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

Mohitosh Kejriwal *Purdue University*

Xiaoxiao Li *Villanova University*

Evan Totty *U.S. Census Bureau*

December 5, 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, diploma signaling, 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 47% of the apparent returns to schooling can be attributed to ability bias while the diploma signaling and human capital components contribute about 39% and 14%, respectively. Furthermore, not accounting for ability bias or degree completion considerably exaggerates the human capital contribution to the returns. Finally, 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.

Keywords: human capital, diploma signaling, ability bias, interactive fixed effects, bias decomposition

JEL Classification: C33, C38, I26

*We are grateful to participants at the Midwest Economics Association, Midwest Econometrics Group, Eastern Economic Association, Southern Economic Association, Villanova School of Business Research Day, Villanova Economics Brown bag, Fudan University, University of Illinois Urbana-Champaign, Loyola University Maryland for useful comments. 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, 2018)

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 about individual skills.[1] As Lovenheim and Turner (2017) emphasize, the relevance of the human capital and signaling models for our education system is, at base, an empirical question.

[1]For detailed discussions of the human capital and signaling frameworks, see, inter alia, Lovenheim and Turner (2017) and Gunderson and Oreopolous (2020).

However, quantifying the relative importance of human capital versus signaling to the returns to education in practice is difficult. In both models, individuals choose the levels of education based on their skill set 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 (2018) 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 obtained from the effect of degree completion (i.e., earning a high school or college degree); thus, we follow Clark and Martorell (2014) and refer to it as the "diploma signaling" component. 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 estimate 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 estimate of the returns which represents the human capital component; (ii) the diploma signaling component obtained from the estimated effect of the degree completion variables; (iii) the ability bias component obtained from the estimated factor structure.[2]

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 diploma signaling component and the estimated factor structure that determines the contribution of the ability bias component.

Our empirical analysis is based on an unbalanced panel data set obtained from the National Longitudinal Survey of Youth (NLSY79). Results from our preferred specification show that about 47% of the apparent returns to schooling can be attributed to ability bias while the diploma signaling and human capital components contribute around 39% and 14%, respectively. The diploma signaling contribution is found to be driven almost entirely by the college degree while the high school degree plays only a very minor role. Furthermore, without accounting for ability, returns to education appear evenly split between human capital and diploma signaling, but controlling for ability shifts the balance sharply to 26/74 in favor of signaling. Likewise, omitting degree

[2]The diploma signaling (or ability bias) components referenced here reflect the contribution of degrees (or ability bias) to the apparent returns to schooling, as obtained through our decomposition as detailed in Section 3.2. This is distinct from the direct returns to degrees themselves, which are captured by the coefficients on the degree completion dummies in an augmented Mincer regression, commonly referred to as the "sheepskin effect." Although we also report and discuss these direct returns when presenting the regression results in Section 5.2, our primary focus is on decomposing the apparent returns to schooling in Section 5.3, which contributes to the empirical discussion of disentangling the roles of human capital, diploma signaling, and ability bias in explaining the returns to education.

completion exaggerates the role of human capital and masks the importance of ability bias to the overall returns to schooling.

A potential limitation of the panel data approach is that identification relies on observing individuals with variation in both education and earnings; consequently, a substantial share of our sample consists of individuals who have accumulated work experience prior to earning their final degrees. This includes both students who work while obtaining a degree and individuals who work between degrees, which may raise concerns about sample selection, since such a sample may differ from the traditional view of students who complete degrees consecutively without working until attaining their final degrees. Moreover, earnings during schooling could be part-time or seasonal work and not fully reflect an individual's earning potential (Card, 1995; Lazear, 1977). We believe that these concerns are alleviated by the facts that (i) we restrict our sample of analysis to observations that 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) as demonstrated in Section 4, a sizeable proportion of students gain work experience prior to degree completion, consistent with previous research documenting the increasing prevalence of student employment (see, e.g., Carnevale et al., 2015); (iii) the summary statistics show that observable characteristics are resilient to the sample restrictions we imposed; (iv) our OLS estimates of sheepskin effects (degree coefficients) align with prior cross-sectional evidence, providing further assurance that our sample is not unusually skewed. A detailed discussion of these issues is included in Section 4.

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 diploma signaling on returns to schooling and 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. Supplementary materials (not for publication) are included in Appendices A-C.

2 Issues in the Existing Literature

Quantifying the relative importance of human capital versus signaling to the observed returns to education remains a critical and challenging issue in economics (Lange and Topel, 2006; Lovenheim and Turner, 2017; Bradley and Green, 2020). The pure human capital model posits that 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 a signal of productivity, thereby improving employment prospects and wage outcomes without necessarily increasing productivity (Spence, 1973; Weiss, 1995). The empirical literature has yet to reach a consensus on the relative contributions of human capital and signaling to the private returns on education. The difficulty in empirically disentangling the contributions emanates from the fact that both models predict a positive relationship between education and wages (Deming, 2022). In what follows, we do not attempt an exhaustive review of the variety of approaches to this debate but rather limit our discussion to studies that serve to underscore our contribution to the literature.

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 that 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. The existence of such diploma 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 (2018) conducts a comprehensive review of the literature and concludes in favor of substantial sheepskin effects as measured by the estimated coefficients on the degree completion dummies, with high school diplomas associated with wage increases of approximately 3-27% and even greater effects for college diplomas, ranging from 24-58% (see Tables E.1 and E.2 of Caplan 2018 for further details). Based on the evidence surveyed, Caplan (2018) suggests a human capital/signaling split of roughly 60/40 for high school and 40/60 for college diplomas.

The interpretation of sheepskin effects as signaling value has faced significant critiques. Chiswick (1973) and Lange and Topel (2006) argue that if graduates are typically more efficient learners who receive a larger productivity boost from schooling than dropouts do, the wage disparities observed between graduates and dropouts might more accurately reflect a self-selection process among students rather than the signaling value of the diploma. Relatedly, degree completion might be associated with other desirable labor market traits (e.g., 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/AFQT/IQ scores, high school rank, and GPA generally continue to report considerable sheepskin effects (e.g., Frazis 1993; Kane and Rouse 1995; Arkes 1999; Light and Strayer 2004).

The employer learning framework offers an alternative perspective to examine this 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 value of education should diminish, implying that sheepskin effects become less pronounced as one's career progresses. Studies adopting this approach generally find a limited role for signaling relative to human capital (e.g., Lange, 2007; Aryal et al., 2022). A major critique of the employer learning literature is its neglect of unmeasurable abilities, which are crucial yet challenging to quantify.[3,4]

Our study contributes to the human capital versus signaling debate by addressing two crucial methodological issues in the extant literature. First, as evidenced by the preceding discussion, a common approach to accounting for ability bias entails the use of observable proxies for ability that are assumed to adequately represent the underlying latent skills that matter for earnings. However, the ability proxy approach has been shown to be ineffective in capturing the unobserved skills. Kejriwal et al. (2024) adopt a factor model framework to conduct a formal evaluation of ability proxies including ASVAB scores, the Rotter Locus of Control Scale, and the Rosenberg Self-Esteem Scale. They provide strong evidence against the efficacy of the proxies in explaining the estimated ability bias. Similar findings based on a structural human capital life cycle model are reported in Das and Polachek (2019). Huntington-Klein (2021) discusses the challenges involved in identifying the relative importance of human capital versus signaling owing in part to the lack of observable proxies that can reliably represent the multidimensional facets of ability. Our approach obviates the reliance on proxies by modeling the ability component via a flexible factor structure which allows multiple dimensions of ability with potentially time varying prices. The factor structure is estimated using panel data on earnings and schooling thereby providing a more data-driven approach to addressing ability bias than the use of proxies.[5]

[3]Other approaches that rely on ability proxies include Bedard (2001) that uses the Knowledge of the World of Work (KWW) test scores and Arteaga (2018) that utilizes Saber 11 test scores. Details are omitted to save space.

[4]Approaches that do not rely on proxies for ability include the regression discontinuity design of Clark and Martorell (2014) and the general equilibrium analysis of Fang (2006).

[5]A related contribution by Hussey (2012) utilizes panel data on MBA graduates and variation in pre-MBA work experience to empirically distinguish between the human capital and signaling models. His evidence supports the signaling hypothesis although he only uses individual fixed effects to control for ability bias.

The second issue pertains to the fact that the estimated contributions of the different components to schooling returns can be very sensitive to the sequence in which covariates are appended to the base model. Specifically, Gelbach (2016) shows that adding covariates sequentially to a base empirical specification can generate very different results regarding the importance of covariates, if the covariates are intercorrelated. He demonstrates the practical relevance of this finding by revisiting the empirical analyses in influential studies such as Neal and Johnson (1996) on the impact of premarket skills on racial wage disparities and Levitt and Syverson (2008) on the value of information in real estate markets. In the present context, the base specification only includes the schooling and experience variables along with demographic controls (if any). As shown in our empirical analysis (Section 5.3), the estimated contributions of human capital, diploma signaling, and ability bias depend quite heavily on the order in which the degree completion variables and the factor structure are added to the base specification. In contrast, our preferred results employ a bias decomposition due to Gelbach (2016) that is invariant to the sequence in which the base specification is augmented to include further covariates. Such a decomposition thus offers a more robust approach to estimating the marginal importance of each component relative to an approach based on sequential covariate addition. Finally, our analysis clarifies how ability bias influences the relative contributions of human capital and diploma signaling to the returns to schooling, and how signaling, in turn, affects the balance between human capital and ability bias.

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 extends the FM framework employed in Kejriwal et al. (2024) to include degree completion indicators and discusses 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 to schooling into its human capital, diploma signaling, and ability bias components. Section 3.3 spells out the research questions of interest and discusses how our framework facilitates their analysis.

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 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 (i.e., observations within an individual are not required to be consecutive.) 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.[6] Both

[6]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.

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.[7] The estimated loadings and factors 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 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

[7] 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 $\Lambda^* = \Lambda A^{-1\prime}$.

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):

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

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 earnings behavior. Rather, we are interested in estimating the average growth rate of earnings with schooling given by (3) employing the interactive fixed effects structure as a device to control for multidimensional 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 (see Appendix A.1 of Kejriwal et al., 2024 for details). 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, $s_{it}e_{it}$, and when degree variables are included, also $z_{h,it}$, and $z_{c,it}$ to proxy for the factors.[8] 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

[8] 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.

shows that the approach works well more generally (e.g., when 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). See Appendix A.2 of Kejriwal et al. (2024) for 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 discrepancies between the CCE and IFE estimators in practice since the latter does not require this condition to be satisfied.[9]

3.2 Decomposition Analysis

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

[9] Our empirical analysis includes the cross-sectional averages of six variables: $y_{it}, s_{it}, s_{it}^2, s_{it}e_{it}, z_{h,it}, z_{c,it}$, when degree variables are included. Thus, the rank condition implies $r \leq 6$. (When degree variables are excluded, $r \leq 4$.) 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 Kejriwal et al., 2020 for details).

between the estimate from the base specification (that includes only schooling and experience variables) and the full specification (that also includes the degree dummies and the factor structure) 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 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 5.3), the estimated contribution can differ considerably depending on the order in which the degree completion variables and the factor structure are appended to the base specification that only includes the human capital and experience variables.

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

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

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

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

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 diploma 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 overall returns to schooling. 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, diploma signaling, and ability bias components to the overall returns to schooling (i.e., $\hat{\beta}_{OLS}$)?

2. What is the role of ability bias in determining the relative contributions of the human capital and diploma signaling components to the overall returns to schooling?

3. What is the role of diploma signaling in determining the relative contributions of the human capital and ability bias components to the overall returns to schooling?

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 diploma signaling components to the overall MRTS. For example, if ability bias is an important determinant of the MRTS but is not controlled for so that one relies on the decomposition (9), the contributions of both the human capital and diploma signaling components to the schooling returns 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 diploma signaling components in terms of their estimated contributions to the MRTS varies depending on whether ability bias is accounted for. Of course, the direction and magnitude of the biases in D_1 and D_2 are specific to the application/data and will be discussed in our empirical analysis.

Finally, to address the third research question, we first estimate a model that excludes the degree 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 controlling for degree completion. If diploma signaling is important to the MRTS, 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.[10]

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) observations that 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

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

observations, where observations within an individual are not required to be consecutive.

Table 1 presents descriptive statistics for not only the main sample of analysis but also how they change as we sequentially add the selection criteria to obtain our final sample.[11] Panel A reports summary statistics calculated at the person-year level ($N*T$), reflecting the full panel dataset, whereas Panel B presents statistics at the individual level (N), summarizing the cross-sectional distribution of characteristics across individuals in the panel. Column (1) presents a baseline sample of 7,849 individuals who are at least 16 years old in a given year and have no missing covariates. Among these individuals, the average age at high school completion is approximately 19, and the average age at college graduation is around 26. Some individuals have no recorded observations before earning a degree, particularly a high school degree. Conditional on having at least one pre-degree observation, the data indicate that working before obtaining a high school or college degree is quite common. Specifically, 39.5% of person-year observations prior to high school graduation show work experience, and 67.5% of person-year observations prior to earning a college degree show work experience. The share of individuals without any work experience before obtaining a high school diploma - representing more traditional students - is approximately 18.2%, and this share drops to just 1.9% for college graduates. These patterns are consistent with existing literature documenting that a substantial share of students engage in work while enrolled in school (see, e.g., Carnevale et al., 2015).

Focusing our attention on white males decreases the sample by more than two-thirds to 2,314 individuals in column (2). It is worth noting that hourly wage earning is 26.5% higher on average for white males compared to the baseline sample. Additionally, the share of observations with work experience prior to degree completion increases - particularly for the high school degree, where 47.3% of person-year records reflect work before graduation. Column (3) imposes wages and hours worked restrictions, but only shrinks the sample by 17 individuals, presumably due

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

to the aforementioned reason that white males above age 16 are strongly attached to the labor market. Moreover, the share of observations with pre-degree work experience remains consistent with column (2) - for instance, 47.6% of person-year observations reflect work prior to high school graduation - suggesting that the prevalence of pre-graduation work experience is not an artifact of the wage and hours restrictions imposed. After trimming the wage outliers in Column (4), we lose another person, and the number of observations decreases to 34,939. 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 1,987 individuals for a total of 30,998 person-year observations, after we carefully exclude individuals with clearly misreported schooling changes. Lastly, 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, our final sample of analysis shrinks to 1,176 individuals for a total of 23,779 person-year observations. In this final sample, 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, amongst which 94% of the person-year observations have a high school diploma, and 24% have a college degree.[12] The share of observations prior to high school graduation reflecting work experience is 50.1%, and 76% prior to college graduation.

While 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 diploma signaling effects.[13] The literature on statistical discrimination suggests that diploma signaling 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

[12] Appendix B provides additional details on the variation in years of schooling and its relation to degree attainment.

[13] 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.

completing 16 or more years of education compared to their white counterparts. Similarly, the diploma signaling effects could vary if diplomas provide different productivity signals for those individuals who have accumulated work experience prior to earning their final degrees. We believe this concern is mitigated by the following facts: (1) As discussed above, working students do not represent a marginal subset but in fact constitute the norm: In our baseline sample of 7,849 individuals, only 18.2% had no work experience prior to obtaining a high school diploma, and just 1.9% lacked such experience before completing a college degree. These figures are consistent with prior research showing that employment during school is both prevalent and growing (see, e.g., Bound et al., 2012; Carnevale et al., 2015). The traditional view - that students finish school before entering the labor market - underrepresents the diversity of modern education-labor pathways. (2) Most observable characteristics in Table 1 remain fairly stable across columns and are resilient to additional restrictions imposed; (3) As discussed in Section 5.2, the magnitude of our OLS estimates of high school and college sheepskin effects aligns well with those reported in the literature based on cross-sectional data, providing further assurance that our sample is not unusually skewed; (4) To the extent one believes diploma signals may be weaker for students with prior work experience—perhaps because employers place less weight on education credentials when work history is available—our estimates may understate the true signaling value of degrees for more "traditional" students. In this sense, the diploma signaling effects we identify may be viewed as a lower bound (Caplan, 2018). Taken together, we believe that sample selection associated with using panel variation is not as severe a threat to our conclusions as might initially appear.

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 different specifications. Section 5.3 reports the results of our decomposition analysis that facilitate estimation of the relative contributions of the different components (human capital, diploma 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, namely, parental education levels and the number of siblings. 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, Tables 3-5 reports parameter estimates using quadratic age controls without demographic covariates. Estimates that incorporate demographic controls, as well as those based on quartic age controls - which yield qualitatively similar results - are provided in Appendix C.

5.2 Parameter Estimates

Table 3 presents 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 in Panel A and B, respectively. 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 Table 3, the OLS estimates of the MRTS without degree controls are between 7.5%-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 as shown in Panel A, the return on a high school degree is approximately 6%, with marginal statistical significance at the 10% level. However, under a quadratic schooling model in Panel B, the estimate drops to about 3% and becomes statistically insignificant.[14] The OLS return to a college degree is substantially higher, at around 24%, and statistically significant at the 1% level. This notable college premium remains robust across the various model specifications as shown in Appendix C, demonstrating a strong and stable increase in earnings that a college degree provides, conditional on the same years of schooling.[15] It is useful to note that the estimated sheepskin coefficient 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 literature, where the year twelve premium varies from 2.6% in

[14]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.2 and C.3.

[15]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 (7.1%) 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.

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 4%-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, 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 of the MRTS 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 causal evidence 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 (Panel A), the results consistently demonstrate both statistically and economically insignificant effects of a high school degree across various estimators. When using quadratic schooling models (Panel B), the IFE estimator without demographic controls [column (4)] yields a relatively large effect (around 7.5%, statistically significant at the 10% level) compared to other estimators as well as specifications that include demographic controls reported in Table C.1 (about 2%-3%, statistically insignificant). The FM returns to 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 (see, e.g., Tobias and Li, 2004). Taken together, our findings corroborate the perspective articulated by Lovenheim and Turner (2017) that the signaling model holds greater relevance for some segments of the education system than others: while the high school diploma carries little independent signaling value, the college degree offers a substantial premium.

Fifth, 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, attaining its highest value of 0.9288 under the quadratic schooling–quadratic age specification estimated with the IFE estimator, as reported in Panel B, column (4), which we therefore regard as our preferred specification. Meanwhile 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 Decomposition Analysis

Our primary objective is to decompose the apparent returns to schooling in order to address the research questions outlined in Section 3.3. To this end, Table 4 details the 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 diploma signaling components (research question 2 in Section 3.3). For brevity, Table 4 only displays results with quadratic age controls but without demographic controls: Columns (1) and (2) present linear and quadratic schooling models, respectively. Robustness results using quartic age specifications without demographic controls are shown in Table C.4, and the corresponding results with demographic controls are reported in Table C.5.

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 the overall MRTS 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 diploma signaling appears slightly more pronounced at about 60% relative to a human capital contribution of around 40%. However, when schooling enters quadratically, the contributions of human capital and diploma signaling to the returns to schooling tend to be roughly equal. These splits are unaffected by the modeling choice for the age controls, and remain consistent whether demographic controls are included or not, as shown in Table C.4 and C.5.

Table 5 displays the 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)], without demographic controls. The structure of Table 5 mirrors that of Table 4, wherein each model setup - linear or quadratic schooling with quadratic age controls - features results for the three different estimators, namely, IFE, CCE, and CCE-2. Table C.6 presents the results using quartic rather than quadratic age specifications (without demographic controls), while Table C.7 documents the corresponding results with demographic controls.

Correcting for ability bias yields rather different results regarding the contributions of human capital and diploma signaling to the overall MRTS. 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) to the overall MRTS is minimal, suggesting that much of the human capital effects documented in previous studies may be overstated and attributable to self-selection. In contrast, the contribution of diploma signaling to the overall MRTS 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. In our preferred specification, which uses the IFE estimator along with a quadratic specification in schooling and age [Table 5, column (4)], the distribution between human capital, diploma signaling, and ability bias is represented as a 14/39/47 split. Averaging across all columns in Table 5, the corresponding shares are 7/36/57, respectively.

It is useful to compare our results in Table 5 with those from a sequential covariate addition approach in order to highlight the practical relevance of our 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-DS-AB approach, first adds the degree indicators followed by the factor structure. The second, labeled the HC-AB-DS approach, first adds the factor structure followed by the degree indicators. For the HC-DS-AB approach, we can write

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

where $MRTS_{ols}^{base}$ is the OLS estimate from the base specification, $MRTS_{ols}^{ds}$ 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 Table 3. The quantities C_2^* and C_3^* are the estimated contributions of the diploma signaling and ability bias components to the baseline

MRTS, respectively, obtained by employing the HC-DS-AB approach. We can easily compute C_2^* and C_3^* from Table 3. For instance, consider the quadratic schooling specification that employs quadratic age controls (Panel B). 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, $C_2/MRTS_{ols}^{base} \simeq 39\%$, $C_3/MRTS_{ols}^{base} \simeq 47\%$ [Table 5 Column (4)]. Thus, the HC-DS-AB approach tends to exaggerate the contribution of the diploma signaling component while understating the contribution of the ability bias component in explaining the apparent returns to schooling. An analogous pattern holds for the other specifications and FM estimators. A similar analysis for the HC-AB-DS approach shows that the relative importance of the diploma signaling and ability bias components depends on the specific FM estimator employed. In particular, the IFE approach overstates (understates) the ability bias (diploma signaling) contribution to the overall MRTS, 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, diploma 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 diploma signaling component also diminishes, though to a lesser extent (i.e., C2 is slightly smaller than D2). Consequently, the distribution between human capital and diploma signaling in explaining the returns to schooling shifts in favor of diploma signaling. For instance, our preferred specification using the IFE estimator combined with a quadratic specification in schooling and age as shown in Table 5, column (4) implies a 26/74 split favoring signaling, as opposed to a nearly equal division when ability controls are omitted [Table 4, column (2)]. These findings highlight the crucial role of accounting for ability bias in accurately assessing the relative contributions of human capital and the diploma signaling components to the overall returns on education.

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

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

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 Table 3. A comparison of G_1 and G_3 with C_1 and C_3 (reported in Table 5) serves to showcase the empirical importance of controlling for degree completion when estimating the relative contributions of human capital and ability bias to the overall MRTS. For example, in our preferred specification employing the IFE estimator in conjunction with a specification that is quadratic in schooling and age [Table 5, column (4)], $C_1/(C_1+C_3) \simeq 23\%$ and $C_3/(C_1+C_3) \simeq 77\%$. In contrast, from columns (1) and (3) of Panel B in Table 3, $G_1/(G_1+G_3) = .0391/0.0834 \simeq 47\%$ and $G_3/(G_1+G_3) = 0.0443/0.0834 \simeq 53\%$, indicating that omitting degree completion considerably overstates the human capital component and understates the ability bias component.

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 to the ongoing debate about the relative importance of human capital and diploma signaling in explaining the returns to education. Our factor model-based analysis of the NLSY79 suggests that the observed returns to schooling are predominantly driven by ability bias, accounting for about 47% of these returns, while actual benefits from degree attainment and human capital accumulation are less pronounced than previously thought. Moreover, the diploma signaling effect is predominantly influenced by college degrees, whereas high school degrees contribute minimally. Importantly, without adjusting for ability, the distribution between human capital and diploma signaling is more evenly split at around 50/50. However, once multidimensional ability is accounted for, this distribution shifts to 26/74, 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. Furthermore, our results suggest the empirical importance of controlling for degree completion when estimating the MRTS, as omitting degree indicators overstates the human capital component and understates the role of ability bias.

Our methodology, which avoids reliance on inadequate ability proxies and does not depend on the sequence of covariate inclusion, offers 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.

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.

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.

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

Huntington-Klein, N. (2021). Human capital versus signaling is empirically unresolvable. *Empirical Economics 60*(5), 2499–2531.

Hussey, A. (2012). Human capital augmentation versus the signaling value of mba education. *Economics of Education Review 31*(4), 442–451.

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.

Kim, D. and T. Oka (2014). Divorce law reforms and divorce rates in the usa: an interactive fixed-effects approach. *Journal of Applied Econometrics 29*(2), 231–245.

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.

Levitt, S. D. and C. Syverson (2008). Market distortions when agents are better informed: The value of information in real estate transactions. *The Review of Economics and Statistics 90*(4), 599–611.

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.

Neal, D. A. and W. R. Johnson (1996). The role of premarket factors in black-white wage differences. *Journal of political Economy 104*(5), 869–895.

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.

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 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) $min(T_i) = 15$
Panel A: Panel Statistics						
Hourly wage inflation-adjusted	19.25	24.35	24.88	16.27	16.16	16.66
	(359.60)	(378.94)	(361.66)	(10.29)	(10.11)	(10.11)
Log hourly wage inflation-adjusted	2.38	2.59	2.66	2.62	2.62	2.66
	(0.74)	(0.76)	(0.74)	(0.58)	(0.57)	(0.56)
Years of school	13.28	13.45	13.54	13.49	13.35	13.28
	(2.31)	(2.41)	(2.41)	(2.37)	(2.21)	(2.15)
Age	33.02	32.65	33.54	33.26	33.31	34.33
	(10.24)	(10.18)	(9.95)	(9.83)	(9.83)	(10.06)
Mother's years of school	11.16	12.03	12.03	12.00	11.97	11.97
	(3.11)	(2.38)	(2.36)	(2.35)	(2.30)	(2.22)
Father's years of school	11.14	12.28	12.28	12.25	12.17	12.19
	(3.89)	(3.36)	(3.31)	(3.29)	(3.23)	(3.16)
Number of siblings	3.58	2.95	2.94	2.95	2.96	2.94
	(2.48)	(1.95)	(1.93)	(1.94)	(1.94)	(1.90)
High school (HS) degree indicator	0.92	0.92	0.94	0.94	0.94	0.94
	(0.28)	(0.27)	(0.24)	(0.24)	(0.23)	(0.23)
College degree indicator	0.22	0.26	0.27	0.27	0.25	0.24
	(0.41)	(0.44)	(0.45)	(0.44)	(0.43)	(0.43)
Number of Observations ($N*T$)	137,734	41,771	35,923	34,939	30,998	23,779
Panel B: Cross-section Characteristics						
Age at HS degree	18.985	18.89	18.88	18.88	18.707	18.716
	(3.14)	(3.046)	(3.027)	(3.028)	(2.644)	(2.682)
Age at college degree	26.311	25.151	25.149	25.149	24.926	25.265
	(6.981)	(4.903)	(4.913)	(4.913)	(4.796)	(5.644)
Observations before HS degree	1.479	1.389	1.376	1.376	1.156	1.138
	(2.863)	(2.752)	(2.732)	(2.733)	(2.425)	(2.449)
Observations before college degree	7.145	6.291	6.277	6.277	5.890	6.373
	(5.818)	(4.513)	(4.518)	(4.518)	(4.467)	(4.852)
Observations before HS degree	3.176	3.011	2.995	2.997	2.723	2.672
(non-zero obs.)	(3.495)	(3.396)	(3.378)	(3.379)	(3.096)	(3.162)
Observations before college degree	7.513	6.554	6.540	6.540	6.201	6.592
(non-zero obs.)	(5.730)	(4.415)	(4.421)	(4.421)	(4.368)	(4.786)
Share of observations working,	0.395	0.473	0.476	0.476	0.465	0.501
before HS degree (non-zero obs.)	(0.392)	(0.401)	(0.401)	(0.401)	(0.407)	(0.411)
Share of observations working,	0.675	0.694	0.695	0.695	0.706	0.760
before college degree (non-zero obs.)	(0.291)	(0.275)	(0.275)	(0.275)	(0.277)	(0.250)
Fraction of individuals without work	0.182	0.149	0.148	0.147	0.144	0.133
before HS degree (non-zero obs.)	(0.385)	(0.356)	(0.355)	(0.354)	(0.351)	(0.339)
Fraction of individuals without work	0.019	0.018	0.018	0.018	0.018	0.010
before college degree (non-zero obs.)	(0.137)	(0.132)	(0.132)	(0.132)	(0.132)	(0.101)
Number of Individuals (N)	7,849	2,314	2,297	2,296	1,987	1,176

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)-(6) sequentially add additional sample criteria until the final sample is shown in column (6). Hourly earnings are adjusted for inflation to 1999 dollars.

Table 2: 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 (instead of quadratic) 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 3: Estimates of the Returns to Schooling - Without Demographic Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	IFE	IFE	CCE	CCE	CCE-2	CCE-2
Panel A: Linear Schooling, Quadratic Age								
Marginal Returns to Schooling	0.0760	0.0302	0.0460	0.0050	0.0437	0.0009	0.0430	0.0030
	(0.0095)	(0.0119)	(0.0054)	(0.0119)	(0.0113)	(0.0150)	(0.0115)	(0.0148)
High School Degree		0.0581		0.0034		0.0073		0.0043
		(0.0338)		(0.0423)		(0.0471)		(0.0488)
College Degree		0.2411		0.1664		0.1676		0.1668
		(0.0397)		(0.0352)		(0.0404)		(0.0360)
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: Quadratic Schooling, Quadratic Age								
Marginal Returns to Schooling	0.0834	0.0410	0.0391	0.0116	0.0421	0.0051	0.0405	0.0075
	(0.0113)	(0.0140)	(0.0153)	(0.0140)	(0.0164)	(0.0218)	(0.0147)	(0.0202)
High School Degree		0.0317		0.0746		0.0323		0.0250
		(0.0396)		(0.0429)		(0.0582)		(0.0550)
College Degree		0.2425		0.1514		0.1400		0.1445
		(0.0399)		(0.0358)		(0.0432)		(0.0349)
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						

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. Columns (3)-(4) are based on Interactive Fixed Effects (IFE) (Bai, 2009), with the estimated number of factors to be 7 and 6, respectively, in Panel A, and 4 and 11 in Panel B. The number of factors is selected following the approach of Kim and Oka (2014). 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 and 6, respectively. Tables C.1–C.3 present robustness checks of schooling return estimates under alternative specifications. Table C.1 replicates our baseline quadratic age specification with demographic controls, while Tables C.2 and C.3 show quartic age specifications without and with demographic controls, respectively.

Table 4: Decomposition Analysis - Without the Factor Structure

	(1) Linear Schooling, Quadratic Age	(2) Quadratic Schooling, Quadratic Age
Without Demographic Controls		
Overall Return to Schooling	0.0760	0.0834
(Total)	(0.0095)	(0.0113)
Human Capital Component	0.0302	0.0410
(D1)	(0.0119)	(0.0140)
Diploma Signaling Component	0.0459	0.0424
(D2)	(0.0077)	(0.0082)
from HS Degree	0.0028	0.0033
	(0.0016)	(0.0041)
from COLL Degree	0.0431	0.0391
	(0.0071)	(0.0064)
D1/Total	40%	49%
	(13%)	(13%)
D2/Total	60%	51%
	(13%)	(13%)

Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.
Note: This table presents the decomposition results obtained without incorporating the factor structure, as specified in equation (9). The overall returns to schooling are OLS estimates based on the baseline specification, which includes only schooling and quadratic experience variables, while excluding degree indicators and the factor structure. For example, the overall MRTS under the linear schooling and quadratic age specification, reported in Column (1), is 7.6%. This estimate corresponds to the MRTS result shown in Table 3, Panel A, Column (1). Table C.4 displays the decomposition results with quartic age specifications without demographic controls. Table C.5 displays the decomposition results with demographic controls.

Table 5: Decomposition Analysis - With the Factor Structure

	(1)	(2)	(3)	(4)	(5)	(6)
	Linear Schooling, Quadratic Age			Quadratic Schooling, Quadratic Age		
	IFE	CCE	CCE2	IFE	CCE	CCE2
Without Demographic Controls						
Overall Return to Schooling	0.0760	0.0760	0.0760	0.0834	0.0834	0.0834
(Total)	(0.0095)	(0.0095)	(0.0095)	(0.0113)	(0.0113)	(0.0113)
Human Capital Component	0.0050	0.0009	0.0030	0.0116	0.0051	0.0075
(C1)	(0.0119)	(0.0150)	(0.0148)	(0.0140)	(0.0218)	(0.0202)
Diploma Signaling Component	0.0299	0.0303	0.0300	0.0322	0.0260	0.0259
(C2)	(0.0069)	(0.0078)	(0.0072)	(0.0081)	(0.0099)	(0.0085)
from HS Degree	0.0002	0.0004	0.0002	0.0078	0.0034	0.0026
	(0.0020)	(0.0023)	(0.0024)	(0.0045)	(0.0061)	(0.0057)
from COLL Degree	0.0297	0.0299	0.0298	0.0244	0.0226	0.0233
	(0.0063)	(0.0072)	(0.0064)	(0.0058)	(0.0070)	(0.0001)
Ability Bias Component	0.0412	0.0448	0.0440	0.0396	0.0524	0.0526
(C3)	(0.0093)	(0.0132)	(0.0122)	(0.0112)	(0.0177)	(0.0162)
C1/Total	7%	1%	4%	14%	6%	9%
	(16%)	(20%)	(19%)	(17%)	(26%)	(24%)
C2/Total	39%	40%	39%	39%	31%	30%
	(9%)	(10%)	(10%)	(10%)	(12%)	(10%)
C3/Total	54%	59%	57%	47%	63%	61%
	(13%)	(18%)	(17%)	(14%)	(22%)	(20%)

Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.
Note: This table presents the decomposition results obtained when incorporating the factor structure, as specified in equation (7). Table C.6 displays the decomposition results with quartic age specifications without demographic controls. Table C.7 displays the decomposition results with demographic controls.

Appendix A: Decomposition Analysis for the General Model

Here we provide details for the decomposition analysis 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

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

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

$$\begin{aligned}
&\Sigma_{ss} = \sum_{i=1}^{N} \sum_{t \in \mathscr{J}_i} \widetilde{S}_{it} \widetilde{S}_{it}', \; \Sigma_{ee} = \sum_{i=1}^{N} \sum_{t \in \mathscr{J}_i} \widetilde{E}_{it} \widetilde{E}_{it}', \; \Sigma_{sy} = \sum_{i=1}^{N} \sum_{t \in \mathscr{J}_i} \widetilde{S}_{it} \widetilde{y}_{it} \\
&\Sigma_{ey} = \sum_{i=1}^{N} \sum_{t \in \mathscr{J}_i} \widetilde{E}_{it} \widetilde{y}_{it}, \; \Sigma_{se} = \sum_{i=1}^{N} \sum_{t \in \mathscr{J}_i} \widetilde{S}_{it} \widetilde{E}_{it}', \; \Sigma_{sz} = \sum_{i=1}^{N} \sum_{t \in \mathscr{J}_i} \widetilde{S}_{it} \widetilde{z}_{it}', \\
&\Sigma_{ez} = \sum_{i=1}^{N} \sum_{t \in \mathscr{J}_i} \widetilde{E}_{it} \widetilde{z}_{it}', \; \Psi_{ss} = \Sigma_{ss} - \Sigma_{se} \Sigma_{ee}^{-1} \Sigma_{se}' \\
&\Psi_{sy} = \Sigma_{sy} - \Sigma_{se} \Sigma_{ee}^{-1} \Sigma_{ey}, \; \Psi_{sz} = \Sigma_{sz} - \Sigma_{se} \Sigma_{ee}^{-1} \Sigma_{ez}
\end{aligned}$$

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 decomposition, we can then write

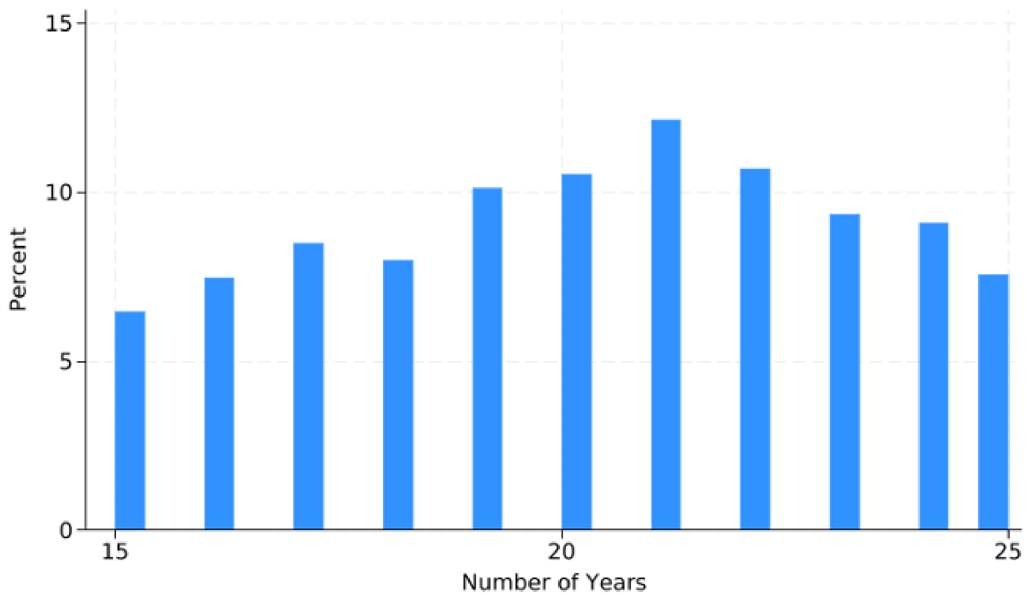
$$\begin{aligned}\hat{\beta}_{OLS} &= \hat{\beta}_{FM} + \Psi_{ss}^{-1}\Psi_{sz}\hat{\delta}_{FM} + \Psi_{ss}^{-1}\sum_{i=1}^{N}\sum_{t\in\mathscr{J}_i}\left[\widetilde{S}_{it} - \Sigma_{se}\Sigma_{ee}^{-1}\widetilde{E}_{it}\right]\hat{\lambda}_i'\hat{f}_t \\ &= C_1 + C_2 + C_3 \end{aligned} \tag{18}$$

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

Delving deeper into the variations of schooling and degree variables provides useful insights. 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% of 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 considerable variation 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, some completed it in less than 12 years, while others took more time - namely, 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 schooling and degree effects (Jaeger and Page, 1996).

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: Joint Distribution of Years of Schooling and Degree Attainment

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: This table presents a detailed cross-tabulation between years of schooling and degree attainment. There is considerable variation 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, 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.

Appendix C - Additional Results

Table C.1: Estimates of the Returns to Schooling - With Demographic Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	IFE	IFE	CCE	CCE	CCE-2	CCE-2
Panel A: Linear Schooling, Quadratic Age								
Marginal Returns to Schooling	0.0715	0.0286	0.0383	-0.0040	0.0408	-0.0067	0.0394	-0.0061
	(0.0128)	(0.0165)	(0.0083)	(0.0116)	(0.0113)	(0.0145)	(0.0116)	(0.0146)
High School Degree		0.0581		0.0080		0.0068		0.0047
		(0.0537)		(0.0520)		(0.0484)		(0.0503)
College Degree		0.2289		0.1684		0.1684		0.1700
		(0.0568)		(0.0384)		(0.0403)		(0.0367)
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						
Panel B: Quadratic Schooling, Quadratic Age								
Marginal Returns to Schooling	0.0788	0.0391	0.0383	-0.0018	0.0374	-0.0026	0.0343	-0.0023
	(0.0158)	(0.0192)	(0.0103)	(0.0156)	(0.0159)	(0.0215)	(0.0147)	(0.0199)
High School Degree		0.0333		0.0221		0.0234		0.0192
		(0.0587)		(0.0542)		(0.0597)		(0.0558)
College Degree		0.2302		0.1487		0.1415		0.1489
		(0.0572)		(0.0382)		(0.0431)		(0.0356)
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. Demographic controls include 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 4 and 8, respectively, in Panel A, and 9 and 10 in Panel B. The number of factors is selected following the approach of Kim and Oka (2014). 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 and 6, respectively.

Table C.2: Estimates of the Returns to Schooling - Without Demographic Controls (Quartic Age)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	IFE	IFE	CCE	CCE	CCE-2	CCE-2
Panel A: Linear Schooling, Quartic Age								
Marginal Returns to Schooling	0.0688	0.0264	0.0448	0.0043	0.0422	0.0004	0.0415	0.0024
	(0.0095)	(0.0118)	(0.0037)	(0.0120)	(0.0113)	(0.0148)	(0.0114)	(0.0147)
High School Degree		0.0354		0.0012		0.0062		0.0026
		(0.0334)		(0.0423)		(0.0474)		(0.0487)
College Degree		0.2312		0.1659		0.1670		0.1663
		(0.0396)		(0.0355)		(0.0403)		(0.0359)
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: Quadratic Schooling, Quartic Age								
Marginal Returns to Schooling	0.0739	0.0358	0.0476	0.0113	0.0403	0.0046	0.0386	0.0069
	(0.0114)	(0.0139)	(0.0169)	(0.0133)	(0.0160)	(0.0216)	(0.0145)	(0.0201)
High School Degree		0.0127		0.0728		0.0310		0.0231
		(0.0387)		(0.0433)		(0.0583)		(0.0547)
College Degree		0.2325		0.1516		0.1392		0.1438
		(0.0400)		(0.0363)		(0.0432)		(0.0348)
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						

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. Columns (3)-(4) are based on Interactive Fixed Effects (IFE) (Bai, 2009), with the estimated number of factors to be 11 and 6, respectively, in Panel A, and 10 and 11 in Panel B. The number of factors is selected following the approach of Kim and Oka (2014). 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 and 6, respectively.

Table C.3: Estimates of the Returns to Schooling - With Demographic Controls (Quartic Age)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	IFE	IFE	CCE	CCE	CCE-2	CCE-2
Panel A: Linear Schooling, Quartic Age								
Marginal Returns to Schooling	0.0653	0.0244	0.0427	-0.0043	0.0408	-0.0067	0.0389	-0.0062
	(0.0128)	(0.0165)	(0.0070)	(0.0116)	(0.0113)	(0.0145)	(0.0116)	(0.0146)
High School Degree		0.0313		0.0066		0.0075		0.0041
		(0.0555)		(0.0518)		(0.0486)		(0.0502)
College Degree		0.2252		0.1681		0.1684		0.1698
		(0.0569)		(0.0388)		(0.0404)		(0.0367)
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						
Panel B: Quadratic Schooling, Quartic Age								
Marginal Returns to Schooling	0.0699	0.0332	0.0438	-0.0016	0.0376	-0.0025	0.0336	-0.0024
	(0.0159)	(0.0192)	(0.0139)	(0.0159)	(0.0159)	(0.0215)	(0.0148)	(0.0199)
High School Degree		0.0107		0.0186		0.0239		0.0187
		(0.0600)		(0.0547)		(0.0599)		(0.0558)
College Degree		0.2264		0.1526		0.1415		0.1486
		(0.0574)		(0.0392)		(0.0431)		(0.0357)
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. Demographic controls include 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 7 and 8, respectively, in Panel A, and 6 and 10 in Panel B. The number of factors is selected following the approach of Kim and Oka (2014). 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 and 6, respectively.

Table C.4: Decomposition Analysis - Without the Factor Structure

	(1) Linear Schooling, Quartic Age	(2) Quadratic Schooling, Quartic Age
Without Demographic Controls		
Overall Return to Schooling	0.0688	0.0739
(Total)	(0.0095)	(0.0114)
Human Capital Component	0.0264	0.0358
(D1)	(0.0118)	(0.0139)
Diploma Signaling Component	0.0424	0.0381
(D2)	(0.0076)	(0.0080)
from HS Degree	0.0016	0.0013
	(0.0015)	(0.0039)
from COLL Degree	0.0408	0.0368
	(0.0070)	(0.0063)
D1/Total	38%	48%
	(14%)	(14%)
D2/Total	62%	52%
	(14%)	(14%)

Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.
Note: This table presents the decomposition results obtained without incorporating the factor structure, as specified in equation (9). The overall returns to schooling are OLS estimates based on the baseline specification, which includes only schooling and quartic experience variables, while excluding degree indicators and the factor structure.

Table C.5: Decomposition Analysis - Without the Factor Structure

	(1)	(2)	(3)	(4)
	Linear Schooling		Quadratic Schooling	
	Age^2	Age^4	Age^2	Age^4
With Demographic Controls				
Overall Return to Schooling	0.0715	0.0653	0.0788	0.0699
(Total)	(0.0128)	(0.0128)	(0.0158)	(0.0159)
Human Capital Component	0.0286	0.0244	0.0391	0.0332
(D1)	(0.0165)	(0.0165)	(0.0192)	(0.0192)
Diploma Signaling Component	0.0428	0.0409	0.0397	0.0367
(D2)	(0.0108)	(0.0108)	(0.0113)	(0.0114)
from HS Degree	0.0026	0.0013	0.0034	0.0011
	(0.0024)	(0.0024)	(0.0059)	(0.0059)
from COLL Degree	0.0402	0.0396	0.0364	0.0357
	(0.0100)	(0.0100)	(0.0090)	(0.0090)
D1/Total	40%	37%	50%	47%
	(20%)	(22%)	(20%)	(24%)
D2/Total	60%	63%	50%	53%
	(20%)	(22%)	(20%)	(24%)

Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.
Note: This table presents the decomposition results obtained without incorporating the factor structure, as specified in equation (9). The overall returns to schooling are OLS estimates based on the baseline specification, which includes schooling and experience variables along with demographic controls, while excluding degree indicators and the factor structure. For example, the overall MRTS under the linear schooling and quadratic age specification, reported in Column (1), is 7.15%. This estimate corresponds to the MRTS result shown in Table C.1, Panel A, Column (1).

Table C.6: Decomposition Analysis - With the Factor Structure

	(1)	(2)	(3)	(4)	(5)	(6)
	Linear Schooling, Quartic Age			Quadratic Schooling, Quartic Age		
	IFE	CCE	CCE2	IFE	CCE	CCE2
Without Demographic Controls						
Overall Return to Schooling	0.0688	0.0688	0.0688	0.0739	0.0739	0.0739
(Total)	(0.0095)	(0.0095)	(0.0095)	(0.0114)	(0.0114)	(0.0114)
Human Capital Component	0.0043	0.0004	0.0024	0.0113	0.0046	0.0069
(C1)	(0.0120)	(0.0148)	(0.0147)	(0.0133)	(0.0216)	(0.0201)
Diploma Signaling Component	0.0294	0.0298	0.0295	0.0313	0.0252	0.0251
(C2)	(0.0069)	(0.0076)	(0.0070)	(0.0080)	(0.0096)	(0.0082)
from HS Degree	0.0001	0.0003	0.0001	0.0074	0.0031	0.0023
	(0.0019)	(0.0021)	(0.0022)	(0.0044)	(0.0059)	(0.0055)
from COLL Degree	0.0293	0.0295	0.0294	0.0240	0.0220	0.0228
	(0.0063)	(0.0071)	(0.0063)	(0.0057)	(0.0068)	(0.0005)
Ability Bias Component	0.0351	0.0386	0.0379	0.0312	0.0441	0.0447
(C3)	(0.0094)	(0.0130)	(0.0121)	(0.0100)	(0.0175)	(0.0161)
C1/Total	6%	1%	3%	15%	6%	9%
	(17%)	(22%)	(21%)	(18%)	(30%)	(27%)
C2/Total	43%	43%	42%	42%	34%	33%
	(10%)	(11%)	(10%)	(11%)	(13%)	(11%)
C3/Total	51%	56%	54%	42%	60%	58%
	(14%)	(20%)	(19%)	(14%)	(25%)	(23%)

Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.
Note: This table presents the decomposition results obtained when incorporating the factor structure, as specified in equation (7).

Table C.7: Decomposition Analysis - 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
With Demographic Controls												
Overall 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)	(0.0128)	(0.0128)	(0.0128)	(0.0128)	(0.0128)	(0.0128)	(0.0158)	(0.0158)	(0.0158)	(0.0159)	(0.0159)	(0.0159)
Human Capital Component	-0.0040	-0.0067	-0.0061	-0.0043	-0.0067	-0.0062	-0.0018	-0.0026	-0.0023	-0.0016	-0.0025	-0.0024
(C1)	(0.0116)	(0.0145)	(0.0146)	(0.0116)	(0.0145)	(0.0146)	(0.0156)	(0.0215)	(0.0199)	(0.0159)	(0.0215)	(0.0199)
Diploma Signaling Component	0.0299	0.0299	0.0301	0.0298	0.0299	0.0300	0.0257	0.0247	0.0255	0.0259	0.0246	0.0253
(C2)	(0.0076)	(0.0076)	(0.0072)	(0.0075)	(0.0075)	(0.0071)	(0.0091)	(0.0098)	(0.0084)	(0.0091)	(0.0096)	(0.0082)
from HS Degree	0.0004	0.0003	0.0002	0.0003	0.0003	0.0002	0.0022	0.0024	0.0019	0.0018	0.0023	0.0018
	(0.0024)	(0.0022)	(0.0023)	(0.0022)	(0.0021)	(0.0021)	(0.0055)	(0.0060)	(0.0056)	(0.0054)	(0.0059)	(0.0055)
from COLL Degree	0.0296	0.0296	0.0298	0.0295	0.0296	0.0298	0.0235	0.0223	0.0235	0.0241	0.0223	0.0234
	(0.0067)	(0.0071)	(0.0064)	(0.0068)	(0.0071)	(0.0064)	(0.0060)	(0.0068)	(0.0002)	(0.0062)	(0.0068)	(0.0003)
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)	(0.0128)	(0.0120)	(0.0087)	(0.0128)	(0.0120)	(0.0118)	(0.0171)	(0.0158)	(0.0121)	(0.0171)	(0.0157)
C1/Total	-6%	-9%	-9%	-7%	-10%	-10%	-2%	-3%	-3%	-2%	-4%	-3%
	(16%)	(21%)	(21%)	(18%)	(23%)	(22%)	(20%)	(27%)	(25%)	(23%)	(31%)	(29%)
C2/Total	42%	42%	42%	46%	46%	46%	33%	31%	32%	37%	35%	36%
	(11%)	(11%)	(10%)	(12%)	(12%)	(11%)	(12%)	(13%)	(11%)	(13%)	(14%)	(12%)
C3/Total	64%	68%	66%	61%	64%	63%	70%	72%	71%	65%	68%	68%
	(13%)	(19%)	(18%)	(14%)	(21%)	(20%)	(16%)	(23%)	(21%)	(18%)	(26%)	(24%)

Source: National Longitudinal Survey of Youth (NLSY79), 1981-2016.
Note: This table presents the decomposition results obtained when incorporating the factor structure (along with demographic controls), as specified in equation (7).