Li, L. Nguyen, and E. Totty (2024). The efficacy of ability proxies for estimating the returns to schooling: A factor model-based evaluation. *Journal of Applied Econometrics 39*(1), 3–21.

Kejriwal, M., X. Li, and E. 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*(5), 548–566.

Koop, G. and J. L. Tobias (2004). Learning about heterogeneity in returns to schooling. *Journal of Applied Econometrics 19*(7), 827–849.

Krueger, A. B. and M. Lindahl (2001). Education for growth: Why and for whom? *Journal of economic literature 39*(4), 1101–1136.

Lange, F. (2007). The speed of employer learning. *Journal of Labor Economics 25*(1), 1–35.

Lange, F. and R. Topel (2006). The social value of education and human capital. *Handbook of the Economics of Education 1*, 459–509.

Lazear, E. P. (1977). Schooling as a wage depressant. *Journal of Human Resources*, 164–176.

Lemieux, T. (2006). The "mincer equation" thirty years after schooling, experience, and earnings. In *Jacob Mincer a pioneer of modern labor economics*, pp. 127–145. Springer.

Light, A. and W. Strayer (2004). Who receives the college wage premium?: Assessing the labor market returns to degrees and college transfer patterns. *Journal of Human Resources 39*(3), 746–773.

Lovenheim, M. and S. Turner (2017). *Economics of Education*. Worth Publishers.

Mincer, J. (1996). Changes in wage inequality, 1970-1990. Technical report, National Bureau of Economic Research.

Murphy, K. M. and F. Welch (1990). Empirical age-earnings profiles. *Journal of Labor economics 8*(2), 202–229.

Park, J. H. (1999). Estimation of sheepskin effects using the old and the new measures of educational attainment in the current population survey. *Economics Letters 62*(2), 237–240.

Pasche, C. (2009). A multiple ability approach to employer learning. *University of Geneva, Economics of Education Project. http://www. educationeconomics. unige. ch/Projets/Competences/Pasche_09. pdf*.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica 74*(4), 967–1012.

Pesaran, M. H. (2015). Testing weak cross-sectional dependence in large panels. *Econometric reviews 34*(6-10), 1089–1117.

Riddell, C. (2008). Understanding 'sheepskin effects' in the returns to education: The role of cognitive skills. *Canada: Department of Economics, University of Toronto*.

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics 87*, 354–374.

Tobias, J. L. and M. Li (2004). Returns to schooling and bayesian model averaging: A union of two literatures. *Journal of Economic Surveys 18*(2), 153–180.

Weiss, A. (1995). Human capital vs. signalling explanations of wages. *Journal of Economic perspectives 9*(4), 133–154.

Westerlund, J. and Y. Petrova (2018). Asymptotic collinearity in cce estimation of interactive effects models. *Economic Modelling 70*, 331–337.

Westerlund, J., Y. Petrova, and M. Norkute (2019). CCE in fixed-T panels. *Journal of Applied Econometrics 34*(5), 746–761.

Westerlund, J. and J.-P. Urbain (2013). On the estimation and inference in factor-augmented panel regressions with correlated loadings. *Economics Letters 119*(3), 247–250.

Zhou, Q. and Y. Zhang (2016). Common correlated effects estimation of unbalanced panel data models with cross-sectional dependence. *Journal of Economic Theory and Econometrics 27*(4), 25–45.

Table 1: Summary of Estimated Specifications

Group	Specification	Controls (beyond the base)	Estimator
Panel A: Linear Schooling			
1	(a) $y_{it} = c_i + s_{it}\beta_1 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + v_{it}$	-	OLS
	(b) $y_{it} = c_i + s_{it}\beta_1 + HS_{it}\beta_2 + COLL_{it}\beta_3 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + v_{it}$	degrees	OLS
2	(a) $y_{it} = c_i + s_{it}\beta_1 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + w_i'\phi_t + v_{it}$	demographics	OLS
	(b) $y_{it} = c_i + s_{it}\beta_1 + HS_{it}\beta_2 + COLL_{it}\beta_3 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + w_i'\phi_t + v_{it}$	degrees, demographics	OLS
3	(a) $y_{it} = c_i + s_{it}\beta_1 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + \lambda_i' f_t + u_{it}$	interactive fixed effects	IFE, CCE, CCE-2
	(b) $y_{it} = c_i + s_{it}\beta_1 + HS_{it}\beta_2 + COLL_{it}\beta_3 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + \lambda_i' f_t + u_{it}$	degrees, interactive fixed effects	IFE, CCE, CCE-2
4	(a) $y_{it} = c_i + s_{it}\beta_1 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + w_i'\phi_t + \lambda_i' f_t + u_{it}$	demographics, interactive fixed effects	IFE, CCE, CCE-2
	(b) $y_{it} = c_i + s_{it}\beta_1 + HS_{it}\beta_2 + COLL_{it}\beta_3 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + w_i'\phi_t + \lambda_i' f_t + u_{it}$	degrees, demographics, interactive fixed effects	IFE, CCE, CCE-2
Panel B: Quadratic Schooling			
5	(a) $y_{it} = c_i + s_{it}\beta_1 + s_{it}^2\beta_2 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + v_{it}$	-	OLS
	(b) $y_{it} = c_i + s_{it}\beta_1 + s_{it}^2\beta_2 + HS_{it}\beta_3 + COLL_{it}\beta_4 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + v_{it}$	degrees	OLS
6	(a) $y_{it} = c_i + s_{it}\beta_1 + s_{it}^2\beta_2 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + w_i'\phi_t + v_{it}$	demographics	OLS
	(b) $y_{it} = c_i + s_{it}\beta_1 + s_{it}^2\beta_2 + HS_{it}\beta_3 + COLL_{it}\beta_4 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + w_i'\phi_t + v_{it}$	degrees, demographics	OLS
7	(a) $y_{it} = c_i + s_{it}\beta_1 + s_{it}^2\beta_2 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + \lambda_i' f_t + u_{it}$	interactive fixed effects	IFE, CCE, CCE-2
	(b) $y_{it} = c_i + s_{it}\beta_1 + s_{it}^2\beta_2 + HS_{it}\beta_3 + COLL_{it}\beta_4 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + \lambda_i' f_t + u_{it}$	degrees, interactive fixed effects	IFE, CCE, CCE-2
8	(a) $y_{it} = c_i + s_{it}\beta_1 + s_{it}^2\beta_2 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + w_i'\phi_t + \lambda_i' f_t + u_{it}$	demographics, interactive fixed effects	IFE, CCE, CCE-2
	(b) $y_{it} = c_i + s_{it}\beta_1 + s_{it}^2\beta_2 + HS_{it}\beta_3 + COLL_{it}\beta_4 + e_{it}\gamma_1 + e_{it}^2\gamma_2 + s_{it}e_{it}\gamma_3 + w_i'\phi_t + \lambda_i' f_t + u_{it}$	degrees, demographics, interactive fixed effects	IFE, CCE, CCE-2

Note: We estimate another set of the above specifications where quartic age controls are employed as the baseline: $e_{it}\gamma_1 + e_{it}^2\gamma_2 + e_{it}^4\gamma_4 + s_{it}e_{it}\gamma_4$.

Table 2: Summary Statistics

	(1) No missing variables and age 16+	(2) White males	(3) Weeks and hours worked restrictions	(4) No outlier wages	(5) No abnormal schooling changes	(6) $min(T_i) = 15$	(7) No missing degree variables
Hourly wage inflation-adjusted	18.80	23.28	24.00	15.94	15.82	16.65	16.66
	(349.95)	(363.95)	(349.57)	(10.18)	(9.99)	(10.09)	(10.11)
Log hourly wage inflation-adjusted	2.37	2.56	2.64	2.60	2.60	2.66	2.66
	(0.74)	(0.76)	(0.73)	(0.58)	(0.57)	(0.56)	(0.56)
Years of school	13.19	13.29	13.39	13.34	13.21	13.29	13.28
	(2.36)	(2.48)	(2.47)	(2.44)	(2.28)	(2.15)	(2.15)
Age	32.66	32.09	33.05	32.78	32.82	34.31	34.33
	(10.20)	(10.11)	(9.92)	(9.79)	(9.78)	(10.05)	(10.06)
Mother's years of school	11.13	11.95	11.96	11.93	11.90	11.97	11.97
	(3.11)	(2.43)	(2.40)	(2.38)	(2.33)	(2.22)	(2.22)
Father's years of school	11.11	12.17	12.18	12.14	12.07	12.19	12.19
	(3.90)	(3.41)	(3.36)	(3.34)	(3.28)	(3.17)	(3.16)
Number of siblings	3.60	3.02	2.99	3.00	3.01	2.94	2.94
	(2.50)	(2.00)	(1.98)	(1.98)	(1.99)	(1.91)	(1.90)
Has a High school degree							0.94
							(0.23)
Has a College degree							0.24
							(0.43)
Number of Individuals (N)	9,224	2,954	2,843	2,842	2,420	1,190	1,176
Number of Observations ($N*T$)	145,473	45,291	38,457	37,452	33,248	24,046	23,779

Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.
Note: Each column reports averages and standard deviations (in parentheses) for the sample specified in the column. Column (1) starts with a basis sample. Columns (2)-(7) sequentially add additional sample criteria until the final sample is shown in column (7). Hourly earnings are adjusted for inflation to 1999 dollars.

Table 3: Estimates of the Return to Schooling - Linear Schooling (Quadratic Age)

	(1) OLS	(2) OLS	(3) IFE	(4) IFE	(5) CCE	(6) CCE	(7) CCE-2	(8) CCE-2
Panel A: Without Demographic Controls								
Marginal Return to Schooling	0.0760	0.0302	0.0460	0.0050	0.0437	0.0009	0.0430	0.0016
	(0.0097)	(0.0122)	(0.0054)	(0.0119)	(0.0126)	(0.0165)	(0.0115)	(0.0146)
High School Degree		0.0581		0.0034		0.0073		0.0023
		(0.0352)		(0.0423)		(0.0504)		(0.0471)
College Degree		0.2411		0.1664		0.1676		0.1700
		(0.0384)		(0.0352)		(0.0401)		(0.0343)
Demos-by-year FE	No	No	No	No	No	No	No	No
Degree Indicators	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R-squared	0.6214	0.6237	0.8704	0.8527	0.7737	0.8151	0.8120	0.8122
CD statistic	25.16	22.966						
Panel B: With Demographic Controls								
Marginal Return to Schooling	0.0715	0.0286	0.0383	-0.0040	0.0408	-0.0067	0.0394	-0.0065
	(0.0098)	(0.0121)	(0.0083)	(0.0116)	(0.0131)	(0.0166)	(0.0116)	(0.0145)
High School Degree		0.0581		0.0080		0.0068		0.0076
		(0.0349)		(0.0520)		(0.0511)		(0.0493)
College Degree		0.2289		0.1684		0.1684		0.1736
		(0.0385)		(0.0384)		(0.0405)		(0.0355)
Demos-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Degree Indicators	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R-squared	0.6229	0.6250	0.6168	0.6873	0.5795	0.6185	0.6168	0.6117
CD statistic	1017.3	1020.5						

Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.
Note: The dependent variable is the log of hourly wage adjusted to 1999 dollars. Our final sample contains $1,176$ individuals with $min(T_i) = 15$ for a total of $23,779$ person-year observations. Standard errors are shown in parentheses and heteroskedasticity-robust for cross-section and clustered at the person level for panel. Person fixed-effects and (quadratic) age controls are controlled for in all specifications. Panel B presents estimates with demographic controls including mother's and father's education levels as well as the number of siblings. Columns (3)-(4) are based on Interactive Fixed Effects (IFE) (Bai, 2009) with the estimated number of factors to be 9 across columns (See Table C.1 for more details). Columns (5)-(6) are based on Common Correlated Effects (CCE) (Pesaran, 2006). Columns (7)-(8) are based on the two-step CCE procedure with the number of factors set to be 4.

Table 4: Estimates of the Return to Schooling - Quadratic Schooling (Quadratic Age)

	(1) OLS	(2) OLS	(3) IFE	(4) IFE	(5) CCE	(6) CCE	(7) CCE-2	(8) CCE-2
Panel A: Without Demographic Controls								
Marginal Return to Schooling	0.0834	0.0410	0.0391	0.0116	0.0421	0.0051	0.0405	0.0042
	(0.0113)	(0.0144)	(0.0153)	(0.0140)	(0.0166)	(0.0226)	(0.0147)	(0.0197)
High School Degree		0.0317		0.0746		0.0323		0.0209
		(0.0397)		(0.0429)		(0.0594)		(0.0526)
College Degree		0.2425		0.1514		0.1400		0.1509
		(0.0384)		(0.0358)		(0.0412)		(0.0337)
Demos-by-year FE	No	No	No	No	No	No	No	No
Degree Indicators	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R-squared	0.6215	0.6238	0.8120	0.9288	0.7959	0.8295	0.8119	0.8117
CD statistic	24.968	22.836						
Panel B: With Demographic Controls								
Marginal Return to Schooling	0.0788	0.0391	0.0383	-0.0018	0.0374	-0.0026	0.0343	-0.0040
	(0.0116)	(0.0144)	(0.0103)	(0.0156)	(0.0171)	(0.0228)	(0.0147)	(0.0197)
High School Degree		0.0333		0.0221		0.0234		0.0234
		(0.0392)		(0.0542)		(0.0605)		(0.0544)
College Degree		0.2302		0.1487		0.1415		0.1539
		(0.0385)		(0.0382)		(0.0418)		(0.0346)
Demos-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Degree Indicators	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R-squared	0.6230	0.6250	0.7042	0.7147	0.6006	0.6326	0.6165	0.6111
CD statistic	1017.6	1021						

Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.
Note: The dependent variable is the log of hourly wage adjusted to 1999 dollars. Our final sample contains $1,176$ individuals with $min(T_i) = 15$ for a total of $23,779$ person-year observations. Standard errors are shown in parentheses and heteroskedasticity-robust for cross-section and clustered at the person level for panel. Person fixed-effects and (quadratic) age controls are controlled for in all specifications. Panel B presents estimates with demographic controls including mother's and father's education levels as well as the number of siblings. Columns (3)-(4) are based on Interactive Fixed Effects (IFE) (Bai, 2009) with the estimated number of factors to be 9 across columns (See Table C.1 for more details). Columns (5)-(6) are based on Common Correlated Effects (CCE) (Pesaran, 2006). Columns (7)-(8) are based on the two-step CCE procedure with the number of factors set to be 4.

Table 5: Bias Decomposition - Without the Factor Structure

	(1)	(2)	(3)	(4)
	Linear Schooling		Quadratic Schooling	
	Age^2	Age^4	Age^2	Age^4
Panel A: Without Demographic Controls				
Naive Return to Schooling	0.0760	0.0688	0.0834	0.0739
(Total)	(0.0097)	(0.0096)	(0.0113)	(0.0112)
Human Capital Component	0.0302	0.0264	0.0410	0.0358
(D1)	(0.0122)	(0.0120)	(0.0144)	(0.0142)
Sheepskin Component	0.0459	0.0424	0.0424	0.0381
(D2)	(0.0040)	(0.0039)	(0.0046)	(0.0045)
from HS Degree	0.0028	0.0016	0.0033	0.0013
	(0.0012)	(0.0011)	(0.0028)	(0.0027)
from COLL Degree	0.0431	0.0408	0.0391	0.0368
	(0.0069)	(0.0035)	(0.0032)	(0.0032)
D1/Total	40%	38%	49%	48%
D2/Total	60%	62%	51%	52%
Panel B: With Demographic Controls				
Naive Return to Schooling	0.0715	0.0653	0.0788	0.0699
(Total)	(0.0098)	(0.0097)	(0.0116)	(0.0115)
Human Capital Component	0.0286	0.0244	0.0391	0.0332
(D1)	(0.0121)	(0.0120)	(0.0144)	(0.0142)
Sheepskin Component	0.0428	0.0409	0.0397	0.0367
(D2)	(0.0039)	(0.0039)	(0.0045)	(0.0045)
from HS Degree	0.0026	0.0013	0.0034	0.0011
	(0.0011)	(0.0011)	(0.0027)	(0.0027)
from COLL Degree	0.0402	0.0396	0.0364	0.0357
	(0.0035)	(0.0035)	(0.0032)	(0.0032)
D1/Total	40%	37%	50%	47%
D2/Total	60%	63%	50%	53%

Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.

Table 6: Table 6: Bias Decomposition - With the Factor Structure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Linear Schooling						Quadratic Schooling					
	Age^2			Age^4			Age^2			Age^4		
	IFE	CCE	CCE2	IFE	CCE	CCE2	IFE	CCE	CCE2	IFE	CCE	CCE2
Panel A: Without Demographic Controls												
Naive Return to Schooling	0.0760	0.0760	0.0760	0.0688	0.0688	0.0688	0.0834	0.0834	0.0834	0.0739	0.0739	0.0739
(Total)	(0.0097)	(0.0097)	(0.0097)	(0.0096)	(0.0096)	(0.0096)	(0.0113)	(0.0113)	(0.0113)	(0.0112)	(0.0112)	(0.0112)
Human Capital Component	0.0050	0.0009	0.0016	0.0043	0.0004	0.0010	0.0116	0.0051	0.0042	0.0113	0.0046	0.0036
(C1)	(0.0105)	(0.0006)	(0.0125)	(0.0105)	(0.0005)	(0.0125)	(0.0130)	(0.0006)	(0.0174)	(0.0122)	(0.0006)	(0.0173)
Sheepskin Component	0.0299	0.0303	0.0305	0.0294	0.0298	0.0299	0.0322	0.0260	0.0265	0.0313	0.0252	0.0256
(C2)	(0.0054)	(0.0004)	(0.0050)	(0.0054)	(0.0004)	(0.0049)	(0.0070)	(0.0003)	(0.0064)	(0.0069)	(0.0004)	(0.0062)
from HS Degree	0.0002	0.0004	0.0001	0.0001	0.0003	0.0000	0.0078	0.0034	0.0022	0.0074	0.0031	0.0019
	(0.0017)	(0.0001)	(0.0019)	(0.0016)	(0.0001)	(0.0017)	(0.0040)	(0.0002)	(0.0047)	(0.0039)	(0.0002)	(0.0045)
from COLL Degree	0.0297	0.0299	0.0304	0.0293	0.0295	0.0299	0.0244	0.0226	0.0243	0.0240	0.0220	0.0237
	(0.0050)	(0.0003)	(0.0045)	(0.0050)	(0.0003)	(0.0044)	(0.0050)	(0.0002)	(0.0042)	(0.0049)	(0.0003)	(0.0041)
Ability Bias Component	0.0412	0.0448	0.0440	0.0351	0.0386	0.0379	0.0396	0.0524	0.0526	0.0312	0.0441	0.0447
(C3)	(0.0093)	(0.0004)	(0.0120)	(0.0094)	(0.0003)	(0.0119)	(0.0112)	(0.0005)	(0.0157)	(0.0100)	(0.0005)	(0.0156)
C1/Total	7%	1%	2%	6%	1%	1%	14%	6%	5%	15%	6%	5%
C2/Total	39%	40%	40%	43%	43%	43%	39%	31%	32%	42%	34%	35%
C3/Total	54%	59%	58%	51%	56%	55%	47%	63%	63%	42%	60%	60%
C1/(C1+C2)	14%	3%	5%	13%	1%	3%	26%	16%	14%	27%	15%	12%
C2/(C1+C2)	86%	97%	95%	87%	99%	97%	74%	84%	86%	73%	85%	88%

Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.

Table 6 Continued: Bias Decomposition - With the Factor Structure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Linear Schooling						Quadratic Schooling					
	Age^2			Age^4			Age^2			Age^4		
	IFE	CCE	CCE2	IFE	CCE	CCE2	IFE	CCE	CCE2	IFE	CCE	CCE2
Panel B: With Demographic Controls												
Naive Return to Schooling	0.0715	0.0715	0.0715	0.0653	0.0653	0.0653	0.0788	0.0788	0.0788	0.0699	0.0699	0.0699
(Total)	(0.0098)	(0.0098)	(0.0098)	(0.0097)	(0.0097)	(0.0097)	(0.0116)	(0.0116)	(0.0116)	(0.0115)	(0.0115)	(0.0115)
Human Capital Component	-0.0040	-0.0067	-0.0065	-0.0043	-0.0067	-0.0067	-0.0018	-0.0026	-0.0040	-0.0016	-0.0025	-0.0043
(C1)	(0.0103)	(0.0006)	(0.0125)	(0.0103)	(0.0006)	(0.0124)	(0.0142)	(0.0007)	(0.0173)	(0.0146)	(0.0007)	(0.0173)
Sheepskin Component	0.0299	0.0299	0.0308	0.0298	0.0299	0.0307	0.0257	0.0247	0.0267	0.0259	0.0246	0.0264
(C2)	(0.0061)	(0.0004)	(0.0051)	(0.0061)	(0.0004)	(0.0050)	(0.0077)	(0.0004)	(0.0064)	(0.0077)	(0.0004)	(0.0063)
from HS Degree	0.0004	0.0003	0.0003	0.0003	0.0003	0.0003	0.0022	0.0024	0.0024	0.0018	0.0023	0.0022
	(0.0020)	(0.0001)	(0.0019)	(0.0019)	(0.0001)	(0.0017)	(0.0048)	(0.0002)	(0.0047)	(0.0047)	(0.0002)	(0.0045)
from COLL Degree	0.0296	0.0296	0.0305	0.0295	0.0296	0.0305	0.0235	0.0223	0.0243	0.0241	0.0223	0.0242
	(0.0055)	(0.0004)	(0.0046)	(0.0056)	(0.0004)	(0.0046)	(0.0051)	(0.0003)	(0.0042)	(0.0053)	(0.0003)	(0.0042)
Ability Bias Component	0.0455	0.0483	0.0471	0.0398	0.0420	0.0412	0.0549	0.0567	0.0561	0.0457	0.0478	0.0478
(C3)	(0.0087)	(0.0004)	(0.0119)	(0.0087)	(0.0004)	(0.0119)	(0.0118)	(0.0005)	(0.0155)	(0.0121)	(0.0005)	(0.0155)
C1/Total	-6%	-9%	-9%	-7%	-10%	-10%	-2%	-3%	-5%	-2%	-4%	-6%
C2/Total	42%	42%	43%	46%	46%	47%	33%	31%	34%	37%	35%	38%
C3/Total	64%	68%	66%	61%	64%	63%	70%	72%	71%	65%	68%	68%
C1/(C1+C2)	-15%	-29%	-27%	-17%	-29%	-28%	-8%	-12%	-18%	-7%	-11%	-19%
C2/(C1+C2)	115%	129%	127%	117%	129%	128%	108%	112%	118%	107%	111%	119%

Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.

Appendix A: Bias Decomposition for the General Model

Here we provide details for the bias decomposition in the general model which includes a quadratic function of schooling and experience. The case with demographic controls included can be handled similarly. The general model can be expressed as

$$y_{it} = c_i + \beta_1 s_{it} + \beta_2 s_{it}^2 + \beta_3 s_{it} e_{it} + \gamma_1 e_{it} + \gamma_2 e_{it}^2 + z_{it}'\delta + v_{it} \tag{14}$$

$$v_{it} = \lambda_i' f_t + u_{it}, \;\; i = 1,...,N; \;\; t = t_i \in \mathscr{J}_i \equiv \{t_i(1), t_i(2), ..., t_i(T_i)\} \tag{15}$$

To begin with, let $S_{it} = (s_{it}, s_{it}^2, s_{it}e_{it})'$, $E_{it} = (e_{it}, e_{it}^2)'$, $\beta = (\beta_1, \beta_2, \beta_3)'$, $\gamma = (\gamma_1, \gamma_2)'$. Consequently, (14) can be written as

$$y_{it} = c_i + S_{it}'\beta + E_{it}'\gamma + z_{it}'\delta + v_{it} \tag{16}$$

As a matter of notation, for any variable (vector or scalar) w_{it}, define its time demeaned version as $\widetilde{w}_{it} = w_{it} - \bar{w}_i$, $\bar{w}_i = T_i^{-1}\sum_{t\in\mathscr{J}_i} w_{it}$. Further, let

$$\Sigma_{ss} = \sum_{i=1}^{N}\sum_{t\in\mathscr{J}_i} \widetilde{S}_{it}\widetilde{S}_{it}', \; \Sigma_{ee} = \sum_{i=1}^{N}\sum_{t\in\mathscr{J}_i} \widetilde{E}_{it}\widetilde{E}_{it}', \; \Sigma_{sy} = \sum_{i=1}^{N}\sum_{t\in\mathscr{J}_i} \widetilde{S}_{it}\widetilde{y}_{it}$$

$$\Sigma_{ey} = \sum_{i=1}^{N}\sum_{t\in\mathscr{J}_i} \widetilde{E}_{it}\widetilde{y}_{it}, \; \Sigma_{se} = \sum_{i=1}^{N}\sum_{t\in\mathscr{J}_i} \widetilde{S}_{it}\widetilde{E}_{it}', \; \Sigma_{sz} = \sum_{i=1}^{N}\sum_{t\in\mathscr{J}_i} \widetilde{S}_{it}\widetilde{z}_{it}',$$

$$\Sigma_{ez} = \sum_{i=1}^{N}\sum_{t\in\mathscr{J}_i} \widetilde{E}_{it}\widetilde{z}_{it}', \; \Psi_{ss} = \Sigma_{ss} - \Sigma_{se}\Sigma_{ee}^{-1}\Sigma_{se}'$$

$$\Psi_{sy} = \Sigma_{sy} - \Sigma_{se}\Sigma_{ee}^{-1}\Sigma_{ey}, \; \Psi_{sz} = \Sigma_{sz} - \Sigma_{se}\Sigma_{ee}^{-1}\Sigma_{ez}$$

We first estimate by OLS the following model:

$$y_{it} = c_i + S_{it}'\beta + E_{it}'\gamma + z_{it}'\delta + v_{it} \tag{17}$$

The resulting estimate of β can be expressed as

$$\hat{\beta}_{OLS} = \Psi_{ss}^{-1}\Psi_{sy}$$

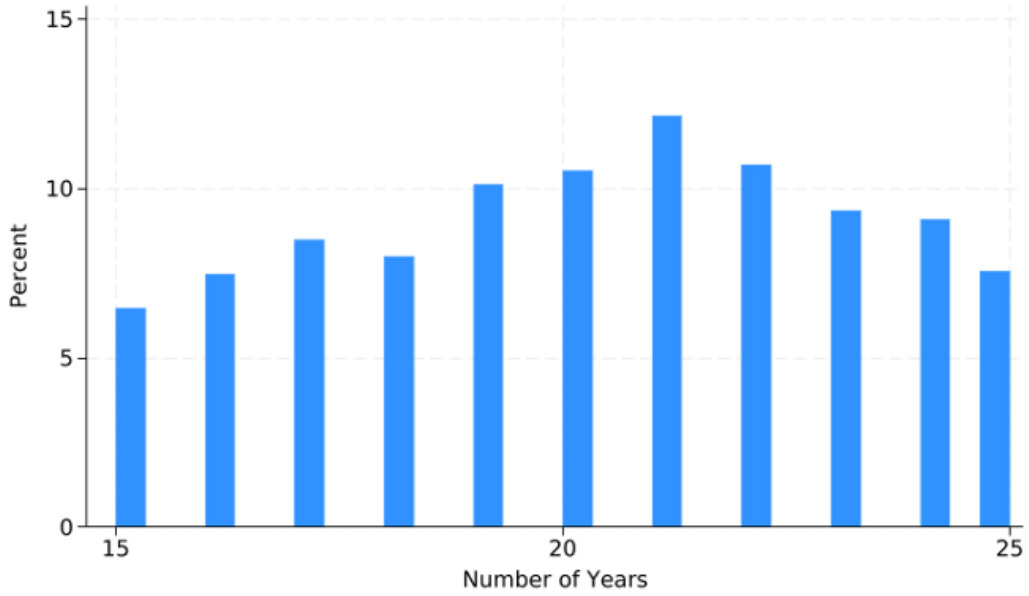
Next, estimate the general model (16) using one of the FM methods. Denote the resulting estimates of β and δ by $\hat{\beta}_{FM}$ and $\hat{\delta}_{FM}$, respectively, and the estimates of λ_i and f_t by $\hat{\lambda}_i$ and $\hat{f}_t$, respectively. Applying the Gelbach bias decomposition, we can then write

$$\begin{aligned}\hat{\beta}_{OLS} &= \hat{\beta}_{FM} + \Psi_{ss}^{-1}\Psi_{sz}\hat{\delta}_{FM} + \Psi_{ss}^{-1}\sum_{i=1}^{N}\sum_{t\in\mathscr{J}_i}\left[\widetilde{S}_{it} - \sum_{se}\sum_{ee}^{-1}\widetilde{E}_{it}\right]\hat{\lambda}_i'\hat{f}_t \\ &= C_1 + C_2 + C_3 \end{aligned} \tag{18}$$

We can now estimate the contribution of each of the three components and evaluate their relative importance as before. The main difference with the simpler model analyzed earlier is that in the general model, the estimates of β and δ are obtained after controlling for experience which is equivalent to running the regressions on the data residualized with respect to experience and its square.

Appendix B - Data

Figure B.1: Distribution of Number of Years



Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.
Note: The figure above shows the distribution of number of years for our analysis sample of 1,176 individuals after applying the sample restrictions including: (1) no missing covariates and at least 16 years old in a given year (2) white males (3) reported working at least 30 weeks and 800 hours in the previous year (4) removal of outlier observations in the top and bottom 1% of the wage distribution (5) no negative or abnormally large schooling changes (6) the minimum number of years within an individual to be 15 where observations are not required to be consecutive (7) no missing information for the timing of when degrees are received. Our final sample contains a total of 23,779 person-year observations in an unbalanced panel, where the maximum number of years is 25 (yearly observations from 1981 to 1994, and biannual observations thereafter from 1996 to 2016).

Table B.1: Variation in Years of Schooling for the Final Sample

A. Changes Per Person

Total Number of Changes	Number of Individuals	Percent (in %)
0	675	57.40%
1	256	21.77%
2	127	10.80%
3	62	5.27%
4	38	3.23%
5	10	0.85%
6	3	0.26%
7	3	0.26%
8	2	0.17%

B. Changes Per Education Level

Education Level after Schooling Change	Number of Cases	Percent (in %)
10	4	0.42%
11	53	5.56%
12	165	17.31%
13	135	14.17%
14	153	16.05%
15	116	12.17%
16	155	16.26%
17	79	8.29%
18	55	5.77%
19	25	2.62%
20	13	1.36%

Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.
Note: Panel A presents the number of schooling changes per person across our final sample. Panel B tabulates the schooling changes by education level after the change. For example, in the first row, we observed 4 instances of an individual's years of schooling changing to 10th grade, and 53 cases changing to 11th grade in the second row, and so on. Thus, Panel B indicates that most of the variation in schooling in our final sample comes from individuals completing grades years 12-16 (i.e., completing high school through 4 years of college). Note, the sum of (total number of changes from Panel A Column 1 $\times$ number of individuals in Panel A Column 2) = 953 which is the total number of cases shown in Panel B.

Table B.2: Distribution in Years of Schooling for a given Degree status

Years of Schooling	No Degree	HS Degree Only	HS & COLL Degree	Total
8	12	1		13
9	15	2		17
10	10			10
11	13	4		17
12	1	557	2	560
13		93	2	95
14		101	6	107
15		29	6	35
16		6	197	203
17			41	41
18			40	40
19			15	15
20			23	23
Total	51	793	332	1,176

Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.
Note:

Appendix C - Result

Table C.1: Estimates of the Return to Schooling - Linear Schooling (Quartic Age)

	(1) OLS	(2) OLS	(3) IFE	(4) IFE	(5) CCE	(6) CCE	(7) CCE-2	(8) CCE-2
Panel A: Without Demographic Controls								
Marginal Return to Schooling	0.0688	0.0264	0.0448	0.0043	0.0422	0.0004	0.0415	0.0010
	(0.0096)	(0.0120)	(0.0037)	(0.0120)	(0.0125)	(0.0165)	(0.0114)	(0.0146)
High School Degree		0.0354		0.0012		0.0062		0.0000
		(0.0345)		(0.0423)		(0.0504)		(0.0471)
College Degree		0.2312		0.1659		0.1670		0.1694
		(0.0381)		(0.0355)		(0.0400)		(0.0341)
Demos-by-year FE	No	No	No	No	No	No	No	No
Degree Indicators	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R-squared	0.6225	0.6246	0.9266	0.8528	0.7738	0.8152	0.8121	0.8123
CD statistic	24.438	23.29						
Panel B: With Demographic Controls								
Marginal Return to Schooling	0.0653	0.0244	0.0427	-0.0043	0.0408	-0.0067	0.0389	-0.0067
	(0.0097)	(0.0120)	(0.0070)	(0.0116)	(0.0131)	(0.0166)	(0.0116)	(0.0145)
High School Degree		0.0313		0.0066		0.0075		0.0067
		(0.0342)		(0.0518)		(0.0512)		(0.0494)
College Degree		0.2252		0.1681		0.1684		0.1734
		(0.0383)		(0.0388)		(0.0405)		(0.0355)
Demos-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Degree Indicators	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R-squared	0.6243	0.6263	0.6731	0.6872	0.5795	0.6185	0.6168	0.6117
CD statistic	1022.6	1025.4						

Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.
Note: The dependent variable is the log of hourly wage adjusted to 1999 dollars. Our final sample contains $1,176$ individuals with $min(T_i) = 15$ for a total of $23,779$ person-year observations. Standard errors are shown in parentheses and heteroskedasticity-robust for cross-section and clustered at the person level for panel. Person fixed-effects and (quartic) age controls are controlled for in all specifications. Panel B presents estimates with demographic controls including mother's and father's education levels as well as the number of siblings. Columns (3)-(4) are based on Interactive Fixed Effects (IFE) (Bai, 2009) with the estimated number of factors to be 9 across columns (See Table C.1 for more details). Columns (5)-(6) are based on Common Correlated Effects (CCE) (Pesaran, 2006). Columns (7)-(8) are based on the two-step CCE procedure with the number of factors set to be 4.

Table C.2: Estimates of the Return to Schooling - Quadratic Schooling (Quartic Age)

	(1) OLS	(2) OLS	(3) IFE	(4) IFE	(5) CCE	(6) CCE	(7) CCE-2	(8) CCE-2
Panel A: Without Demographic Controls								
Marginal Return to Schooling	0.0739	0.0358	0.0476	0.0113	0.0403	0.0046	0.0386	0.0036
	(0.0112)	(0.0142)	(0.0169)	(0.0133)	(0.0165)	(0.0226)	(0.0145)	(0.0196)
High School Degree		0.0127		0.0728		0.0310		0.0184
		(0.0388)		(0.0433)		(0.0594)		(0.0526)
College Degree		0.2325		0.1516		0.1392		0.1501
		(0.0382)		(0.0363)		(0.0412)		(0.0336)
Demos-by-year FE	No	No	No	No	No	No	No	No
Degree Indicators	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R-squared	0.6225	0.6247	0.9150	0.9287	0.7960	0.8296	0.8121	0.8118
CD statistic	24.283	23.162						
Panel B: With Demographic Controls								
Marginal Return to Schooling	0.0699	0.0332	0.0438	-0.0016	0.0376	-0.0025	0.0336	-0.0043
	(0.0115)	(0.0142)	(0.0139)	(0.0159)	(0.0171)	(0.0228)	(0.0148)	(0.0197)
High School Degree		0.0107		0.0186		0.0239		0.0226
		(0.0384)		(0.0547)		(0.0607)		(0.0545)
College Degree		0.2264		0.1526		0.1415		0.1536
		(0.0384)		(0.0392)		(0.0418)		(0.0346)
Demos-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Degree Indicators	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R-squared	0.6243	0.6263	0.6561	0.7158	0.6006	0.6326	0.6164	0.6111
CD statistic	1022.8	1025.8						

Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.
Note: The dependent variable is the log of hourly wage adjusted to 1999 dollars. Our final sample contains $1,176$ individuals with $min(T_i) = 15$ for a total of $23,779$ person-year observations. Standard errors are shown in parentheses and heteroskedasticity-robust for cross-section and clustered at the person level for panel. Person fixed-effects and (quartic) age controls are controlled for in all specifications. Panel B presents estimates with demographic controls including mother's and father's education levels as well as the number of siblings. Columns (3)-(4) are based on Interactive Fixed Effects (IFE) (Bai, 2009) with the estimated number of factors to be 9 across columns (See Table C.1 for more details). Columns (5)-(6) are based on Common Correlated Effects (CCE) (Pesaran, 2006). Columns (7)-(8) are based on the two-step CCE procedure with the number of factors set to be 4.