Interpreting Sheepskin Effects in the Returns to Education

Alfonso Flores-Lagunes
Department of Economics
University of Arizona
alfonso@eller.arizona.edu

Audrey Light
Department of Economics
Ohio State University
light.20@osu.edu

March 2007

We thank Lung-fei Lee and Ian Walker for useful discussions, Stephen Bronars for comments on an earlier draft, and Alex Shcherbakov for providing excellent research assistance. Flores- Lagunes gratefully acknowledges financial support from the Industrial Relations Section at Princeton University.

Abstract

Researchers often identify sheepskin effects by including degree attainment (*D*) and years of schooling (*S*) in a wage model, yet the source of independent variation in these measures is not well understood. We argue that *S* is negatively correlated with ability among degree-holders because the most able graduate the fastest, while a negative correlation exists among dropouts because the most able benefit from increased schooling. Using data from the NLSY79, we find that wages decrease with *S* among degree-holders and increase with *S* among dropouts. The independent variation in *S* and *D* needed for identification is not due to reporting error. Instead, we conclude that skill varies systematically among individuals with a given degree status.

JEL Classification: I21, J24.

Keywords: returns to schooling, returns to degree attainment, sheepskin effects, measurement error.

Alfonso Flores-Lagunes
Department of Economics
University of Arizona
alfonso@eller.arizona.edu

Audrey Light
Department of Economics
Ohio State University
light.20@osu.edu

1. Introduction

A central issue in labor economics is why credentialed workers (those with high school diplomas or college degrees) earn more than their non-credentialed counterparts. Such sheepskin effects are consistent with sorting models of education (Arrow 1973, Spence 1973, Stiglitz 1975, Weiss 1983) in which employers use credentials to identify workers with desirable traits that cannot be directly observed.[1] However, sheepskin effects are also generated by human capital models (Becker 1964, Card 1999) if good learners are the ones who stay in school long enough to earn credentials, or if "lumpiness" in the learning process leads to more skill acquisition in degree years than in preceding years (Chiswick 1973, Lange and Topel 2006). Despite difficulties in distinguishing between these two competing models, the "sorting versus human capital" debate has dominated the sheepskin literature for over 30 years.

Largely overlooked in this debate is the role of functional form in the interpretation of sheepskin effects. In the earliest empirical studies (Hungerford and Solon 1987, Taubman and Wales 1973), sheepskin effects were identified by including a nonlinear function of highest grade completed (S) or categorical measures of degree attainment (D) in a log-wage model. More recently, analysts have taken advantage of richer survey data to implement a different identification strategy: rather than include *S or D* in their wage models, they control for both *S and D* (Arkes 1999, Ferrer and Riddell 2002, Frazis 1993, Jaeger and Page 1996, Park 1999). The interpretation of the resulting sheepskin effects—defined as the wage gap between credentialed and non-credentialed workers *conditional* on years of schooling—is the focus of our analysis.

When both S and D are included in a wage model, sheepskin effects are identified because individuals with a given amount of schooling differ in their degree attainment or, stated differently, because years of schooling vary among individuals with a given degree. We begin by considering how individuals' schooling decisions could generate the necessary variation in S and D. Among existing human capital and sorting models, only Weiss's (1983) "sorting-cum-learning" model explains why S and D might vary independently.[2] In Weiss's model, individuals attend school for S years and then take a test. High-ability individuals pass the test and earn a

[1]Following Weiss (1995), we use the term "sorting models" to refer to both signaling and screening versions of the models.

[2]In other models (*e.g.,* Arrow 1973, Becker 1964, Card 1999, Spence 1973, Stiglitz 1975) schooling attainment is a one-dimensional construct; whether it is measured as highest grade completed or highest degree received is left to survey designers and data analysts.

degree, while low-ability individuals terminate their schooling without a degree. While this behavioral framework justifies the inclusion of *S* and *D* in a wage model, it is inconsistent with the fact that individuals take varying amounts of time to earn *identical* degrees.

Given the lack of compelling explanations for the type of variation in *S* and *D* seen in the data, we present a human capital model in which (*i*) individuals differ in ability, (*ii*) high-ability individuals acquire more skill than low-ability individuals during each year of school, and (*iii*) degrees are awarded once a given skill threshold is reached. In addition to predicting that high-ability individuals earn degrees and low-ability individuals do not, our model shows that ability is *negatively* correlated with time spent in school among degree-holders; everyone in this population reaches the same level of achievement, but the most able reach the threshold in the shortest time. Among individuals who do not earn a degree, however, those with the *most* ability stay in school the longest because they benefit from additional skill investments. The model can generate this independent, systematic variation in *S* and *D* regardless of whether students are uncertain about their abilities or the labor market returns to schooling.

Our schooling model provides a rationale for including both *S* and *D* in the wage function. In particular, it leads us to specify a wage function in which the *S* slope varies across degree categories, and it predicts that the *S* slope is negative for degree holders (*e.g*., college graduates) and positive for nondegree-holders (*e.g*., college dropouts). In contrast, earlier studies include independent dummy variables for each degree category (*D*) and for schooling (*S*) (Arkes 1999, Ferrer and Riddell 2002, Frazis 1993, Jaeger and Page 1996, Park 1999), or they specify a fully-interacted model with a dummy variable for every *S-D* cell (Jaeger and Page 1996, Park 1999). In the absence of an explicit theoretical justification for these functional forms, it is difficult to interpret the estimates.[3]

In estimating our wage model with data from the 1979 National Longitudinal Survey of Youth, we consider two additional issues. First, we acknowledge that the independent variation in *S* and *D* used for identification can arise from reporting errors as well as from the optimizing

[3]In fact, existing estimates appear to be highly sensitive to functional form. When controlling for (non-interacted) dummy variables for each degree level and each year of schooling, Jaeger and Page (1996) predict a gap in log wages of 0.16 between bachelor's degree holders and college dropouts. When controlling for dummy variables for every *S-D* cell, they predict the same gap in log wages (holding *S* constant at 16) to be 75% higher.

behavior described by our model. Because models that control for both *S* and *D* rely on variation in *S within* each degree category, the estimates are more vulnerable to "noise" than are estimates that rely on the total variation in the data. To contend with measurement error, we extend the switching regression model with imperfect regime classification information proposed by Lee and Porter (1984) to the case where a polychotomous variable (*D*) is potentially misclassified. We also reestimate our wage equations with *S* and *D* data that are judged to be "clean" to determine whether seemingly error-ridden observations are driving our results. While measurement error in both *S* and *D* has been widely explored (Ashenfelter and Krueger 1994, Black *et al.* 2000, Bound *et al*. 2001, Flores-Lagunes and Light 2006, Kane *et al*. 1999), we believe this is the first analysis to ask whether "noise" is an important source of the independent variation in *S* and *D* used to identify sheepskin effects.

Second, we argue that "highest grade completed" is not always the preferred measure of time spent in school. Among degree holders, we wish to know how long it takes to earn a given credential. However, time to completion may not be fully captured by "highest grade completed" if the latter measures credits earned toward a degree—for example, high school graduates may report having completed grade 12 even when they earn their diploma a year early. Among dropouts, where our goal is to measure the skill acquisition that takes place prior to school exit, "highest grade completed" (that is, progress made towards a degree) is likely to be a better measure than "time to school exit." In light of these concerns, we use both highest grade completed and age at school exit (conditional on work experience gained while in school) as alternative measures of *S* in our wage models.

Our estimates reveal that wage-*S* slopes vary across degree categories in the systematic manner predicted by our model. When measured as age at school exit, we find that each year of *S* is associated with 3% lower wages among high school and college graduates; when measured as highest grade completed, each year of *S* is predicted to raise wages by 1% among high school dropouts and 4% among college dropouts. These findings are largely invariant to measurement error in *S* and *D*. Instead, the independent variation in *S* and *D* observed in the data appears to reflect important skill differences among individuals with a common degree status. By recasting sheepskin effects (degree effects conditional on *S*) as "time in school" effects conditional on *D*, we learn that dropouts who stay in school the longest are the most highly skilled of their "type," as are graduates who complete their degrees in the shortest time.

2. Schooling Model

Our objective is to understand why time spent in school (*S*) varies among individuals with a given degree status (*D*). An obvious explanation is that some students take longer than others to complete a degree because they devote time to employment, either by combining part-time school with employment or by interrupting their enrollment to enter the labor market. In estimating our wage models, we include detailed measures of work experience to account for this phenomenon.[4] Therefore, our purpose here is to explain why time to school exit varies among students with identical levels of employment effort.

Our model is a straightforward extension of Becker's (1964) schooling model in which individuals terminate their schooling when the marginal benefit equals the marginal cost. The Becker model provides a single observed dimension—years of schooling (*S*)—in which to assess individuals: the more *S* a worker has, the more skill and ability he is expected to embody. To augment this framework, we simply recognize that the institutional process of awarding degrees introduces well-defined points at which individuals terminate their schooling. Our key assumptions are that acquired skill is *identical* among individuals who earn a given degree, and that high-ability individuals reach this skill threshold *faster* than low-ability individuals. Consider a situation where only one degree is available, so all labor market entrants are classified as graduates or dropouts. Dropouts make their schooling decision precisely as described in the Becker model; among these individuals, *S* and ability are positively correlated. Graduates terminate their schooling once they acquire enough skill to earn the degree; among this group, the most able earn their degrees in the shortest time.

Before proceeding to the details of our model, a few comments are in order. First, the preceding discussion assumes that individuals face no *ex ante* uncertainty about their own abilities and, therefore, about whether they are capable of earning the degree. In formalizing our model, we demonstrate that the predictions can continue to hold when we introduce uncertainty. Second, the assumption that acquired skill is identical among degree-holders abstracts from the fact that school characteristics and programs of study can affect how much is learned, conditional on student ability and time spent in school. An extensive literature explores the effects of school quality and college major on subsequent earnings (*e.g.*, Altonji *et al.* 2005,

---

[4]Light (1998, 2001) demonstrates that wage gains associated with schooling can be dramatically overestimated if in-school work experience is ignored.

Arcidiacono 2004, Brewer *et al.* 1999, Dale and Krueger 2002), but we do not attempt to incorporate those additional sources of heterogeneity in our analysis. Third, we use a human capital model, rather than a sorting model, to explain why years of school varies among individuals within a given degree category. We believe this human capital framework is informative, but we do not argue that it provides the *only* explanation for observed variation in the data. A sorting model could argue that students use both degree attainment *and* speed of completion (perhaps along with their school quality, grade point average, *etc*.) as a multi-dimensional signal of their ability (Groot and Oosterbeek 1994).

A. No uncertainty

We consider a group of individuals who are heterogeneous with respect to their innate ability, which is known *ex ante*. These individuals face a three-period lifespan during which they can pursue one of three paths: work for three periods, attend school for one period and work for the remaining two periods, or attend school for two periods and work for the final period. Time spent in school (S=0,1,2) combines with ability (A) to produce skill according to the simple production function $K=AS$. Once skill reaches the threshold $\overline{K}$, a degree is awarded. Thus, $\overline{S}=\overline{K}/A$ is the exogenous time needed to graduate, and individuals are destined to be dropouts if they lack the ability needed to graduate within two years (*i.e*., if $A<\overline{K}/2$). Given the discrete nature of the decision-making, we assume that A is distributed in such a way that individuals fall into one of three categories: $\overline{S}=1$, $\overline{S}=2$, or $\overline{S}>2$.[5]

After completing school, individuals' labor market earnings (Y) are given by the function

$$Y=\gamma A+\beta_S K, \tag{1}$$

where γ and β_S are the (known) rates at which innate skill and acquired skill are rewarded in the labor market. Given that a second year of school doubles an individual's skill level but decreases his work life from two periods to one, we must allow the second period of school to have a greater rate of return than the first $(\beta_2>\beta_1)$ to ensure that S=2 will be a feasible option for some individuals.[6] Equation (1) indicates that ability has both a direct effect on earnings and an indirect effect via its role in converting

[5]The assumption of a discrete ability distribution allows us to abstract from mid-period graduation dates and post-graduate schooling (including the option of earning a second degree), although these complexities will be incorporated in the empirical analysis.

[6]More generally, we maintain the assumption that $\beta_2>a\beta_1>b\gamma$, where the positive parameters a and b are large enough to ensure that the decision to stay in school is driven entirely by person-specific factors (*viz.*, ability and cost).

time spent in school into additional skill.

Given a direct, per-period cost of schooling of C and a discount rate of r, the present value of each investment option (no school, one year of school, or two years of school) is:

$$V_{s=0} = \gamma A + \frac{\gamma A}{1+r} + \frac{\gamma A}{(1+r)^2} \tag{2a}$$

$$V_{s=1} = -C + \frac{\gamma A}{1+r} + \frac{\gamma A}{(1+r)^2} + \frac{\beta_1 A}{1+r} + \frac{\beta_1 A}{(1+r)^2} \tag{2b}$$

$$V_{s=2} = -2C + \frac{\gamma A}{(1+r)^2} + \frac{\beta_2 2A}{(1+r)^2}. \tag{2c}$$

In the absence of uncertainty, there is no sequential decision-making involved. At the beginning of their lives, individuals simply choose the investment path that maximizes the present value of lifetime earnings.

For individuals with $A < \overline{K}/2$, the alternatives are to attend no school (*S*=0), dropout after one year, or dropout after two years. One year of school is preferred to zero years as long as $V_{s=1} \geq V_{s=0}$, or

$$A \geq \frac{C(1+r)^2}{\beta_1(2+r) - \gamma(1+r)^2}. \tag{3}$$

The denominator in equation (3) is positive as long as the labor market return to skills acquired in school is "large" relative to the return to ability that must be foregone while enrolled; following footnote 6, we assume this condition holds. Thus, conditional on market factors β_1, γ and r, the likelihood of attending school for one year increases in ability and decreases in out-of-pocket costs. As long as A and C are not positively correlated (which is highly implausible, given that merit-based financial aid and other factors are expected to produce a negative correlation), students who drop out after one year have more ability than those who do not attend school at all.

Dropouts choose to enroll for two years rather than one year if $V_{s=2} \geq V_{s=1}$, or

$$A \geq \frac{C(1+r)^2}{2\beta_2 - \beta_1(2+r) - \gamma(1+r)} \tag{4}$$

where, again, the denominator is positive as long as the second year of school has a sufficiently large labor market return relative to the return to existing skills that must be held out of the labor market for an additional year. If the denominator in equation (4) is smaller than the denominator in (3)—that is, if the discounted marginal return to schooling is smaller for the second year than for the first—then we predict that students who drop out after two years are more able than those who drop out a year earlier. In short, we obtain the same results as any orthodox schooling model with heterogeneous ability (Becker 1964, Card 1999): in the absence of a perverse situation in which the most able face the highest out-of-pocket costs, *time in school (s) is positively correlated with ability.*

Individuals in the intermediate portion of the ability distribution ($\overline{K}/2 \leq A < \overline{K}$ or, given our discrete ability distribution, $A = \overline{K}/2$) face the same three alternatives as do low-ability individuals, but they are able to earn a degree if they choose two years of school. Inequalities (3) and (4) describe the conditions that lead them to choose one year of school instead of no school, or two years instead of one. However, as long as the market factors $(\beta_2, \beta_1, \gamma, r)$ and costs (C) are not systematically less favorable for these individuals than for their less able counterparts, their decision is clear: if inequalities (3) and (4) are met for *anyone* with $A < \overline{K}/2$, then they are met for *everyone* with $\overline{K}/2 \leq A < \overline{K}$. Individuals with intermediate ability choose to stay in school for as long as the "best" dropouts, but unlike dropouts they leave school with skill level $\overline{K}$ and a degree.

The remaining individuals in the population are the most able ($A = \overline{K}$), and their alternatives are to attend no school or to attend for one year and earn a degree. Again, inequality (3) holds for *everyone* in this group unless they face an unusually high out-of-pocket expense, discount rate, or direct return to ability (γ). Thus, these individuals choose to attend school for one year, during which they acquire $\overline{K}$ units of skill and earn a degree. This gives us our final result: *among individuals who earn a degree, time in school (S) is negatively correlated with ability.*

B. Uncertainty

To illustrate how imperfect information affects schooling decisions, we focus on a situation where individuals are initially uncertain about their true ability (A). Specifically, we assume that perceived ability at the outset of the decision-making period is $A_0 = A + \varphi$, where φ is a mean-zero, constant-variance error term. We assume that uncertainty is resolved after one period—that is, after one year of "experiencing" one's ability in school or at work, the variance of φ falls to zero.

Rather than detail the sequential decision-making that arises in this uncertain environment, we simply consider the types of "mistakes" that can be made. In the absence of uncertainty, we found that individuals at the bottom of the ability distribution $(A < \overline{K}/2)$ sort themselves perfectly on (true) ability in deciding whether to enroll for zero, one, or two years before entering the labor market as dropouts. The presence of uncertainty can induce individuals with low A but an extremely large φ to drop out after one year, even though they would choose S=0 if true ability were known *ex ante*. Similarly, individuals with relatively high A for this population but a negative φ might choose $S=0$ instead of enrolling for one or two years before dropping out. In short, temporary uncertainty can cause a subset of high-ability dropouts to have the shortest enrollment durations and low-ability dropouts to stay in school the longest.

Individuals who are capable of earning a degree ($A \geq \overline{K}/2$) can contribute to the spillover of high-ability dropouts into the S=0 category. In the absence of uncertainty, we found that these individuals *always* choose to enroll in school and ultimately earn their degrees. The only "mistake" these individuals can make in the face of uncertainty is to choose $S=0$ as the result of an extreme, negative draw from the error distribution. Among those who enroll in school (all of whom graduate in either one or two years), our earlier result is unaffected: high-ability individuals earn their degrees faster than low-ability individuals. Thus, uncertainty regarding one's innate ability only affects the dropout population by causing some high-ability individuals to stay in school longer than some of their low-ability counterparts—but as long as extreme errors are uncommon, this type of spillover will not reverse the positive correlation between ability and enrollment duration among dropouts.[7]

3. Econometric strategy

In section 2, we assumed that individuals choose their enrollment duration (S) to maximize lifetime earnings that depend on innate ability (A) and acquired skill (K). Econometricians cannot observe A and K directly, but we *can* observe S and degree attainment

[7]Uncertainty about labor market returns to schooling (β_s) can also alter the dropout population. High-ability individuals who would normally earn degrees can potentially drop out of school if they underestimate the value of additional skill investments, and dropouts can leave school too soon or too late (relative to the "perfect information" optimum). However, we expect this type of uncertainty to have relatively little effect on the relationship between ability and enrollment duration as long as errors are not systematically correlated with ability.

(D). Under the assumptions of our model, these observed factors fully describe ability and skill. Table 1 summarizes the case with no uncertainty: among individuals without a degree (D=0), increasing S from 0 to 1 to 2 corresponds to monotonic increases in both A and K. Degree-holders (D=1) necessarily have more ability and more skill than those without a degree; everyone in this population has $\overline{K}$ units of skill, but those with S=1 have more A than those with S=2. Given this mapping between ability/skill and schooling attainment, we can specify a wage function that is equivalent to (1), but that conditions on observed factors S and D.

Our model calls for a functional form that allows wages to increase with each successive degree category and change with S within each category. Thus, we specify the following wage function:

$$Y_i = \sum_{k=1}^{4} \alpha_k D_{ki} + \sum_{k=1}^{4} \delta_k D_{ki} S_i + \pi Z_i + u_i, \qquad (5)$$

where Y_i is the natural logarithm of the average hourly wage for individual i, D_{ki} is a vector of dummy variables identifying four degree categories, S_i is enrollment duration, Z_i is a vector of additional covariates including cumulative labor market experience, and u_i represents unobserved factors. In contrast to our earlier, simplifying assumption that workers earn a single degree upon reaching a given skill threshold, we now allow for two successive degrees. Specifically, D_k distinguishes between high school dropouts (D_1), high school graduates (D_2), college dropouts (D_3) and college graduates (D_4).[8] Our model predicts that log-wages increase monotonically with each successive degree $(\alpha_1 < \alpha_2 < \alpha_3 < \alpha_4)$, and that the S slope is positive for dropouts $(\delta_1 > 0,\ \delta_3 > 0)$ and negative for degree-holders $(\delta_2 < 0, \delta_4 < 0)$.

It is worth reiterating that our specification is not generally used in the sheepskin literature. The orthodox approach—often dictated by a lack of independent data on S and D—is to use a spline function or step function in S and omit separate measures of D (Belman and Heywood 1991, Hungerford and Solon 1987, Lange and Topel 2006). Among studies that control for both S and D, most include degree dummies and an independent (noninteracted) function of S (Arkes 1999, Ferrer and Riddell 2002, Frazis 1993, Jaeger and Page 1996, Park 1999). This is equivalent to imposing the restriction $\delta_1 = \delta_2 = \delta_3 = \delta_4$, although some studies

[8]In section 4 we discuss our decision to use these four degree categories and describe the data used to estimate equation (5).

relax our restriction that *Y* is a linear function of *S*. Jaeger and Page (1996) and Park (1999) propose alternative specifications that allow for interactions between *S* and *D*, but they use an extremely flexible functional form that includes a parameter for every year-of-schooling/degree combination. We propose equation (5) as the most parsimonious way to capture the degree-specific intercepts and *S* slopes that are consistent with our schooling model.

As long as the covariates $D_{ki}, S_i,$ and Z_i are complete and accurate representations of the factors that determine wages in the labor market, we can use ordinary least squares (OLS) to estimate the parameters in equation (5). While we maintain the assumption that our covariates are sufficient statistics for innate ability and acquired skill (and that wages are based on these factors), we cannot assume that our survey data are reported without error. Unfortunately, the instrumental variables and generalized method of moments estimators that are often used to account for measurement error in *S* or *D* (Ashenfelter and Krueger 1994, Black *et al*. 2000, Flores-Lagunes and Light 2006, Kane *et al.* 1999) are inappropriate for our application because they allow only *one* covariate to be reported with error, and they require two, independent reports for the error-ridden variable.[9] In the absence of a tractable estimation strategy that accounts for measurement error in both *S* and *D*, we tackle them separately. First, we use an extension of the switching regression model with imperfect sample separation information proposed by Lee and Porter (1984) to account for errors in *D* only. As described in section 4, we then use a sample of seemingly "clean" responses to determine whether measurement error in *S* has a substantial effect on our estimates.

To account for the potential misclassification of degree status, we replace equation (5) with the following degree-specific equations:

$$Y_i = \alpha_k + \delta_k S_i + \pi Z_i + u_{ki} \qquad \text{for } k = 1,2,3,4 \tag{6}$$

where, as in (5), we allow the intercept and *S* slope (but not the *Z* slopes) to vary with degree status. We define the probability that a given wage belongs to "regime *k*" (*i.e.,* the worker's true degree status is *k*) as λ_k, where $\sum \lambda_k = 1$. Our information about the worker's "regime" (degree status) comes from the information he reports to the survey; we refer to this self-reported degree

[9]We have an independent report of high school graduation status for a subset of respondents for whom high school transcripts were collected, but the NLSY79 does not collect similar validation data for college attendance and degrees. Similarly, we have sibling-reported "highest grade completed" for respondents with in-sample siblings, but these reports pre-date final schooling attainment for many respondents.

status as D_i^R and to the "true" degree status as D_i^T; both variables have an integer value between one and four to correspond to our four-category degree taxonomy. We also define the elements of the four-by-four transition matrix describing the relationship between D_i^T and D_i^R as $p_{kl} = \Pr(D^R = l \mid D^T = k)$ for k,l=1,…4. We impose the identifying assumptions that measurement errors in self-reported degree status $(D_i^R - D_i^T)$ are independent of the wage-equation errors u_{ki} conditional on D_i^T, and that the four u_{ki} are normally distributed, mean zero, and equal variance.[10]

Given these definitions and assumptions, we can specify a likelihood function that depends on the probability density function of log-wages in each degree category, the degree probabilities (λ_k), and the misclassification probabilities (p_{kl}). We then use the expectation-maximization (EM) algorithm to produce maximum-likelihood estimates of the parameters of our model. This strategy is a straightforward extension of Lee and Porter (1984), who use a two-regime model (k=1,2). It produces estimates for the degree-specific intercepts (α_k), degree-specific S slopes (δ_k), and additional coefficients (π) that are comparable to what is obtained with OLS, but that account for possible discrepancies between the degree status known to us and the "true" status.[11] As noted earlier, we maintain the assumption that S and Z are reported without error, and informally test the assumption about S by comparing estimates based on a "clean" sample with estimates that include seemingly-erroneous reports of S and D.

4. Data

A. Sample Selection and Variable Definitions

We use data from the 1979 National Longitudinal Survey of Youth (NLSY79) to estimate the wage functions described in section 3. The NLSY79 began in 1979 with a sample of 12,686 youth born in 1957-64, and it remains in progress today. Respondents were interviewed annually from 1979 to 1994 and biennially thereafter; 2004 is the last year for which

---

[10]The equal variance assumption can be relaxed, but it is a useful empirical restriction that aids with convergence in estimation (see Morduch and Stern 1997). Given our generalization to multiple regimes, we also impose the restriction $p_{14} = p_{41} = 0$. This restriction is also used in Douglas *et al.* (1995).

[11]Intuitively, maximizing the likelihood function of this imperfect (degree) information model amounts to using the optimal rule for classifying observations into degrees given the available information (Lee and Porter 1984).

we have data. The NLSY79 is an ideal source of data for our analysis because respondents report their highest grade completed, dates of enrollment, and degree attainment; the survey also provides unusually detailed information on labor market activities, which enables us to net out the potentially confounding effects of work experience gained while in school. Details on the survey can be found in Center for Human Resource Research (2004).

The first step in our data creation process is to identify each respondent's chronological sequence of diplomas and degrees received, along with the date when each credential was awarded. If an individual attends high school, college, or graduate school without earning a diploma or degree, we include his attendance spell and dropout date in the degree sequence. We use this information to place each respondent into one of four categories: high school dropout, high school graduate, college dropout, or college graduate (*i.e.,* bachelor's degree recipient). We do not form post-college categories because our theoretical model assumes that individuals holding a given degree are homogeneous with respect to acquired skill, and we lack the sample size to define separate categories for holders of master's, professional, and doctoral degrees. In addition, we are unable to identify the degree programs being pursued by graduate school dropouts.[12] We also choose not to segment the college dropout category into two-year college dropouts, two-year college graduates, and four-year college dropouts because these groups are conceptually indistinct, given the frequent use of two-year colleges as "stepping stones" to four-year colleges (Hilmer 1997, Light and Strayer 2004, Rouse 1995). For example, we would hesitate to argue that a student who earns an associate's degree in two years and then spends one year at a four-year college differs in ability from a student who enrolls at a four-year college for three years. We substantiate this decision in section 5 by demonstrating that eliminating two-year degree-holders from the sample does not significantly affect our estimates. We also show that our inferences do not depend on whether we treat individuals who pass the general educational development test (GED) as high school dropouts (our default classification) or high school graduates.

After classifying each sample member with respect to highest degree (D), the second step is to identify the first wage earned after the degree is awarded or the individual drops out of school. Because college graduates can attend graduate school after receiving their bachelor's

[12]Individuals who attend graduate school remain in our sample as college graduates if they are nonenrolled for at least six months between college and graduate school, and report a wage during the interim.

degree, we require individuals to be nonenrolled when the wage is reported (see footnote 12). The log of this initial, post-school average hourly wage, divided by the consumer price index (CPI-U), is the dependent variable used to estimate our wage models.

The third step is to define the remaining covariates to be included in our wage model. During each interview, respondents are asked to report their current "highest grade completed" if they have attended school since the last interview. We use the value reported during the first interview after the degree or dropout date as one measure of "time in school," which we now refer to as $S1$. As an alternative measure of "time in school," we use the degree or dropout date in conjunction with the respondent's birth date to determine the age (measured to the nearest month) at which he or she left school; we refer to this variable as $S2$. We also define a set of dummy variables indicating the calendar year during which the wage was earned, and a dummy variable indicating whether the respondent is male. After experimenting with alternative demographic compositions, we opt to use a pooled sample of men and women but restrict the sample to individuals who are "white," by which we mean neither black nor Hispanic. It is common in the sheepskin literature to focus on white men (Hungerford and Solon 1987) or to estimate separate wage models for different race-sex groups (Belman and Heywood 1991, Jaeger and Page 1996). We find that our estimates are largely insensitive to whether we focus on white men or a combined sample of white men and women, but they change quantitatively when nonwhites are added to the sample. Because an examination of race differences in sheepskin effects is beyond the scope of our study, we simply eliminate blacks and Hispanics from our sample.

In addition to D, $S1$ or $S2$, calendar year and gender, we also control for cumulative labor market experience. We use each respondent's first-reported post-school wage as the dependent variable, so experience varies much less than it would if we measured wages later in the career. Nonetheless, it is essential that we accurately measure *in-school* work experience in order to "net out" this source of variation in enrollment duration. We use the detailed work history information available in the NLSY79 to construct a measure of cumulative weeks worked from the 18th birthday to the date when the wage is earned. In addition, we use the work experience reported by 16 and 17 year olds (which is reported only by respondents who are younger than 18 when the survey begins) to compute average work effort at age 16-17 as a fraction of work effort at age 18, by degree status. We then use these averages to assign every sample member a measure of predicted, early experience. We control for actual experience since age 18 and its square, along with predicted "early" (pre-age 18) experience in each of our wage models.

This data construction process produces a sample containing a single wage observation for each of 5,153 sample members. Our sample excludes the 5,176 blacks and Hispanics among the original 12,585 NLSY79 respondents, as well as 2,357 whites for whom the degree status is unknown or a post-school wage is unavailable; only 138 of these individuals are excluded because they enroll in graduate school after receiving a bachelor's degree.

In addition to analyzing the entire sample of 5,153 whites, we also examine a subsample in which the schooling and degree variables are judged to be "clean." To construct this subsample, we exploit the fact that degree attainment and highest grade completed should conform to certain institutional norms if respondents consider their progress toward a degree when reporting their highest grade completed. We expect high school dropouts to complete grade 11 or lower, high school graduates to complete grade 12, college dropouts to complete at least grade 12 but less than grade 16, and college graduates to complete grade 16. In forming a "clean" sample, we eliminate individuals if their reported S-D combination is sufficiently far from these expectations: we require $S \leq 12$, S=11-13, S=12-16, and S=15-19 for respondents in the high school dropout, high school graduate, college dropout, and college graduate categories, respectively. This strategy eliminates 3% of observations in each high school category, 4% in the college dropout category, and 2% in the college graduate category. The remaining sample consists of 4,998 individuals for whom the S and D data are not necessarily error-free, but are invariably less error-ridden than the data in the larger sample. By comparing estimates for our two samples, we can assess the effect of measurement error on the estimates.[13] However, we do not construct a similar "clean" sample using our alternative measure of S (age at school exit) because part-time and discontinuous enrollment often delay school exit. We control for these delays by including detailed experience measures in our wage model, but we lack clear expectations of the unconditional relationship between age at school exit and degree attainment.

B. Summary Statistics

---

[13]Our strategy is less demanding of the data than those requiring validation data (*e.g.*, Freeman 1984), and more flexible than those requiring relatively simple functional forms in order to jointly estimate measurement error and outcome models (Black *et al*. 2000, Flores-Lagunes and Light 2006, Kane *et al*. 1999).

Table 2 contains sample means and standard deviations for the variables used to estimate the wage models. Table 3 contains a cross-tabulation of "highest grade completed" ($S1$) and degree status (D). It is clear from these distributions that $S1$ varies considerably within D category. A comparison of the coefficient of variation across columns reveals that $S1$ varies far more within each dropout category than within each degree category; this conforms to the notion that $S1$ measures progress made towards a degree and is therefore relatively homogenous among degree-holders. However, if we instead ask how often $S1$ falls within the "expected" range for that particular degree category (less than 12 for high school dropouts, 12 for high school graduates, 13-15 for college dropouts and 16 for college graduates) we find the most noise among the college-goers: $S1$ is "as expected" for 92-93% of individuals in each high school category but only 87% of individuals in each college category. These patterns suggest that there is ample variation with which to identify independent wage effects of $S1$ within each degree category—especially among dropouts—but that misreporting might be a particularly important source of this variation among college-goers.

Table 4 replicates table 3, except we now use "age at school exit" ($S2$) in place of highest grade completed. Using the coefficients of variation for comparison, it is clear that this alternative measure of time in school varies far more within degree category than does highest grade completed. While there is no "expected" age at which individuals complete each degree category, given that school exit dates can be extended by part-time or interrupted enrollment, it is interesting to note that only 75% of high school dropouts leave school by age 18, only 66% of high school graduates earn their degrees at age 18, and only 35% of college graduates earn their degrees at age 22. In short, age at school exit diverges from "$S1 + 6$" as degree attainment increases. The estimates presented in section 5 will reveal whether the marginal wage effects of $S1$ and $S2$ differ once we condition on the in-school work experience that explains much of the divergence in these two measures.

5. Findings

Table 5 presents estimates of the degree-specific intercepts (α_k) and S-slopes (δ_k) for a variety of wage model specifications, all of which use "highest grade completed" ($S1$) as our measure of time in school; additional estimates for each specification are in table A1. We discuss these estimates before turning to the corresponding estimates in tables 6 and A2 in which $S1$ is replaced with age at school exit ($S2$).

Column 1 of table 5 contains OLS estimates of the model given by equation 5 in section 3. For comparison, column 1′ contains estimates of a model that restricts the four degree-specific slope coefficients to be equal. The column 1′ specification is representative of much of the existing literature (Arkes 1999, Ferrer and Riddell 2002, Frazis 1993, Jaeger and Page 1996) in which the goal is simply to identify degree effects conditional on $S1$. Based on the column 1′ estimates, we would predict that the gap in log wages between high school graduates and high school dropouts is 0.04, the corresponding gap among college graduates and dropouts is 0.15, and an additional year of school is associated with a 2% wage boost for all workers. When we allow the relationship between $S1$ and log wages to vary across degree categories (column 1), we estimate sheepskin effects that are almost twice as large as what is seen in column 1′ (0.07 for a high school degree and 0.28 for a college degree), and we reject at a 1% significance level the null hypothesis that the slope coefficients are equal across degree categories.

Moreover, the estimates in column 1 provide moderate support for our theoretical argument that wages increase (decrease) with time in school among dropouts (graduates). The estimated slope coefficients are 0.042 for the college dropout category and an imprecisely estimated -0.012 for college graduates; these point estimates are consistent with the notion that time in school is positively correlated with skill for dropouts but negatively correlated with innate ability for degree-holders. In fact, the estimated intercepts and slopes lead us to predict that college dropouts and graduates who complete 18 years of school will earn the same wage, assuming they are identical in other dimensions. For high school dropouts and graduates, however, the estimated slope coefficients are both essentially zero. As shown in table 3, high school graduates have the least variation in $S1$ of any degree category, and the least deviation from their "expected" level of $S1$. It appears that we lack enough variation in this particular schooling measure to identify a wage difference between "early" high school graduates and their "on time" or "late" graduating counterparts. While the estimated slope coefficient for high school dropouts is positive, as predicted, we do not find a significant wage boost for workers who advance from, say, 10 to 11 years of school. It may simply be the case that pre-market skills are relatively unimportant for the types of jobs held by high school dropouts.

In columns 3 and 4 of table 5, we assess the effects on our OLS estimates of reclassifying certain degree types. In column 3, we move GED recipients from the high school dropout category to the high school graduate category. This dramatically increases the predicted log-wage gap associated with earning a high school degree (doubling it from 0.07 in column 1 to 0.14 in column 3) but has no effect on the slope coefficients. In column 4, we eliminate two-year degree holders from the sample rather than include them in the college dropout category. Eliminating these relatively high wage earners leads to a slight decrease in the estimated intercept and slope coefficient for college dropouts, but does not qualitatively affect our findings.

Our next task is to assess the effects of measurement error on our estimates. In column 2 of table 5, we report maximum likelihood estimates for our version of the Lee-Porter switching regressions model described in section 3.[14] The estimated intercepts and coefficients based on this estimation strategy are virtually identical to what we obtain using OLS (column 1). Moreover, as reported in table A1, the estimated probability that a reported degree is the "true" degree is 0.96 or higher for all degree categories. These estimates are consistent with the finding of Kane *et al.* (1999) that survey respondents tend to report their degree status with a high degree of accuracy. As an alternative strategy for assessing the effects of measurement error, we compute OLS estimates using a "clean" sample that excludes observations where the reported *S-D* combination is highly implausible (*e.g.,* no high school diploma but S=16, or S=12 and a bachelor's degree). The differences between these estimates (shown in columns 5′ and 5) and the corresponding estimates in columns 1′ and 1 are not significantly different from zero. Despite the fact that the standard errors increase (as expected) when we switch to the clean sample, it is worth noting that each estimated slope coefficient in column 5 has the sign predicted by our model.

Next, we ask how our estimates change when we replace "highest grade completed" ($S1$) with "age at school exit" ($S2$) as our measure of time in school. Table 6 contains estimates for wage model specifications that use this alternative measure, but are otherwise identical to the corresponding specifications 1-4 in table 5. Specifications 5 and 5′ are omitted from table 6 because we lack priors on the unconditional relationship between degree and age at school exit needed to select a "clean" sample.

In comparing the models shown in columns 1′ and 1 of table 6, we again find that the data reject at a 1% significance level the equal slope restriction imposed by specification 1′. Using the preferred specification 1, we find that replacing "highest grade completed" with "age at school exit" has little effect on the estimated intercepts (sheepskin effects), although we now

[14] For this model, the standard errors are obtained using 300 bootstrap replications.

predict a larger wage gap between college dropouts and high school graduates than what is seen in table 5. However, switching schooling measures has a dramatic effect on the estimated slope coefficients, all of which are significantly smaller when we use *S*2 than when we use *S*1. In column 1 of table 5, we saw that the estimated slope coefficient is positive (and significant) for college dropouts and negative (but insignificant) for college graduates; we obtain the same signs in table 6, but now the estimated slope coefficient for dropouts is zero (0.004 with a standard error of 0.005) while the estimated slope coefficient for graduates is a precisely estimated -0.026. For the two high school degree categories, the estimated slope coefficients are essentially zero in table 5, but negative and statistically significant in table 6. When we reclassify GED recipients (column 3), eliminate two-year degree holders (column 4) or use a switching regressions model to account for misclassification of degree status (column 2), our estimates remain largely unchanged relative to the column 1 estimates.

To understand why our estimates are sensitive to the manner in which we measure time in school, it is useful to consider the two degree categories (high school and college graduates) separately from the two dropout categories. Even if holders of a given degree do not have *identical* levels of acquired skill, as assumed by our theoretical model, they complete similar programs and earn a similar number of credits. Consider one individual who earns a high school diploma at age 18, and another who earns the same diploma a year earlier. Both should report their highest grade completed as 12 to reflect the fact that they completed the 12th and final year of their program, but their reported school exit date should differ because one of them completed the program more quickly than the other. In short, "age at school exit" is a more informative measure of what we wish to know about degree recipients—namely, the speed with which they complete a common grade or degree program. Therefore, it comes as no surprise that the estimated *S*2 slopes in table 6 (based on "age at school exit") predict that degree-holders earn approximately 3% less for every extra year they take to graduate, whereas the estimated slope coefficients in table 5 (based on "highest grade completed") lack statistical significance.

In contrast, we believe "highest grade completed" is a more informative measure than "age at school exit" for both dropout categories. Our goal is to measure the amount of school completed (*i.e.*, credits earned) in order to control for heterogeneity in skill among individuals with a given nondegree status. If reported accurately, "highest grade completed" is likely to be the preferred measure of schooling attainment for this purpose, given that future dropouts may "drag out" the time to school exit by failing courses, being truant, and otherwise spending time neither learning nor acquiring work experience. We believe the estimated slope coefficients in table 5, which imply that college dropouts earn 4% more for every year spent in school, are preferred for assessing the effects of time spent in school among dropouts. As noted earlier, the fact that high school dropouts do not appear to earn more when they complete an additional grade is consistent with the notion that pre-market skill is not important for the jobs they typically hold.

We can offer additional evidence to substantiate the argument that "age at school exit" is the preferred measure for degree-holders in the sense that it measures (inverse) innate ability, whereas "highest grade completed" is the preferred measure of skill and ability among dropouts. In 1980, over 90% of NLSY79 respondents completed the Armed Services Vocational Aptitude Battery (ASVAB). NLSY79 users have access not only to individual ASVAB scores, but also to scores for the Armed Forces Qualification Test (AFQT), which are computed from respondents' scores on four parts of the ASVAB (word knowledge, paragraph comprehension, arithmetic reasoning, and mathematics knowledge). Because AFQT scores are considered to be good measures of pre-market skills (Neal and Johnson 1996), we assess their correlation with both measures of time spent in school for each degree-specific subsample.[15] Among individuals in both degree categories, age-adjusted AFQT scores are strongly, negatively correlated with "age at school exit" but not with "highest grade completed." Within both dropout categories, age-adjusted AFQT scores are strongly, positively correlated with "highest grade completed," whereas the correlations with "age at school exit" are small and positive for high school dropouts and zero for college dropouts.

6. Concluding Comments

Our analysis begins with the observation that researchers often identify sheepskin effects by controlling for both degree attainment (*D*) and years of schooling (*S*) in a wage model, despite the lack of compelling explanations for why these two measures of schooling attainment would vary independently. We argue that individuals with a given degree are roughly homogenous with respect to acquired skill, but because the more able can earn their degrees relatively quickly, *S* is negatively correlated with innate ability among this population. Conversely, individuals

---

[15] Respondents' ages ranged from 16 to 23 when the ASVAB was administered in 1980, so we use deviations between raw scores and age-specific means.

who drop out of a given degree program vary considerably with respect to both innate ability and acquired skill, and *S* is positively correlated with these traits. Our simple extension of Becker's schooling model justifies the inclusion of both *D* and *S* in a wage model, and suggests that the schooling slopes should be allowed to differ across degree categories.

In estimating wage models that control for both *D* and *S* using data from the 1979 National Longitudinal Survey of Youth, we identify a number of patterns that are consistent with our model. First, we find that the data resoundingly reject the restriction that schooling slopes are equal across degree categories—in other words, it is important to include degree dummies (*D*) and *S-D* interactions, rather than simply controlling for *S* and *D*. Second, we find that estimated sheepskin effects—*e.g.,* predicted log-wage gaps between high school graduates and high school dropouts, and between college graduates and college dropouts—are considerably larger when we allow *S* slopes to differ with degree status than when we do not. Third, when we measure *S* as "age at school exit," we predict that each additional year of school lowers wages by about 3% among high school and college graduates. When we measure *S* as "highest grade completed," an additional year of school is predicted to raise wages by 4% among college dropouts, but only by an imprecisely estimated 1% among high school dropouts. These findings are consistent with our model, although the near-zero schooling slope for high school dropouts suggests that they tend to hold jobs where skill acquired via additional schooling is not valued. Fourth, our estimates prove to be largely invariant to our attempts to account for measurement error in self-reported *S* and *D*, which suggests that the independent variation in these two dimensions of schooling attainment is not dominated by "noise."

The fact that our alternative measures of time spent in school ("age at school exit" and "highest grade completed") appear to capture different information for degree holders and dropouts is a useful finding in its own right. We argue that high school and college graduates are "expected" to complete grades 12 and 16, respectively, and that, as a result, the age at which they earn their degrees is a more informative measure of ability than is their highest grade completed. Conversely, "highest grade completed" is a useful measure of progress made toward a degree among dropouts, whereas variation in "age at school exit" might reflect time spent neither gaining work experience (which we are able to control for separately) nor learning. Our estimates suggest that individuals within a given degree or dropout category vary considerably with respect to their ability and/or acquired skill, and that additional measures of schooling attainment are useful for explaining variation in post-school wages. While ours is not the first study to view schooling attainment as a multi-dimensional construct, we suspect there is far more to be learned by exploring heterogeneity in completion patterns among individuals with a given level of schooling attainment.

References

Altonji, Joseph G., Todd E. Elder and Christopher R. Taber. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113 (February 2005): 151-184.

Arcidiacono, Peter. "Ability Sorting and the Returns to College Major." *Journal of Econometrics* 121 (July/August 2004): 343-375.

Arkes, Jeremy. "What Do Educational Credentials Signal and Why Do Employers Value Credentials?" *Economics of Education Review* 18 (February 1999): 133-141.

Arrow, Kenneth J. "Higher Education as a Filter." *Journal of Public Economics* 2 (July 1973): 193-216.

Ashenfelter, Orley and Alan Krueger. "Estimates of the Economic Returns to Schooling from a New Sample of Twins." *American Economic Review* 84 (December 1994): 1157-1173.

Becker, Gary. *Human Capital; a Theoretical and Empirical Analysis, with Special Reference to Education*. New York: Columbia University Press (for the National Bureau of Economic Research), 1964.

Belman, Dale and John S. Heywood. "Sheepskin Effects in the Returns to Education: An Examination of Women and Minorities." *Review of Economics and Statistics* 73 (November 1991): 720-24.

Black, Dan A., Mark C. Berger, and Frank A. Scott. "Bounding Parameter Estimates with Nonclassical Measurement Error." *Journal of the American Statistical Association* 95 (September 2000): 739-748.

Bound, John, Charles Brown, and Nancy Mathiowetz. "Measurement Error in Survey Data." In James J. Heckman and Edward E. Leamer (editors), *Handbook of Econometrics*, Volume 5. Amsterdam: Elsevier Science, 2001.

Brewer, Dominic J., Eric R. Eide and Ronald G. Ehrenberg. "Does It Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings." *Journal of Human Resources* 34 (Winter 1999): 104-123.

Card, David. "The Causal Effect of Education on Earnings." In Orley Ashenfelter and David Card (editors), *Handbook of Labor Economics,* Volume 3. Amsterdam: Elsevier Science B.V., 1999.

Center for Human Resource Research. *NLSY79 User's Guide*. 2004. Columbus, OH: The Ohio State University, CHRR NLS User Services.

Chiswick, Barry. "Schooling, Screening, and Income." In Lewis Solmon and Paul Taubman (editors), *Does College Matter?* New York: Academic Press, 1973.

Dale, Stacy Berg and Alan B. Krueger. "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables." *Quarterly Journal of Economics* 117 (November 2002): 1491-1527.

Douglas, Stratford, Karen S. Conway, and Gary Ferrier. "A Switching Frontier Model for Imperfect Sample Separation Information: With an Application to Constrained Labor Supply." *International Economic Review* 36 (May 1995) 503-526.

Ferrer, Ana M. and Craig W. Riddell. "The Role of Credentials in the Canadian Labour Market." *Canadian Journal of Economics* 35 (November 2002): 879-905.

Flores-Lagunes, Alfonso and Audrey Light. "Measurement Error in Schooling: Evidence from Samples of Siblings and Identical Twins." *Contributions to Economic Analysis and Policy* 5 (issue 1, 2006).

Frazis, Harvey. "Selection Bias and the Degree Effect." *Journal of Human Resources* 28 (Summer 1993): 538-554.

Freeman, Richard B. "Longitudinal Analyses of the Effects of Trade Unions." *Journal of Labor Economics* 2 (January 1984): 1-26.

Groot, Wim and Hessel Oosterbeek. "Earnings Effects of Different Components of Schooling: Human Capital Versus Screening." *Review of Economics and Statistics* 76 (May 1994): 317-321.

Hilmer, Michael J. 1997. "Does Community College Attendance Provide a Strategic Path to a Higher Quality Education?" *Economics of Education Review* 16(1): 59-68.

Hungerford, Thomas and Gary Solon. "Sheepskin Effects in the Returns to Education." *Review of Economics and Statistics* 69 (February 1987): 175-77.

Jaeger, David A. and Marianne E. Page. "Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education." *Review of Economics and Statistics* 78 (November 1996): 733-740.

Kane, Thomas J., Cecilia Elena Rouse and Douglas Staiger. "Estimating Returns to Schooling When Schooling is Misreported." NBER Working Paper 7235, July 1999.

Lange, Fabian and Robert Topel. "The Social Returns to Education and Human Capital." In Eric Hanushek and Finis Welch (editors), *Handbook of the Economics of Education,* Volume 1. Amsterdam: North-Holland, 2006.

Lee, Lung-fei and Robert H. Porter. "Switching Regression Models with Imperfect Sample Separation Information with an Application to Cartel Stability." *Econometrica* 52 (March 1984): 391-418.

Light, Audrey. "Estimating Returns to Schooling: When Does the Career Begin?" *Economics of Education Review* 17 (February 1998): 31-45.

———. "In-School Work Experience and the Returns to Schooling." *Journal of Labor Economics* 19 (January 2001): 65-93.

Light, Audrey and Wayne Strayer. "Who Receives the College Wage Premium? Assessing the Labor Market Returns to Degrees and College Transfer Patterns." *Journal of Human Resources* 39 (Summer 2004): 746-773.

Morduch, Jonathan J. and Hal S. Stern. "Using Mixture Models to Detect Sex Bias in Health Outcomes in Bangladesh." *Journal of Econometrics* 77 (1997): 259-276.

Neal, Derek A. and William R. Johnson. "The Role of Pre-Market Factors in Black-White Wage Differences." *Journal of Political Economy* 104 (October 1996): 869-895.

Park, Jin-Heum. "Estimation of Sheepskin Effects Using the Old and the New Measures of Educational Attainment in the Current Population Survey." *Economics Letters* 62 (February 1999): 237-240.

Rouse, Cecilia Elena. 1995. "Democratization or Diversion? The Effect of Community Colleges on Education Attainment." *Journal of Business and Economic Statistics* 13(2):217-24.

Spence, Michael. "Job Market Signaling." *Quarterly Journal of Economics* 87 (August 1973): 355-374.

Stiglitz, Joseph E. "The Theory of 'Screening,' Education, and the Distribution of Income." *American Economic Review* 65 (June 1975): 283-300.

Taubman, Paul J. and Terence J. Wales. "Higher Education, Mental Ability, and Screening." *Journal of Political Economy* 81 (January/February 1973): 28-55.

Weiss, Andrew. "A Sorting-cum-Learning Model of Education." *Journal of Political Economy* 91 (June 1983): 420-442.

———. "Human Capital vs. Signaling Explanations of Wages." *Journal of Economic Perspectives* 9 (Fall 1995): 133-154.

Table 1: Relationship between Observed Measures of Schooling Attainment and Unobserved Ability and Skill

	Variables that are observed by employers		Variables that are unobserved by employers	
Worker type	Degree status (D)	Time in school (S)	Innate ability (A)	Acquired skill (K)
1	0	0	A_1	0
2	0	1	$A_2 > A_1$	A_2
3	0	2	$A_3 > A_2$	$2A_3$
4	1	2	$A_4 = \overline{K}/2 > A_3$	$\overline{K}$
5	1	1	$A_5 = \overline{K}$	$\overline{K}$

Note: These relationships are predicted by the "no uncertainty" schooling model presented in section 2.

Table 2: Summary Statistics for Selected Variables

Variable	Mean	S.D.
Ln(average hourly wage)	2.01	.57
Highest grade completed (*S*1)	12.95	2.17
Age at school exit (*S*2)	20.17	3.15
Degree		
High school dropout (D_1)	.11	
High school diploma (D_2)	.42	
College dropout (D_3)	.28	
College graduate (D_4)	.19	
Actual experience[a]	.99	1.97
Actual experience squared	4.85	24.96
Early experience[b]	.06	.19
1 if male	.50	
Number of individuals	5,153	

[a]Hours worked from 18th birthday to date wage was earned, divided by 2,000.
[b]Estimated hours worked between 16th and 18th birthday.
Note: Sample consists of whites only.

Table 3: Highest Grade Completed by Highest Degree Received

Highest grade completed	High school dropout	High school graduate	College dropout	College graduate	All degree levels
2-8	119 (21.1)	4 (0.2)	1 (0.1)	1 (0.1)	125 (2.4)
9-10	258 (45.8)	57 (2.6)	4 (0.3)		319 (6.2)
11	139 (24.7)	77 (3.5)	6 (0.4)		222 (4.3)
12	30 (5.3)	2,035 (93.4)	339 (23.9)	4 (0.4)	2,408 (46.7)
13	6 (1.1)	4 (0.2)	377 (26.5)		387 (7.5)
14	9 (1.6)	1 (0.1)	381 (26.8)	6 (0.6)	397 (7.7)
15	1 (0.2)		138 (9.7)	7 (0.7)	146 (2.8)
16	1 (0.2)		139 (9.8)	864 (87.2)	1,004 (19.5)
17			17 (1.2)	69 (7.0)	86 (1.7)
18-20			19 (1.3)	40 (4.0)	59 (1.1)
All grades [% row total]	563 [10.9]	2,178 [42.3]	1,421 [27.6]	991 [19.2]	5,153
Coefficient of variation	16.28	4.13	10.52	4.48	16.73

Note: The table shows the number of sample members reporting each *S*-*D* combination. Percents of column totals are in parentheses.

Table 4: Age at School Exit by Highest Degree Received

Highest grade completed	High school dropout	High school graduate	College dropout	College graduate	All degree levels
12-16	165 (29.3)	17 (0.78)			182 (3.5)
17	131 (23.3)	94 (4.3)	3 (0.2)		228 (4.4)
18	126 (22.4)	1,426 (65.5)	71 (5.0)	4 (0.4)	1,627 (31.6)
19	36 (6.4)	507 (23.3)	289 (20.3)		832 (16.2)
20	14 (2.5)	77 (3.5)	314 (22.1)	4 (0.4)	409 (7.9)
21	9 (1.6)	13 (0.6)	250 (17.6)	22 (2.2)	294 (5.7)
22	5 (0.9)	7 (0.3)	179 (12.6)	344 (34.7)	535 (10.4)
23	5 (0.9)	7 (0.3)	126 (8.9)	383 (38.7)	521 (10.1)
24+	72 (12.8)	30 (1.4)	189 (13.3)	234 (23.6)	525 (10.2)
All ages [% row total]	563 [10.9]	2,178 [42.3]	1,421 [27.6]	991 [19.2]	5,153
Coefficient of variation	23.62	7.09	14.87	7.79	15.59

Note: The table shows the number of sample members reporting each *S-D* combination. Percents of column totals are in parentheses.

Table 5: OLS and Maximum Likelihood Estimates of Alternative Wage Models Using "Highest Grade Completed" as Measure of Time in School

	Full sample				Drop 2-year degrees	Clean sample	
Variable	(1′)	(1)	(2)	(3)[a]	(4)[a]	(5′)	(5)
Degree category							
High school dropout (D_1)	-.027	-.079	-.079	-.150	-.080	-.027	-.103
	(.029)	(.042)	(.061)	(.052)	(.042)	(.032)	(.048)
High school graduate (D_2)	.011	-.009	-.007	-.006	-.017	.009	-.021
	(.017)	(.025)	(.029)	(.023)	(.024)	(.017)	(.049)
College dropout (D_3)	.057	.045	.043	.047	.024	.053	.043
	(.018)	(.018)	(.016)	(.018)	(.019)	(.018)	(.018)
College graduate (D_4)	.204	.320	.320	.322	.314	.207	.304
	(.026)	(.063)	(.078)	(.063)	(.064)	(.029)	(.091)
Time in school ($S1$)	.024					.025	
	(.006)					(.007)	
$S1$ interacted with							
High school dropout (D_1)		.009	.009	.007	.011		.003
		(.011)	(.019)	(.014)	(.011)		(.013)
High school graduate (D_2)		.007	.009	.008	.008		-.003
		(.018)	(.023)	(.015)	(.018)		(.047)
College dropout (D_3)		.042	.042	.042	.034		.042
		(.008)	(.009)	(.008)	(.009)		(.009)
College graduate (D_4)		-.012	-.012	-.012	-.011		-.005
		(.019)	(.024)	(.019)	(.019)		(.028)
Estimation method[b]	OLS	OLS	ML	OLS	OLS	OLS	OLS

[a]Column 3 moves GED recipients from category D_1 to D_2. Column 4 omits students who earn two-year degrees from category D_3.

[b]OLS estimates correspond to model 5; in columns 1a and 5a, the slope restriction $\delta_1 = \delta_2 = \delta_3 = \delta_4$ is imposed. Maximum likelihood estimates correspond to the switching regressions model 6; see section 3 for details.

Note: Standard errors are in parentheses; in column 2, standard errors are based on the bootstrap with 300 replications. Additional parameter estimates are reported in table A1.

Table 6: OLS and Maximum Likelihood Estimates of Alternative Wage Models Using "Age at School Exit" as Measure of Time in School

Variable	Full sample				Drop 2-year degrees
	(1′)	(1)	(2)	(3)[a]	(4)[a]
Degree category					
High school dropout (D_1)	-.128	-.131	-.131	-.161	-.143
	(.023)	(.023)	(.025)	(.026)	(.023)
High school graduate (D_2)	-.037	-.075	-.075	-.058	-.081
	(.017)	(.021)	(.018)	(.019)	(.020)
College dropout (D_3)	.062	.045	.045	.048	.023
	(.018)	(.018)	(.017)	(.018)	(.019)
College graduate (D_4)	.273	.321	.321	.326	.318
	(.022)	(.032)	(.043)	(.032)	(.034)
Time in school (*S*2)	-.011				
	(.004)				
*S*2 interacted with					
High school dropout (D_1)		-.012	-.012	-.016	-.014
		(.005)	(.006)	(.006)	(.005)
High school graduate (D_2)		-.031	-.031	-.023	-.032
		(.008)	(.007)	(.006)	(.008)
College dropout (D_3)		.004	.004	.005	-.002
		(.005)	(.007)	(.005)	(.006)
College graduate (D_4)		-.026	-.026	-.026	-.027
		(.009)	(.015)	(.009)	(.009)
Estimation method[b]	OLS	OLS	ML	OLS	OLS

[a]Column 3 moves GED recipients from category D_1 to D_2. Column 4 omits students who earn two-year degrees from category D_3.

[b]OLS estimates correspond to model 5; in columns 1a and 5a, the slope restriction $\delta_1 = \delta_2 = \delta_3 = \delta_4$ is imposed. Maximum likelihood estimates correspond to the switching regressions model 6; see section 3 for details.

Note: Standard errors are in parentheses; in column 2, standard errors are based on the bootstrap with 300 replications. Additional parameter estimates are reported in table A2.

Table A1: Additional Estimates for Specifications in Table 5
("Highest grade completed" is used as measure of time in school)

Variable	(1′)	(1)	(2)	(3)	(4)	(5′)	(5)
Actual experience	.091	.090	.090	.089	.087	.092	.090
	(.009)	(.009)	(.010)	(.009)	(.009)	(.009)	(.009)
Actual experience squared	-.003	-.003	-.003	-.003	-.003	-.003	-.003
	(.001)	(.001)	(.006)	(.001)	(.001)	(.001)	(.001)
Early experience	.014	.013	.013	.012	.018	.010	.009
	(.033)	(.033)	(.035)	(.033)	(.034)	(.033)	(.033)
1 if male	.159	.160	.160	.160	.164	.160	.161
	(.012)	(.012)	(.011)	(.012)	(.012)	(.012)	(.012)
σ_u^2	.176	.176	.175	.176	.175	.175	.175
Number of observations	5,153	5,153	5,153	5,153	4,908	4,998	4,998

Note: Each specification also includes calendar year dummies. In column 2, the estimated probability of being degree k (λ_k) and corresponding standard errors is 0.108 (.006), 0.413 (.011), 0.287 (.008) and 0.192 (.006) for k=1,2,3 and 4; the estimated probabilities that the reported degree status is "true" (p_{kk}) are all 0.96 or higher.

Table A2: Additional Estimates for Specifications in Table 6
("Age at school exit" is used as measure of time in school)

Variable	(1′)	(1)	(2)	(3)	(4)
Actual experience	.105	.102	.102	.100	.102
	(.009)	(.009)	(.011)	(.009)	(.010)
Actual experience squared	-.004	-.004	-.004	-.004	-.004
	(.001)	(.001)	(.008)	(.001)	(.001)
Early experience	.035	.036	.036	.034	.044
	(.033)	(.033)	(.034)	(.033)	(.034)
1 if male	.156	.157	.157	.158	.163
	(.012)	(.012)	(.012)	(.012)	(.012)
σ_u^2	.177	.176	.175	.175	.175
Number of observations	5,153	5,153	5,153	5,153	4,908

Note: Each specification also includes calendar year dummies. In column 2, the estimated probability of being degree k (λ_k) and corresponding standard errors is 0.109 (.004), 0.422 (.007), 0.277 (.006) and 0.192 (.005) for k=1,2,3 and 4; the estimated probabilities that the reported degree status is "true" (p_{kk}) are all 0.96 or higher.