Student Abilities During the Expansion of U.S. Education, 1950-2000*

Lutz Hendricks† Todd Schoellman‡

February 4, 2009

Abstract

The U.S. experienced a dramatic expansion of education during the post-war period. We argue that the rise of schooling was accompanied by a large decline in average student ability for every level of schooling, with important implications for wage measurement. This paper quantifies the importance of changing abilities for explaining two major wage facts: the slowdown in wages since 1960, and the rise in the college wage premium since 1970.

We develop a model of school choice with ability heterogeneity and imperfect information about ability. In a cross-section, higher ability students are more likely to remain in school. In the time series, declining costs of schooling and rising returns to schooling induce more, lower-ability workers to remain in school. We calibrate our model to match facts on the post-war relationships between schooling, wages, and proxies for ability. We find that observed wages understate the growth of skill prices by 31 to 58 percentage points. We also find that the observed rise in the college skill premium can be entirely attributed to changes in the relative ability composition of college and high school graduates.

JEL: I2, J24.

Key words: Education. Ability. Skill premium.

---

*For helpful comments we thank seminar participants at the Cleveland Fed, the University of Georgia, the Clemson University Bag Lunch, and the Triangle Dynamic Macro Workshop. The usual disclaimer applies.

†University of North Carolina, Chapel Hill, CESifo, Munich, and CFS, Frankfurt; lutz@lhendricks.org

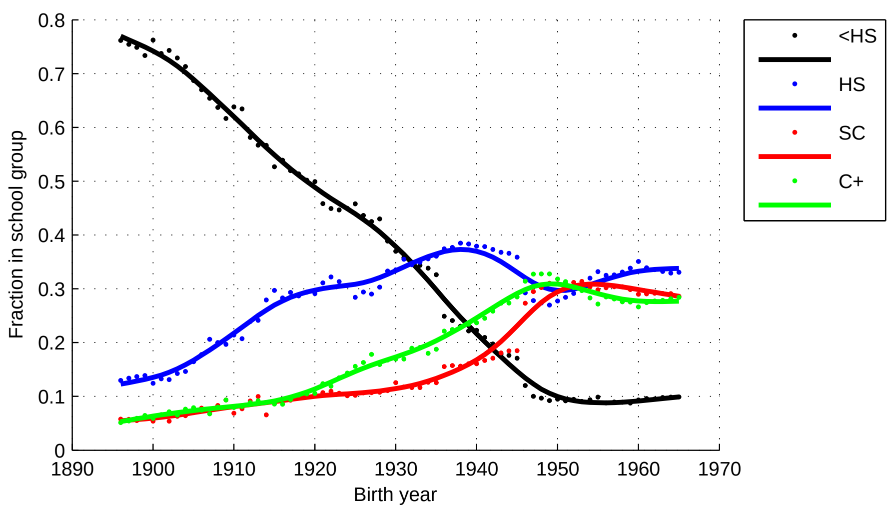
‡Clemson University; tschoel@clemson.edu

1 Introduction

The question. Over the course of the 20th century, the U.S. experienced a large increase in educational attainment. Figure 1 illustrates this trend. For each cohort born between 1896 and 1965, Figure 1 displays the fraction of men in each of four education categories: high school dropouts (<HS), high school graduates (HS), some college (SC), and college graduates (C+).[1] Of the men born in 1900, only about 25% completed high school. By 1965, high school graduation had become all but universal and the median person attended at least some college.

This dramatic increase in education raises the possibility that the average ability of students at a given level of schooling may have declined (Laitner 2000; Juhn et al. 2005). In the 1920's, the typical college student belonged to a small elite. Today, half of all students attend at least some college. Motivated by these observations, this paper attempts to answer the question: How much did the changing ability of students affect the growth rates of wages and human capital?

Figure 1: Education by Birth Cohort



Motivation. We think of measured hourly wages as the product of human capital and unobserved skill prices. If student abilities decline over time, conditional on schooling,

[1]See section A1 for details on how these statistics are constructed. Goldin & Katz (2008) summarize the expansion of U.S. education.

skill prices grow faster than measured wages (Laitner 2000). Existing measures of wage growth, which abstract from changing student abilities, are then downward biased. The main purpose of this paper is to measure the magnitude of this bias.

Similarly, existing measures of skill premia abstract from changing abilities. A large literature has documented a dramatic increase in the wages of college graduates relative to high school graduates since 1950 (Goldin & Katz 2008). If the ability composition of college graduates rises relative to high school graduates, existing measures of skill biased wage growth are upward biased. The second purpose of this paper is to measure the magnitude of this bias.

Since measured wages are the product of human capital and skill prices, any bias in the estimation of skill price growth implies an offsetting bias in the growth rate of human capital. Our third objective is to measure the magnitude of this bias.

Approach. Comparable, high quality measures of student abilities by education that span long time periods do not exist. Moreover, given that ability is measured with error, the effect of ability on wages is difficult to estimate. We therefore use a quantitative model of school choice to measure the dispersion of ability, the correlation between ability and schooling, and the effect of ability on wages.

Our model features finitely lived individuals of heterogeneous abilities. Ability determines the returns to schooling, but workers observe only a noisy signal of their own ability. They then choose between discrete schooling levels. Obtaining a higher level of schooling involves foregoing current earnings and paying some costs, but leads to higher future earnings. Our model is structured so that, with perfect information, higher ability workers go to school longer. Given that students are imperfectly informed about their own ability, educational sorting by ability is imperfect.

The key model parameters are the dispersion of ability and of the ability signal observed by students. Conventional measures of wage and human capital growth are biased if ability is highly dispersed and educational sorting is strong (the students' ability signals are precise).

In our model, two exogenous driving forces lead to rising educational attainment over time: (i) changes in the costs of schooling and (ii) changes in the skill specific wages. As education rises, the ability composition of each education level changes.

We calibrate the model to match data on U.S. educational attainment, wages, and abilities over the period 1950 to 2000. Specifically, using U.S. Census data, we estimate the educational attainment and wages earned by the cohorts born between 1916 and 1965. Using NLSY79 data, we estimate the joint distribution of schooling, measures of ability, and wages for the 1960 cohort. Finally, we estimate the dispersion of the persistent component of wages from PSID data. These form our calibration targets.

Findings. In our preferred calibration, ability heterogeneity implies a standard deviation of log wages near 0.5. This large ability dispersion, together with the dramatic expansion in U.S. education observed in the data, implies substantial declines in the mean abilities of students at all education levels. Except for college graduates, the model implies that unobserved skill prices grew at least twice as fast as measured wages. We also find substantial changes over time in the relative abilities of workers with different education. Notably, our model attributes the *entire* growth in the college wage premium since 1950 to the rising relative abilities of college graduates versus high school graduates. The role of ability changes is smaller for the relative wages of high school dropouts and college dropouts.

We find that our results remain robust when we vary the dispersion of abilities and the degree of educational sorting by ability. Our results are only overturned when ability dispersion is so small that wage variation is mostly due to luck rather than ability. However, in that case the model is at odds with the empirical relationships between wages, schooling, and abilities that we document in the paper.

Outline. The paper is structured as follows. Section 2 describes the model and derives its implications for the relationship between individual abilities and school choices. The calibration procedure and targets are presented in Section 3, followed by our findings and the sensitivity analysis in Section 4. The final section concludes.

1.1 Related Literature

A small number of previous studies have addressed the question we pose. Laitner (2000) studies a model of human capital investment with workers of heterogeneous abilities. His model qualitatively accounts for changes in relative wages and in wage inequality observed in U.S. post-war data. We move beyond a qualitative analysis and quantify the importance of changing abilities for movements in measured wages.

Finch (1946) and Taubman & Wales (1972) collect aptitude or achievement test score data from several studies to identify changes in student abilities over time. We discuss their findings further in Section 4. Juhn et al. (2005) question the comparability of these studies on the grounds that they pool data based on different aptitude tests and covering different samples. Bishop (1989) addresses the comparability problem by using the Iowa Test of Educational Development, which has been administered to 95% of Iowa schools since 1940. Unfortunately, Bishop's data contain no information about the relative scores of different education groups. A related literature documents that students with higher test scores are more likely to continue their education (Heckman & Vytlacil 2001; Cunha et al. 2005).

A fundamental problem with all of the test scores reported in the literature is that they partly measure human capital produced in school rather than innate abilities. Even IQ scores are strongly affected by schooling (Winship & Korenman 1997). Moreover, students arguably know far more about their abilities than only their test scores (see Cunha et al. 2005, who also propose methods for identifying students' information). Our model addresses these issues by explicitly modeling test scores as noisy signals of students' information about their abilities.

Juhn et al. (2005) propose an approach that avoids measuring abilities entirely. They investigate whether more educated cohorts earn lower wages in a given Census year and find a weak effect. Juhn et al.'s approach faces a number of challenges. Given that cohort education rises smoothly over time, it is difficult to disentangle the effects of experience, cohort quality and time varying skill prices. The identifying variation in their approach comes from the relative wage movements of young (educated) and old (less educated) cohorts. An alternative interpretation for such wage movements has been proposed by Card & Lemieux (2001). They show that the rising skill premium during the 1980s affected young and old workers differently and interpret this as evidence in favor of imperfect substitutability between young and old workers. We avoid this issue by focusing our analysis on workers within a 10 year age window.

Our work is also related to the large literature that documents the evolution of skill premia in the U.S. and proposes a range of explanations. We refer the reader to Goldin & Katz (2008) for references. Our analysis complements this literature. It suggests that the changing ability composition of workers masks some movements of relative wages during the post-war period.

Finally, our work has implications for why U.S. educational attainment increased dramatically since 1950. We plan to explore this question in more detail in future work. Rangazas (2002) and Restuccia & Vandenbroucke (2008) present alternative macroeconomic studies of this issue, which abstract from heterogeneity in student ability.

2 A Model of School Choice

Outline. We develop a model of school choice to measure the changing ability composition of workers with different educational attainment since 1950. A broad outline of the model is as follows. The economy is inhabited by cohorts of finitely lived individuals. At birth, each individual is endowed with an *ability to learn*. Based on a noisy signal of ability, students choose among a discrete number of schooling levels. Choosing more schooling raises expected lifetime earnings, but incurs higher schooling costs. A key assumption is that highly able students produce more human capital per year of schooling. This leads students who receive a favorable signal of their ability to choose longer schooling. However, since information about ability is noisy, workers do not always choose the ex-post optimal level of education and the correlation between ability and schooling is less than perfect.

The model accommodates two driving forces for the increase in schooling attainment over time: changes in the relative costs of different education levels, and changes in the relative wages earned by different school levels. The change in average ability will depend on the source of the underlying change in attainment.

Demographics. Time is discrete, starts at year $t = 0$, and continues forever. Each year, a cohort of new workers of unit measure is born. Workers live for T periods. They are indexed by their birth cohort τ and their age v, with period t then given by $t = \tau + v - 1$.

Timing. The timing of events over an individual's lifetime is as follows. At birth, each worker is endowed with an ability to learn or generate human capital from time spent in school (A). Ability is not directly observed. Instead, the worker draws a noisy signal of A, denoted A^*. Based on this signal, the worker chooses one of S schooling levels, indexed by s.

She then spends T_s periods in school producing human capital, after which she enters the labor market and works an exogenous number of hours in each period of life, earning an exogenous hourly wage. The remainder of her time endowment is consumed as leisure. Students sell the expected present value of lifetime earnings in an actuarially fair asset market. The resulting payment is spent on consumption over the life-cycle.

The choice of schooling is irreversible. Students cannot return to school upon learning their true ability. We view this assumption as a simplified version of an environment where workers slowly learn about the true value of their education while on the job. As long as ability is not revealed too quickly, older workers will find returning to school generally undesirable. This assumption also motivates why our analysis focuses on 35-44 year old workers.

We assume that the ability distribution is log-Normal: $a \equiv \log(A) \sim N(\mu_a, \sigma_a)$. Conditional on a, the distribution of the signal is also log-Normal: $a^* \equiv \log(A^*) \sim N(a, \sigma_{a^*})$. We refer to the special case $\sigma_{a^*} = 0$ as the perfect information case. The dispersions of ability σ_a and noise σ_{a^*} are key parameters of the model.

Schooling. Upon completing school level s, a worker with ability A is endowed with

$$h(s, A) = (A)^{\eta_s} \tag{1}$$

units of skill s human capital. $\eta_s > 0$ determines the elasticity of human capital with respect to learning ability. This human capital production function matches two key facts in the data. First, if wages are proportional to human capital then wages are log-Normally distributed in our model, which is roughly consistent with the data. Second, if η_s is increasing in s, then wage gains from schooling and schooling attainment increase with ability. This leads to a positive correlation between ability and schooling (see Section 2.2), consistent with the positive correlation between aptitude test scores and schooling we document in U.S. data (see Section 3.1.2).

Work. At ages 1 through T_s, students are in school and do not work. After graduation, the worker supplies $h(s, A)\ e_{s,v}\ l_{s,v}$ units of type s labor. $e_{s,v}$ is an exogenous age-efficiency profile. $l_{s,v}$ is an exogenous age hours profile. $l_{s,v} = 0$ while in school. $1 - l_{s,v}$ is consumed as leisure.

A worker with school type s earns a wage of $x_t z_{s,t}$ per efficiency unit of work time. Wages are in units of the consumption good. Alternatively, we may think of workers as producing $x_t z_{s,t}$ units of consumption. x_t is a skill neutral level of wages (or labor productivity). It grows exogenously at the constant rate $g(x)$.$z_{s,t}$ determines the relative wage (or productivity) of type s labor. It grows at the constant rate $g(z_s)$. Movements in relative wages (differences in $g(z_s)$ across skill levels) are often associated with skill-biased technological change (e.g., Bound et al. 1992).

Preferences. Individuals order paths of consumption ($c_{s,q,v}$) and leisure according to

$$\sum_{v=1}^{T} \beta^v \left[\log(c_{s,q,v}) + \xi \log(1 - l_{s,v})\right] - \chi_{s,\tau} \tag{2}$$

where $q = (A^*, \tau)$ denotes the worker's type, $\beta > 0$ is the discount factor, and $\xi > 0$ determines the relative weight of leisure in the period utility function. $\chi_{s,\tau}$ is the utility cost of schooling. School costs measure the relative preferences of workers for time spent in school versus work, the relative preferences of workers for college versus high school occupations, and the relative financial costs of different education levels. Since workers have access to complete consumption insurance, we need not specify preferences over uncertain consumption streams.

Discussion. A number of our modeling assumptions deserve comment. Our notion of ability encompasses any endowment that affects an individual's wages or returns to schooling. It includes cognitive skills, noncognitive skills that are valued in the labor market, preferences that affect learning effort in school, and more. We do not take a position on which of these individual endowments are important or how they are developed.

Complete consumption insurance simplifies the analysis without affecting the findings in an obvious direction. Whether borrowing constraints are important for school choice is a controversial issue in the literature. Cameron & Taber (2004) find no evidence of borrowing constraints in the U.S. However, their evidence does not apply to the early cohorts contained in our data.

The assumption that students do not perfectly observe their abilities generates imperfect educational sorting. The degree of sorting is important for the quantitative results (see Section 4.2). Other forms of heterogeneity could generate imperfect sorting, such as dispersion in school costs. We lack clear evidence to distinguish these alternatives. This issue should be explored in future work.

We assume that hours worked vary across education groups. This is necessary so that the model can simultaneously account for variation in wages and lifetime earnings.

Some authors argue that a rising skill premium may reflect an increase in the rental price of high ability labor relative to low ability labor (Juhn et al. 1993; Murane et al. 2005). In assuming that earnings depend on human capital, but not directly on ability, we abstract from this possibility.

The distribution of abilities is assumed to be time invariant. We are not aware of data that speak to changes in the dispersion of abilities. However, the Flynn effect (Flynn 1984) suggests that intelligence test scores trend up at a rate of around one standard deviation every fifty years. As discussed in Flynn (1999), it is not clear whether the trend in test scores represents a trend in intelligence or in test taking skills. We discuss the implication of trending mean ability further in Section 2.3.

2.1 Worker's Problem

Workers choose schooling s and a consumption path $c_{s,q,v}$ to maximize (2) subject a budget constraint which equates the present value of consumption to the expected value of lifetime earnings:

$$\sum_{v=1}^{T} \frac{c_{s,q,v}}{R^v} = \int_{q'} \Pr(A|A^*)\ Y\left(s,A\right)\ dA \tag{3}$$

where

$$Y\left(s,A\right) = \sum_{v=T_s+1}^{T} \frac{x_{\tau+v-1} z_{s,\tau+v-1}\ l_{s,v} e_{s,v}\ h\left(s,A\right)}{R^v} \tag{4}$$

denotes the present value of lifetime earnings. At birth, individuals have access to complete markets where they can buy and sell instruments that pay off conditional on the different realizations of true ability. We assume these instruments have actuarially fair prices in the sense that the price of a state A contingent bond equals $\Pr(A|A^*)$. R is the exogenous gross interest rate.

2.2 Optimal Consumption and Schooling

Next, we derive expressions that characterize the worker's consumption and schooling decisions. We can solve the worker's problem in two steps: first, we find the optimal allocation of consumption over time given school choice; then we find the school choice that maximizes lifetime utility.

The lifetime consumption profile obeys the standard Euler equation

$$c_{s,q,v+1} = \beta R c_{s,q,v} \tag{5}$$

which implies a present value of lifetime consumption given by $c_{s,q,1}\Lambda$ where $\Lambda = \sum_{v=1}^{T} \beta^{v-1}/R$ is a present value factor. The budget constraint then implies a level of consumption given by

$$c_{s,q,1} = \Lambda^{-1}\ E\left\{Y\left(s, A\right)|A^{*}\right\} \tag{6}$$

Lifetime utility is then given by

$$\begin{aligned} V\left(s, q\right) &= \sum_{v=1}^{T} \beta^{v}\left[\log\left(c_{s,q,1}\right) + \left(v-1\right)\log\left(\beta R\right) + \xi\log\left(1-l_{s,v}\right)\right] - \chi_{s,\tau} & (7)\\ &= R\Lambda\beta\log\left(\Lambda^{-1}\ E\left\{Y\left(s, A\right)|A^{*}\right\}\right) - \hat{\chi}_{s,\tau} & (8)\end{aligned}$$

where

$$\hat{\chi}_{s,\tau} = \chi_{s,\tau} - \sum_{v=1}^{T}\beta^{v}(v-1)\log(\beta R) + \xi\sum_{v=1}^{T}\beta^{v}\log(1-l_{s,v}) \tag{9}$$

is an aggregate of all the school-specific terms that are constant across workers. Optimal school choice satisfies

$$s = \arg\max V\left(s, q\right) \tag{10}$$

Educational sorting. We derive conditions under which the model implies positive sorting. Each worker's school choice is determined by the value gap $V\left(s+1, q\right) - V\left(s, q\right)$. If this gap is positive, the worker prefers $s+1$ over s. How does the gap change with the ability signal? Note that $Y\left(s, A\right)$ equals $h\left(s, A\right)$ times a constant that does not depend on ability. Therefore

$$\begin{aligned} \frac{\partial V\left(s, q\right)}{\partial a^{*}} &= R\Lambda\beta\frac{\partial\log\left(E\left\{Y\left(s, A\right)|A^{*}\right\}\right)}{\partial a^{*}} & (11)\\ &= R\Lambda\beta\frac{\partial\log\left(E\{A^{\eta_s}|A^{*}\}\right)}{\partial a^{*}} & (12)\end{aligned}$$

Consider first the case of *perfect information* where $A = A^{*}$. In this case

$$\frac{\partial V\left(s, q\right)}{\partial a} = R\Lambda\beta\eta_{s} \tag{13}$$

The gains to higher levels of schooling are increasing in ability if and only if $\eta_{s+1} > \eta_{s}$. We assume this property through the rest of the paper. The properties of the model are then as follows: if a student with ability A is indifferent between s and $s+1$, students with higher ability prefer $s+1$, and students with lower ability prefer s. Depending on parameters it is possible that no student is indifferent between two education categories and hence no students choose, for instance, to attend college. However, given that we study broad education categories, we do not consider such parameter configurations. Therefore, we conclude: students perfectly sort by their ability. Students with higher ability go to school longer.

Consider next the case of *imperfect information*. Maintaining the assumption that $\eta_{s+1} > \eta_s$, students segment on expected ability, which in this model is driven by their ability signal. Given a student with expected ability $E(A|A^*)$ who is indifferent between schooling levels s and $s+1$, students with higher expected ability prefer $s+1$, and students with lower expected ability prefer s. Since the ability signal is the only determinant of expected ability, and expected ability is strictly increasing in the signal, students segment perfectly by ability signal. Students who receive higher ability signals go to school longer.

2.3 The Rise in Schooling

Our model offers two reasons why schooling may rise over time: changes in the relative costs of schooling ($\chi_{s,\tau}$) and changes in relative wages ($z_{s,t}$).To gain insight into the determinants of educational attainment, consider the indifference condition

$$\begin{aligned} V(s+1,q) - V(s,q) &= \qquad (14) \\ R\Lambda\beta \left[\log\left(E\left[Y(s+1,A)|A^*\right]\right) - \log\left(E\left[Y(s,A)|A^*\right]\right)\right] + \hat{\chi}_{s,\tau} - \hat{\chi}_{s+1,\tau} &= 0 \qquad (15) \end{aligned}$$

This determines the ability level of the marginal household who is just indifferent between choosing s or $s+1$. It is useful to write lifetime earnings as

$$Y(s,A) = (A)^{\eta_s}\ x_{\tau+34}\ z_{s,\tau+34}\ e_{s,35}\ M_s \qquad (16)$$

where

$$M_s = \sum_{v=T_s+1}^{T} \frac{x_{\tau+v-1} z_{s,\tau+v-1} l_{s,v} e_{s,v}}{R^v x_{\tau+34} z_{s,\tau+34} e_{s,35}} \qquad (17)$$

Lifetime earnings have three components: human capital $(A)^{\eta_s}$, the wage earned per unit of human capital at age 35 (an arbitrary, fixed age), and the time invariant ratio of lifetime earnings to the wage per hour at age 35, M_s. The model therefore implies that schooling is time invariant if (i) schooling costs grow at the same rate for all levels, so that $\hat{\chi}_{s+1,\tau} - \hat{\chi}_{s,\tau}$ is constant over time, and (ii) relative skill prices do not change, so that $\log(z_{s+1,t}) - \log(z_{s,t})$ is constant over time. Skill neutral wage growth ($g(x) > 0$) does not affect schooling decisions.

However, if $\hat{\chi}_{s+1,\tau} - \hat{\chi}_{s,\tau}$ falls over time, then attainment level $s+1$ will become relatively less costly and students will tend to go to school longer. Similarly, if $g(z_{s+1}) > g(z_s)$, relative wages of more skilled workers rise over time and workers remain longer in school. Lacking data on the relative importance of the school costs and wage growth for changes in U.S. educational attainment, we calibrate the processes governing $\hat{\chi}_{s,\tau}$ and $z_{s,t}$.

The model can accommodate two alternative causes of rising education. First, an increase in the average ability of students (μ_a) is isomorphic to a particular parameterization of skill biased wage growth in our model. That is, increasing μ_a by $\Delta\mu_{a,t}$ and reducing $\log(z_{s,t})$ by $\eta_s^{-1}\Delta\mu_{a,t}$ leaves all wages and therefore schooling decision unchanged. Since all η_s in our calibrated model are close to each other, the implied changes in relative skill prices, $(\eta_{s+1} - \eta_s)\Delta\mu_{a,t}$ are small. Growth in μ_a therefore has approximately the same effect as skill neutral wage growth ($g(x)$). It does not affect our conclusions about the changes in relative wages and relative human capital growth rates.

Increases in education quality (η_s) are isomorphic to skill biased wage growth. In (14), an increase in $\eta_{s+1} - \eta_s$ has a similar effect to an increase in $\log(z_{s+1,t}) - \log(z_{s,t})$. Both increase the incentives to attend schooling and both raise the relative wages paid to skilled labor. The difference is that rising school quality has a stronger effect on the highly able, while rising wages affect all workers symmetrically. Therefore, the implications for the dispersion of wages within school groups differ. However, since we calibrate our model to match data on mean wages, we cannot distinguish school quality growth from relative wage growth.

3 Calibration of the Model

Model parameters. The parameters to be calibrated determine worker preferences (β), schooling technologies (η_s) and costs ($\hat{\chi}_s$), prices ($R, x_{2000}, g(x), z_{s,2000}, g(z_s)$), and the moments of the ability related distributions ($\mu_a, \sigma_a, \sigma_{a^*}$). For all level parameters that grow over time, we take the base year to be 2000. Since our data show a marked slowdown in wage growth during the 1970s, we assume that x grows at rate $g_1(x)$ until 1975 and at rate $g_2(x)$ thereafter.

Calibration approach. Some of the parameters are fixed on the basis of outside evidence. We let a model period correspond to one year, and assume that workers live for $T = 70$ years. We set the interest rate to 5% ($R = 1.05$) and fix $\beta = 1/R$.

A number of parameters may be normalized. Since $z_{s,t}$ determines only relative wages, we can normalize $z_1 = 1$ and $g(z_1) = 0$. Since abilities have no units until converted into productivities, η_s and σ_a always appear as the product $\eta_s \sigma_a$. We may therefore normalize $\eta_1 = 1$. Finally, choosing units of A allows us to normalize $\mu_a = 1$.

The remaining parameters are chosen to match the following observations:

1. From the 1950-2000 waves of the U.S. Census, we estimate the education attained and wages earned by the cohorts born between 1906 and 1965.

2. From the NLSY79, we estimate the joint distribution of schooling, wages, and a noisy measure of ability for the 1960 birth cohort (AFQT scores).

3. From the PSID, we estimate the dispersion of the permanent component of wages.

Our calibration algorithm simulates life histories for the cohorts born between 1906 and 1965. The algorithm searches over the space of the parameters to minimize the weighted sum of squared deviations between model and data moments. The school cost parameters $\hat{\chi}_{s,\tau}$ are chosen as residuals to replicate exactly the educational attainment of each cohort. This method attributes short-term variation in attainment to costs rather than to relative wage changes, which seems plausible for events such as the Vietnam War. Fitting educational attainment exactly is important since variations in ability by cohort-education status are the focus of our interest. The remainder of this section describes how the data moments are constructed.

3.1 Data Moments

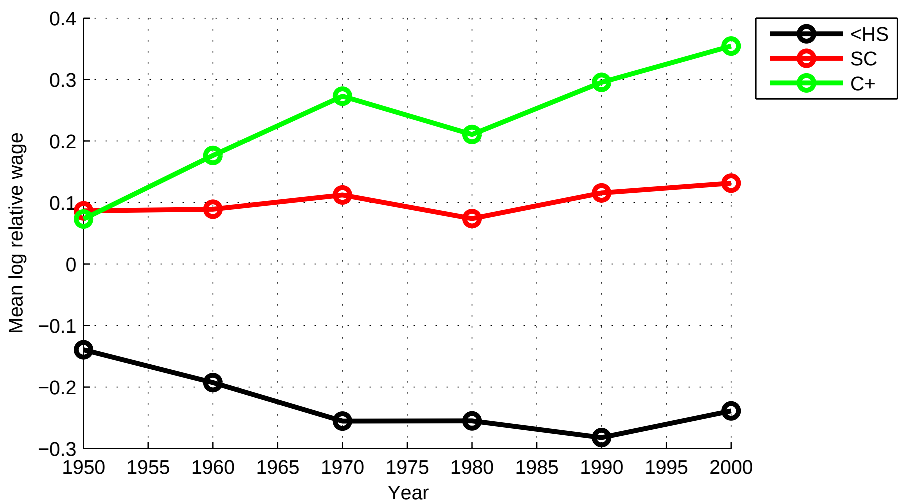
3.1.1 Cohort Education and Wages

We estimate educational attainment and wages by cohort from the 1950 to 2000 waves of the IPUMS database (Ruggles & Sobeck 1997). We do not include 1940 because it is a war year. The sample includes all men aged 15-70 who are not in school, who do not live in group quarters, and who report positive wage and salary income. We construct measures of educational attainment and of real hourly wages for each cohort born between 1906 and 1965. Each cohort is observed exactly once between the ages of 35 and 44. See Appendix A1 for details.

Cohort educational attainment. Figure 1 in the Introduction shows the fraction of persons in each birth cohort that reports a given schooling level. Similar data have been reported, for example, by Goldin & Katz (2008). The solid lines represent Hodrick-Prescott filtered data. To highlight the long-run trends, we include 1940 data in this figure, even though we do not use them in the calibration.

Relative wages. For each Census year, Figure 2 shows the mean log wage of each school group relative to high school graduates. Our data replicate the main features previously documented by Goldin & Katz (2008). Since 1950, we observe a sharp increase in the college wage premium and a decline in the relative wages of high school dropouts.

Figure 2: Skill Premia by Census Year



3.1.2 Education and Aptitudes

We use NLSY79 data to measure the degree of educational sorting by ability and the covariation of wages with ability. The NLSY79 is a representative, ongoing sample of persons born between 1957 and 1964. We retain all men who participated in the ASWAB battery of aptitude tests, which we interpret as a noisy signal of ability. We include members of the minority samples, but use weights in our analysis to offset the oversampling of minorities. For each person, we construct measures of real hourly wages at age 35 and of educational attainment. The details are given in Appendix A2.

Table 1: Schooling and AFQT: NLSY79 data

AFQT quintile	<HS	HS	SC	C+
1	0.499	0.200	0.091	0.014
2	0.278	0.273	0.167	0.036
3	0.140	0.255	0.240	0.133
4	0.067	0.192	0.287	0.266
5	0.015	0.080	0.215	0.550
Fraction	0.208	0.385	0.173	0.234
N	1404	2034	888	881

Note: Fraction of persons falling in each AFQT quintile, conditional on schooling. "Fraction" denotes the fraction of persons completing each school level. N is the number of observations.

Aptitudes. Our proxy for ability is the 1980 Armed Forces Qualification Test (AFQT) percentile rank (variable R1682). The AFQT aggregates a battery of aptitude test scores into a scalar measure. The tests cover numerical operations, word knowledge, paragraph comprehension, and arithmetic reasoning (see NLS User Services 1992 for details). We remove age effects by regressing AFQT scores on the age at which the test was administered (in 1980). We transform the residual so that it has a standard Normal distribution, which conforms with our model.

We interpret AFQT scores as a noisy signal of the worker's ability signal A^*. In particular, we assume that the distribution of AFQT obeys $AFQT \sim N(a^*, \sigma_{AFQT})$. Thus, workers are assumed to have better ability information than the econometrician. σ_{AFQT} is a calibrated parameter.

Schooling and ability. Table 1 characterizes educational sorting by ability. For each school class, the table shows the fraction of persons falling into each ability quintile. Given that our model implies perfect educational sorting by ability signal A^*, the data of Table 1 mainly contain information about the precision of the AFQT signal (σ_{AFQT}).

The table shows evidence of strong sorting. Half of high school dropouts fall into the lowest AFQT quintile, whereas half of college graduates fall into the highest quintile. This is consistent with Heckman & Vytlacil (2001). Using other measures of ability, Taubman & Wales (1972) and Herrnstein & Murray (1994) suggest that sorting may have been weaker for earlier birth cohorts. We explore this possibility in Section 4.3.

Table 2: Wage regressions: NLSY79 data

	<HS	HS	SC	C+
β	0.055	0.090	0.055	0.103
σ_β	0.026	0.016	0.028	0.040
R^2	0.01	0.03	0.02	0.03
N	594	1094	390	353

Note: The table shows the results from regressing log wages at age 35 on AFQT score separately for each schooling group. β is the estimated return to schooling, σ_β is its standard error. N is the number of observations.

Wages and ability. Table 2 reports the results from regressing log wages at age 35 on AFQT within school classes. AFQT is transformed so that it has a standard Normal distribution in the population. This makes the results comparable with the literature and conforms with our model. A one standard deviation increase in AFQT is associated with a 6% to 10% increase in wages. This is consistent with other estimates of wages on AFQT that control for schooling. Bowles et al. (2002) survey 24 studies with a mean regression coefficient of 0.07.

Note that the regression coefficient is not given a structural interpretation in our analysis. We only use it to describe the data. We are interested in how the conditional mean of wages varies with measured ability, not in the "direct" effect of ability on wages, holding other characteristics constant. For this reason, we do not include controls in the wage regression. When we calibrate the model, we simulate AFQT scores and run a regression of exactly this form for the pool of workers who attain each education level.

3.1.3 Permanent wage dispersion

The last data moment used in the calibration characterizes the dispersion of the permanent component of wages. In our model, the variances of wages and lifetime earnings, conditional on schooling, are both proportional to $\eta_s^2 \sigma_a^2$. Since our model abstracts from luck and other transitory shocks to earnings, it is important that we purge transitory variation from the wage data. Following Guvenen (2007) and others, we do so by estimating the variance of permanent component of wages.

We think of log-wages of individual j at time t as being generated by an autoregressive

Table 3: Wage regressions: PSID data

Schooling	$\sigma_{Y(s)}$	σ_α	ρ	$\sigma_{\hat{\varepsilon}}$	σ_ε
<HS	0.389	0.335	0.887	0.171	0.296
HS	0.390	0.270	0.973	0.110	0.327
SC	0.359	0.285	0.881	0.192	0.268
C+	0.444	0.242	0.969	0.154	0.335

Note: The table shows the estimated coefficients obtained from wage regressions using PSID data. $\sigma_{Y(s)}$ is the standard deviation of lifetime earnings.

earnings process with an individual-specific fixed component:

$$\log(w_{j,t}) = \alpha_j + X_{j,t}\beta + \zeta_{j,t} + \varepsilon_{j,t} \tag{18}$$

where X is the vector of the individual's characteristics, β is a vector of constants, ε is a transitory shock, and ζ is a persistent shock which evolves according to an AR(1) process. The moment of interest is the variance of α_j. We estimate this income process using PSID data. Results are given in Table 3. Details are available in Appendix A3.

In using this approach, we assume that variation in α_j is predictable at the time schooling decisions are made. The findings of Geweke & Keane (2000), Cunha, Heckman & Navarro (2005) suggest that a large share of variation in lifetime earnings is predictable.

3.2 Model Parameters

The calibrated parameters are given in Table 4. The key parameters are the standard deviations of the various ability measures and the η_s which transform abilities into productivities. We find that all η_s are close to unity, which simplifies the interpretation of the findings. Approximately, abilities correspond to productivities in all school groups.

The standard deviation of abilities (σ_a) is near 0.5. Since this parameter is central for our findings, it is useful to consider whether its magnitude appears reasonable. The model implies that, for given schooling, a person at the 95^{th} ability percentile earns 2.5 times more than a person of median ability. The range between the 95^{th} and the 5^{th} percentile amounts to a 6.3 fold wage gap. The calibrated value of σ_a is also close to the standard deviation of log adjusted wages within age / school groups. It varies across cohorts, but without a clear trend. According to the estimates of Section 3.1.3, about 60% of this wage dispersion

Table 4: Model Parameters

Moment	Value	Moment	Value
x_{2000}	7.906	$g_{x,<1975}$	2.67%
$z_{2000,2}$	0.776	$g_{x,>=1975}$	-0.13%
$z_{2000,3}$	0.556	g_{z_2}	0.43%
$z_{2000,4}$	0.415	g_{z_3}	0.15%
η_1	1.0000	g_{z_4}	0.33%
η_2	1.0001	σ_a	0.558
η_3	1.0004	σ_{a^*}	0.220
η_4	1.0007	σ_{AFQT}	0.848

is permanent. As our sensitivity analysis reveals, this calibration target, which corresponds to $\eta_s^2 Var(a|s)$ in the model, is largely responsible for the calibrated value of σ_a.

The combination of large ability dispersion and sizeable changes in educational attainment over time raises the concern that our model may imply large trends in the variance of wages over time. This is not the case, mainly because educational sorting by ability is highly imperfect. The largest movement is for high school dropouts; wage variation declines by 22% for this group while rising modestly for the others.

It is tempting to argue that empirical estimates imply a much weaker relationship between ability and wages than our model does. In the estimates reviewed by Bowles et al. (2002), a one standard deviation increase in ability is, on average, associated with a wage gain of only 7%. However, these estimates rely on noisy measures of ability and suffer from attenuation bias.[2] By construction, our model is consistent with the empirical relationship between wages and measured ability proxies because this is one of our calibration targets.

A set of parameters that is closely related to our findings is the set of wage growth rates, $g(x) + g(z_s)$. In the model, skill prices grow between 60% and 80% between 1950 and 2000. The wage growth rates we estimate from the Census data are substantially smaller. Our model attributes the differences to changes in worker abilities over time. This foreshadows the finding, discussed in Section 4, that measured wage growth substantially underestimates the growth of skill prices.

[2] Bishop (1989) corrects for measurement error and estimates that a one standard error increase in ability leads to a 19% increase in wages. He assumes that the variance of measurement error is known and given by the KR-20 reliability of the PSID's GIA score.

3.3 Model Fit

In this section, we evaluate the model's ability to replicate the calibration targets. Since the model exactly replicates educational attainment by cohort, this is not shown.

Wages. The first set of calibration targets consists of mean log wages of persons aged 35-44 in each Census year. Figure 3 compares the model predictions with the data described in Section 3.1.1. Each line represents a school group. The model accounts well for the low frequency wage movements seen in the data. A sharp decline in wage growth is visible after 1970, which motivates our assumption that skill neutral wage growth slows in 1975.

Figure 3: Mean Log Wages, Model and Data

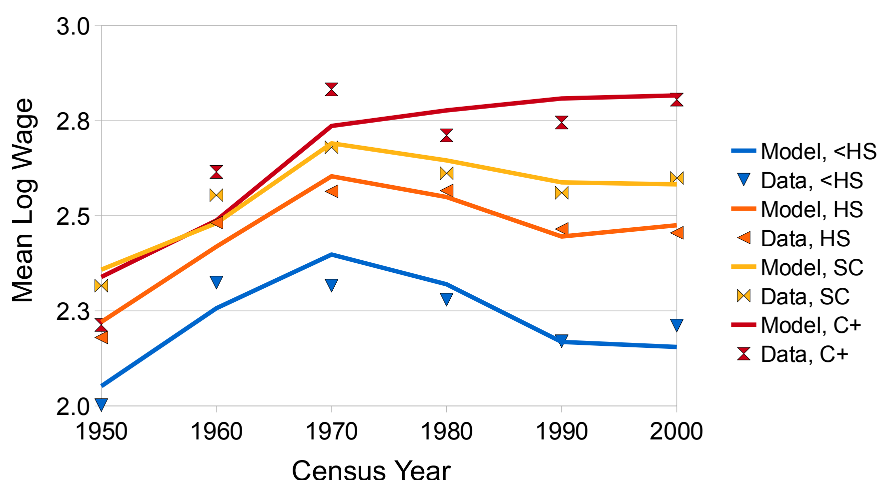
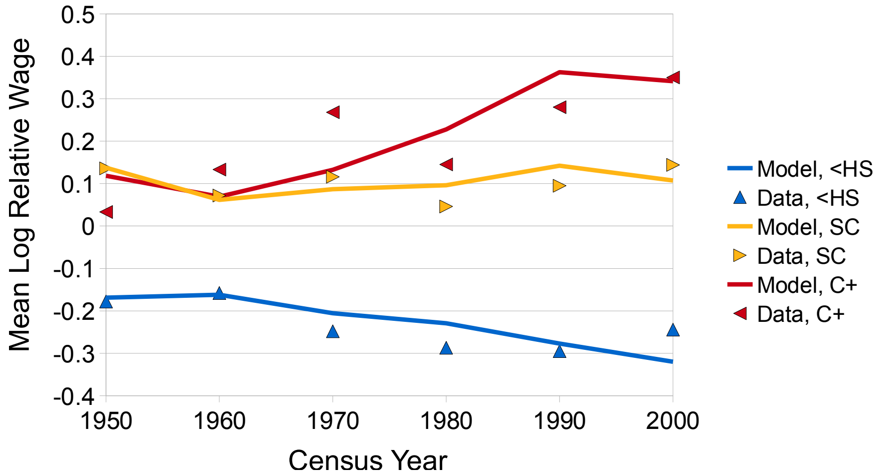


Figure 4 displays the same wage data in the form of skill premia relative to high school graduates. The model accounts well for the long-run trends in relative wages. Two main discrepancies are visible. The model misses the large drop in the college premium between 1970 and 1980. This drop could be due to the changes in the coding of schooling discussed in Appendix A1. The model overstates relative college wages in 1950. In the data, the mean college wage in 1950 is below the mean wage earned by a college dropout. This anomalous result may be due to the small size of the 1950 sample (see Appendix A1).

Education and aptitudes. The second set of calibration targets characterizes the joint distribution of AFQT, schooling, and wages discussed in Section 3.1.2. Figure 5 shows the density of AFQT scores by schooling level for the model and the NLSY79 data. Overall,

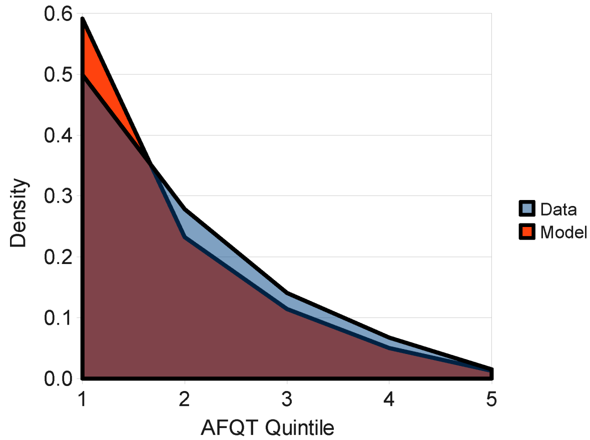
Figure 4: Returns to Schooling, Model and Data



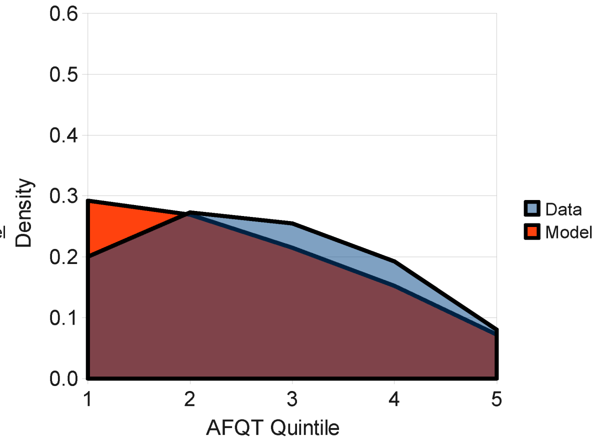
the model accounts reasonably well for the data. The main discrepancy is the too large fraction of low AFQT persons among high school graduates.

Figure 6 displays the results of regressing log wages at age 35 on AFQT, which is scaled to have a standard Normal distribution. Separate regressions are estimated for each school group. In the model, the returns to AFQT are generally lower than in the data (described in Section 3.1.2), but especially so in the middle school groups. To understand this, note that the returns in the model are $\eta_s \frac{da}{dAFQT}$. η_s is similar for all s. Low returns in the middle groups therefore mean more noise in AFQT as a measure of ability. One reason for this is that ability noise is less likely to change schooling in the lowest and highest education classes. For college graduates, even large positive noise does not change schooling. Similarly, for high school dropouts, the same is true for large negative noise. Hence, the relationship between AFQT and ability is stronger in the lowest and highest education groups.

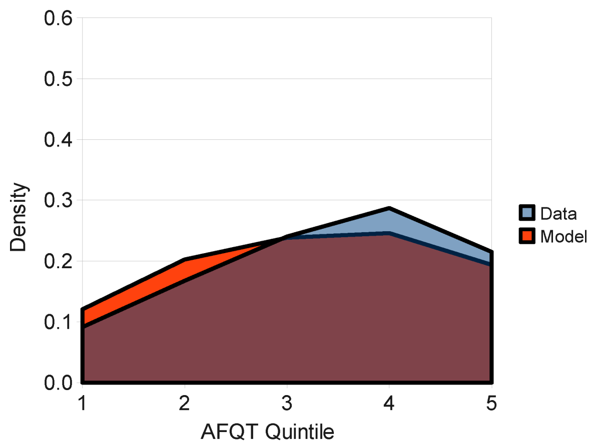
Permanent wage dispersion. The final calibration target is the dispersion of the permanent component of wages estimated in Section 3.1.3. Figure 7 compares the model standard deviations with the data for each school group. The model generates roughly the right amount of dispersion on average, but overstates the dispersion for college graduates. The model implies larger dispersion for the outer school groups because they contain most of the extreme ability draws.



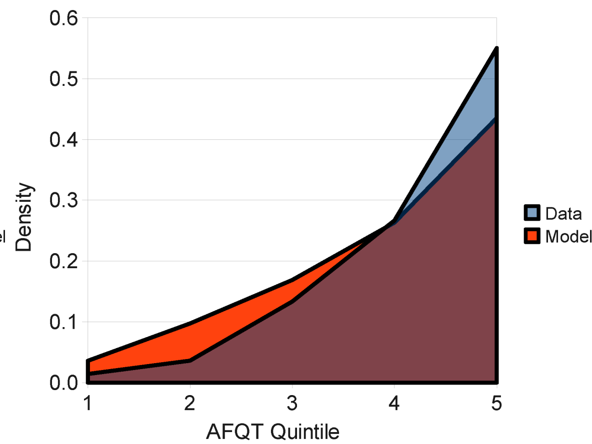
(a) High School Dropouts



(b) High School Graduates



(c) Some College



(d) College Graduates

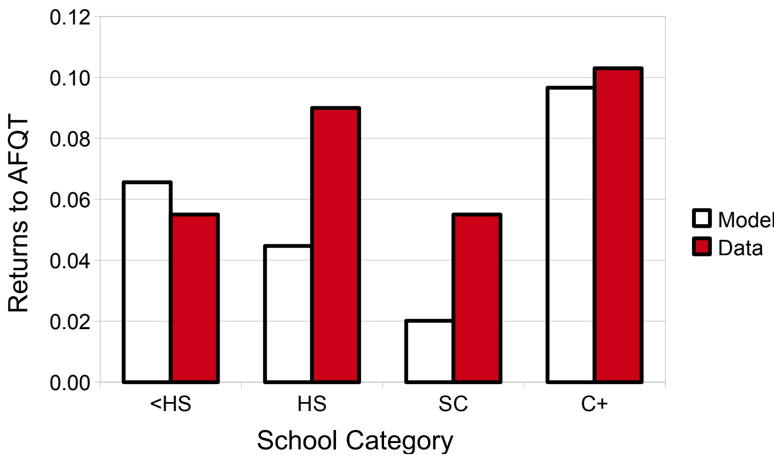
Figure 5: AFQT Distribution for Different School Attainments

4 Results

The main question we address in this paper is: As education expands, the abilities of workers with any given education level decline. How much of the observed movements of wages and skill premia in the U.S. since 1950 are due to these changing abilities?

To answer this question, Figure 8 shows mean the log abilities implied by the model for each birth cohort in our dataset. Since all η_s are close to one, ability units are close to log wage units. The model implies large declines in abilities for all school groups, but especially for the high school and some college groups. For some school groups, mean ability changes by more than total measured wages. Notably, mean log ability of high school graduates drops by 0.5, while wages grow by 0.275. The model therefore implies that large shares of

Figure 6: Returns to AFQT



wage and skill premium movements are due to abilities rather than skill prices.

Figure 9 illustrates the role of ability changes for skill premia. For each school group, the figure shows two paths of relative wages. The solid lines are the paths of measured wages predicted by the calibrated model. These were already shown in Figure 4. The dashed lines show the evolution of relative skill prices, $z_{s,t} - z_{2,t}$. All wages are in logarithms and expressed relative to high school graduates.

The main message of Figure 9 is that accounting for changes in worker ability leads to large revisions in the estimated changes of skill premia. Conventional measures of wages attribute the entire change in hourly earnings (the solid line) to skill price movements. In our model, the average ability of workers in all school groups declines as education expands over time. However, relative to high school graduates, mean abilities rise for all other school groups. As a result, the growth rates of skill premia are below the growth rates of measured relative wages.

The discrepancy is particularly striking for the college wage premium, which is the relative wage that has experienced the largest changes since 1950 and has received the most attention in the literature. Even though the relative wages of college graduates rose by 32% since 1950, our model implies that the relative price of college educated labor *declined* by 5%. The gap between measured wages and skill prices is due to a large increase in the average ability of college graduates relative to high school graduates.

Table 5 summarizes the changes in relative wages over the period 1950-2000. Measured relative wages declined for high school and college dropouts, but rose for college graduates

Figure 7: Dispersion of Permanent Earnings

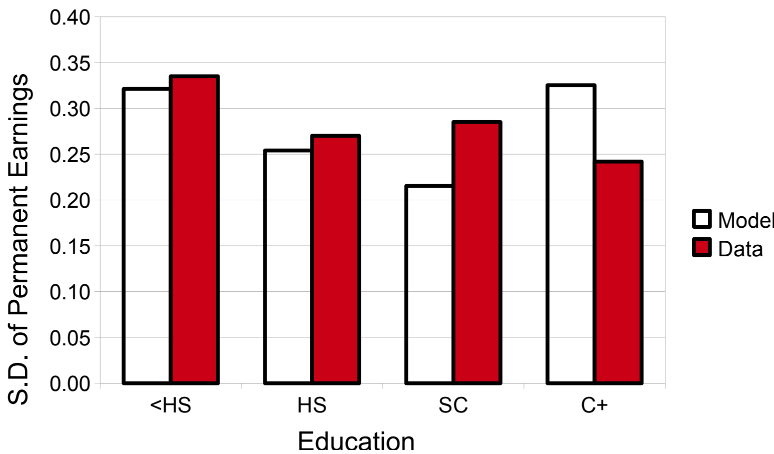


Table 5: Decomposing Changing Wage Premia

	Data Wages	Model Wages	Skill Price	Ability
<HS-HS	-0.07	-0.15	-0.21	0.06
SC-HS	0.01	-0.03	-0.14	0.11
C+-HS	0.32	0.22	-0.05	0.27

Note: The first two columns compare the Census data and model-predicted changes in wage premia from 1950-2000, against high school graduates. The last two columns decompose the model-predicted changes into the underlying sources: changing relative skill prices and changing relative ability.

(column "data wages"). The "model wages" column shows the corresponding changes in relative measured wages implied by the model. The last two columns decompose the model's predicted wage changes into two components: that due to changing relative skill prices, and that due to changing relative ability. In each case, skill prices rise more slowly than measured wages, but the discrepancy is particularly large for college graduates.

Table 6 decomposes the changes in wage *levels* into the contributions of skill prices and ability changes. Its layout is analogous to that of Table 5. The main message is that large declines in abilities mask substantial growth in skill prices. To illustrate, the measured mean wage of a high school graduate rose by 27% between 1950 and 2000. Our model implies that the rental price of high school labor grew roughly three times as fast (by 84%). However, the expansion of schooling led to a drop in the average ability (column "difference"), which erodes much of the growth in wages.

Figure 8: Mean Log Ability of Selected Cohorts

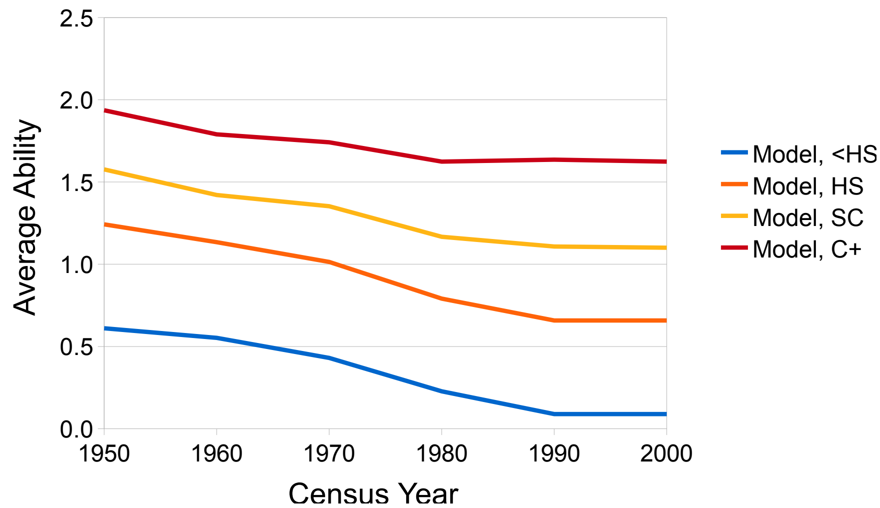


Table 6: Decomposing Changing Wage Levels

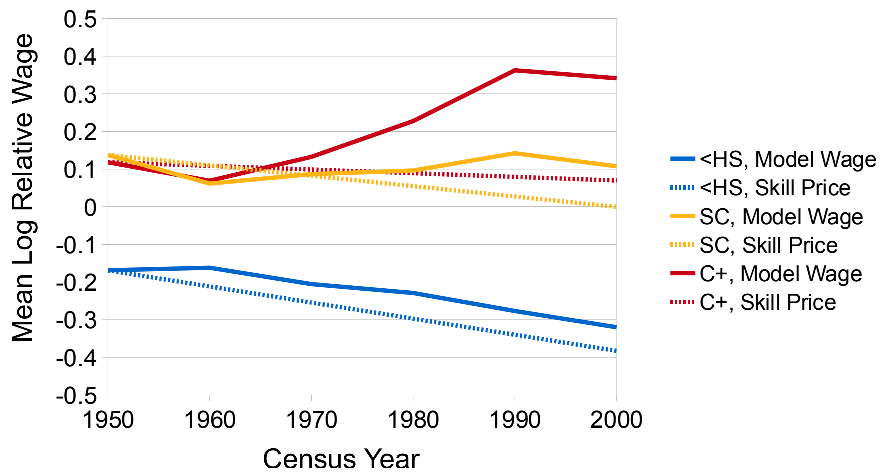
	Data Wages	Model Wages	Skill Price	Ability
<HS	0.21	0.1	0.62	-0.52
HS	0.27	0.25	0.84	-0.58
SC	0.28	0.22	0.7	-0.48
C+	0.59	0.48	0.79	-0.31

Note: The first two columns compare the Census data and model-predicted changes in average wages from 1950-2000. The last two columns decompose the model-predicted changes into the underlying sources: changing skill prices and changing ability.

A large literature has pointed out that real wages have barely increased since about 1960 (Katz & Autor 1999). Our findings support Laitner's (2000) suggestion that skill prices may have grown substantially, but measured wages are pushed down by the declining abilities of the students attending each school level.

It would be helpful to compare our findings with direct evidence on the changes in student abilities over time. Unfortunately, such evidence is scarce. Taubman & Wales (1972) collect the results of several studies. Test scores are expressed as percentile ranks and therefore contain no information about the trend in mean ability. Taubman & Wales find that the ability gap between college and high school students widened between 1925 and 1960. Herrnstein & Murray (1994) report a similar finding. This is consistent with our model's implications, shown in Figure 9.

Figure 9: Wage Premia and Relative Skill Prices, Model



The studies collected by Finch (1946) indicate that the average ability of high school graduates remained roughly constant between 1916 and 1942. By contrast, our model implies that average student ability declines over time.

A major caveat applies to all of the available time series evidence on student test scores. The nature of the tests and the student populations covered vary over time. Assumptions and adjustments are needed to compare statistics from different time periods (see Juhn et al. 2005). Moreover, since tests are generally taken when students have completed at least some high school, it is not clear to what extent test scores reflect student abilities as opposed to skills learned in school (Winship & Korenman 1997).

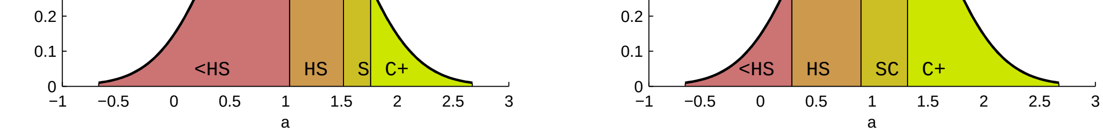
4.1 Intuition

In this section, we provide intuition for our findings. The basic mechanism that drives the results is that rising schooling inevitably lowers the average abilities of students at all levels. How strong this mechanism is depends on the dispersion of abilities and on the degree of educational sorting by ability.

An example with perfect sorting. The intuition for our findings is complicated by the interaction between ability heterogeneity and imperfect educational sorting by ability. Fortunately, a version of the model with perfect information about ability ($\sigma_{a^*} = 0$) has similar implications and is easier to understand. In this section, we present the implications of this model, assuming that all calibrated parameters are the same as in the baseline case.

Figure 10a shows the schooling decisions and abilities of the 1915 birth cohort, observed in the 1950 Census. The bell shaped curve is the Normal density of a with the calibrated standard deviation σ_a. The vertical lines represent the ability cutoffs that delineate the schooling levels. These are taken from the data, since the model exactly replicates them. For example, the least able 53% of the population choose not to complete high school. Their mean a equals 0.58, against a population mean of 1. This can be computed directly as the mean of a right truncated Normal distribution with standard deviation σ_a.

Now forward to the year 2000. Figure 10b shows the school choices for the 1965 birth cohort. Only 10% of the population fail to graduate from high school. Losing the right tail of the ability distribution reduces the mean a for high school dropouts to 0.02. The decline in mean ability is 0.56, compared with 0.51 in the calibrated model. Other education groups yield similar results. The baseline model yields ability changes that are similar to the perfect sorting example. Note that the resulting wage changes are similar to the ability changes because all η_s are close to 1.



(a) 1915 Cohort

(b) 1965 Cohort

Figure 10: Abilities and Schooling

The perfect sorting example shows why the relative ability of college graduates rises over time. The expansion of education means that new, lower ability students are added to both the high school and college populations over time. However, high school abilities are also impacted by the tendency for the most able high school students to join the some college group, while the most able college students have no further education to aspire to. Moreover, the some college group here acts as a wedge between the high school and college groups, and it has expanded over time, separating their abilities. This reasoning suggest that the increase in relative college abilities is a robust feature of our model; it is the quantitative magnitude that may vary.

Educational sorting. Given the calibrated parameters, the model implies strong sorting. Figures 11 shows distributions of ability by school level for the 1960 cohort. 90% of college graduates are drawn from the top 2 ability quintiles. More than 80% of high school dropouts are drawn from the lowest ability quintile. This happens even though the ability signal observed by the agent is quite noisy. Recall that roughly 30% of the signal's standard deviation is noise.

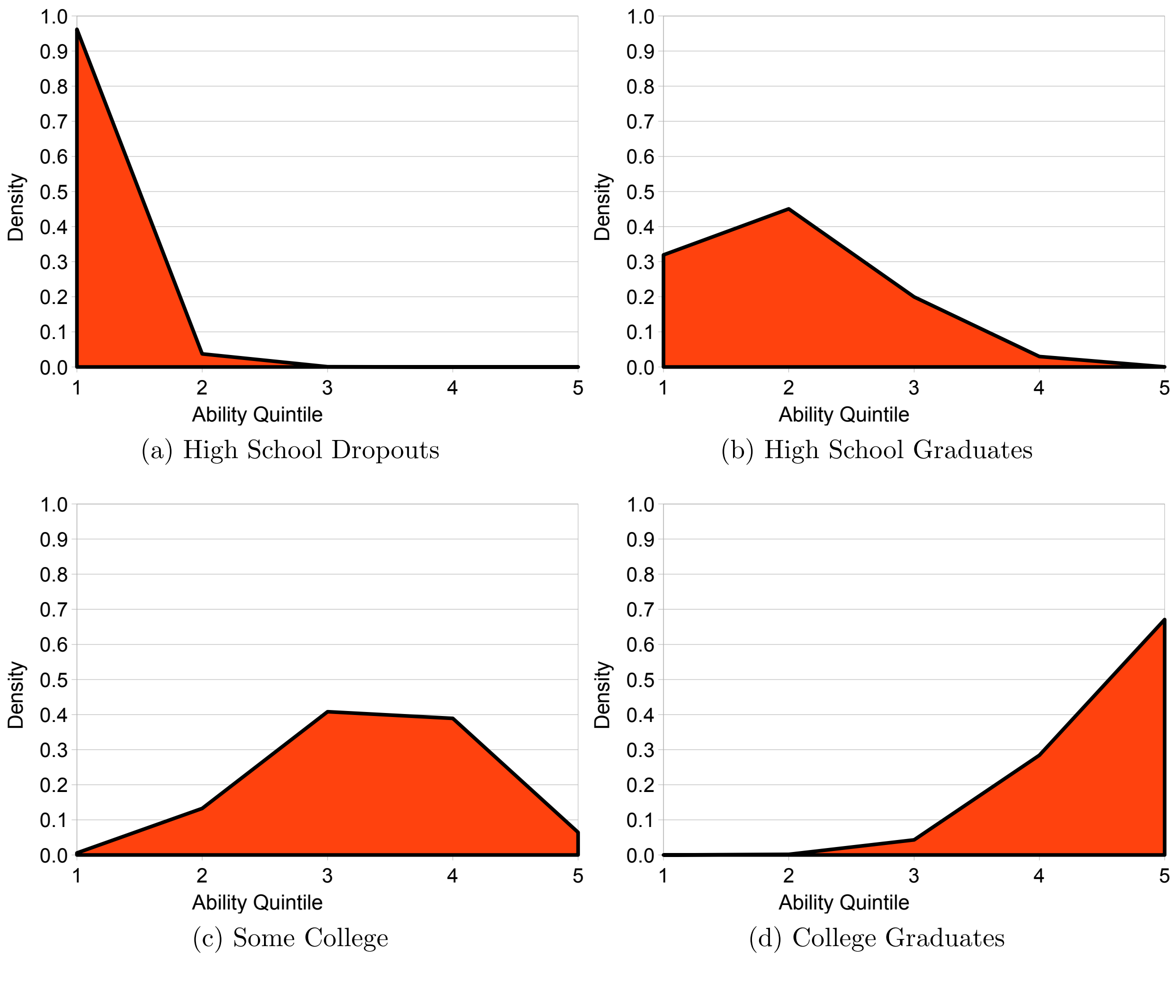


Figure 11: Ability Distribution for Different School Attainments

Table 7: Ex Ante and Ex Post Optimal Schooling

	Actual School Choice			
Ex-Post Optimal	<HS	HS	SC	C+
<HS	81.9%	11.2%	0.0%	0.0%
HS	18.1%	74.4%	23.3%	0.6%
SC	0.0%	14.0%	59.7%	17.4%
C+	0.0%	0.3%	17.0%	82.0%

Note: Each row shows person in one school group. Each column shows the fraction of persons who would have chosen other school levels, given perfect knowledge of their abilities.

Another way of assessing the degree of sorting examines how many persons ex post regret their schooling choices. Table 7 shows the fraction of persons in each school class who would revise their choices upon learning their true abilities. More than 70% of high school dropouts and of college graduates would not revise their decisions. However, roughly half of those completing high school or some college would.

It is useful to compare our findings with those of Navarro (2008). Based on a structural model of school choice, Navarro estimates that 13% of high school graduates and 16% of college graduates would revise their schooling choices upon learning about factors that affect their earnings. Conversely, 81% of the variance of lifetime earnings is predictable at age 18 for college graduates and 44% for high school graduates. The comparison suggests that our calibration may imply too much noise and hence too weak sorting by ability. To the extent this is the case, our findings are conservative.

4.2 Robustness

This section examines the robustness of our findings. The main parameters of concern are the dispersion of abilities (σ_a) and the amount of noise in the worker's ability signal (σ_{a^*}) which governs educational sorting. We study how these parameters affect the main findings as well as the model's ability to attain the calibration targets. We vary either σ_a or σ_{a^*} across a grid. For each value, we calibrate the wage parameters x_t and $z_{s,t}$ to minimize the deviation between model and data wages. All other parameters remain fixed at their baseline values.

Varying noise in the ability signal. Figure 12 shows the effect of varying σ_{a^*} from a case that is close to perfect information to a case with 75% more noise. Of particular interest are larger values of σ_{a^*} because we are concerned that our model could overstate the role ability changes, if educational sorting were too strong. The top panels of Figure 12 show how varying σ_{a^*} affects the model's ability to attain the calibration targets. This is summarized by showing the average "wage return" of AFQT (the coefficient of regressing log wages on AFQT) and the average variance of permanent wages. Both averages are taken across school groups. The lower two panels illustrate how the main results change. They show the growth rate of the high school skill price ($\log x_t + \log z_{2,t}$) and of the college wage premium ($\log z_{4,t} - \log z_{2,t}$) between 1950 and 2000.

The main insight is that varying σ_{a^*} has little effect on the findings. Ability noise raises the dispersion of permanent wages. This is due to more mixing of high and low ability workers within school groups. However, the wage and skill premium changes are not strongly affected.

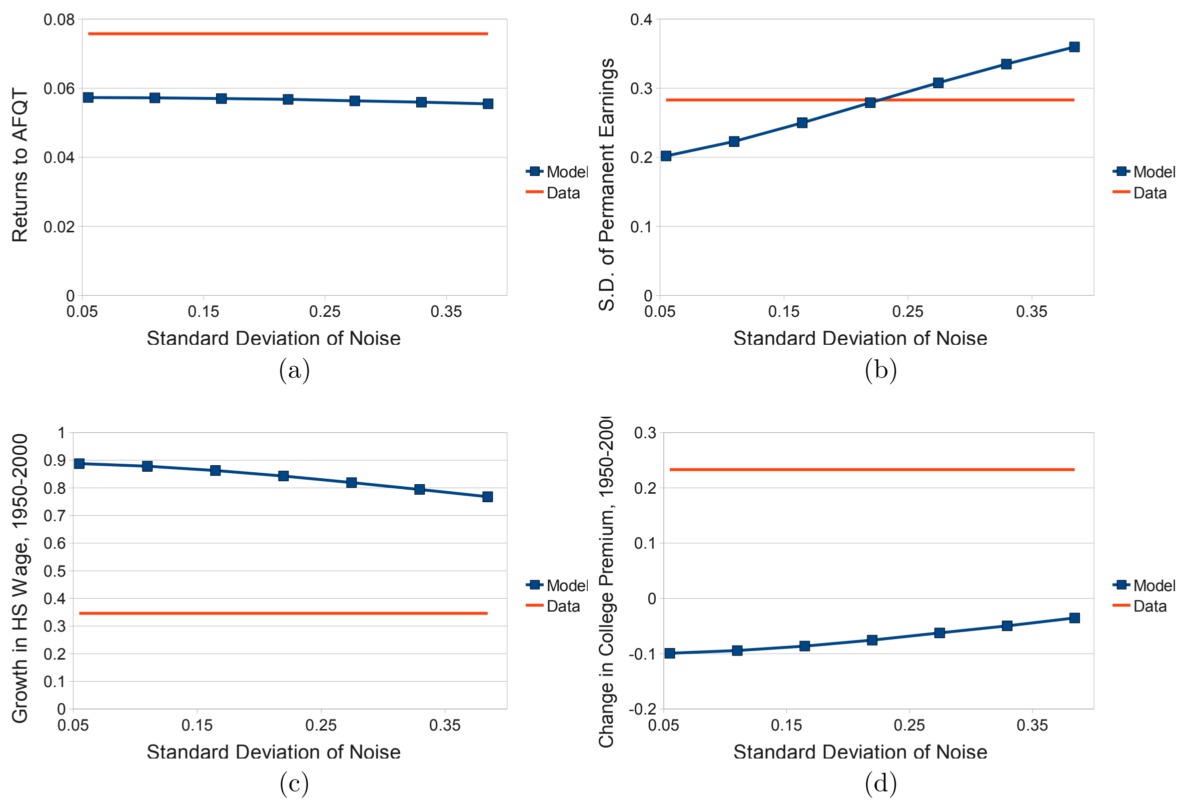


Figure 12: Changing Variance of Noise

Varying the dispersion of abilities. Figure 13 shows model economies with ability dispersion (σ_a) ranging from 25% to 100% of the baseline calibration. Higher ability dispersion would strengthen our results, so we focus on lower values. Similar to Figure 12, the top two panels show how well the model attains the calibration targets. The data clearly favor larger values of σ_a. Even the calibrated value is too low to fully account for the wage returns to AFQT and for the variance of permanent wages. Smaller values of σ_a increase the gap between model and data dramatically.

The lower two panels of Figure 13 show that lower ability dispersion strongly increases skill price growth rates. Reducing σ_a by half relative to the baseline value substantially reduces the effect of ability changes on high school wage growth (panel c) and cuts the effect on the college wage premium by two-thirds (panel d). Clearly, our results are sensitive to the calibrated value of σ_a. However, reducing σ_a by half also dramatically reduces the contribution of abilities to wages and to wage dispersion. We conclude that even a model with ability dispersion that seems much too low compared to our data targets still generates a substantial downward bias in measured high school wages and a substantial upward bias in the relative returns to college.

Jointly varying σ_{a^*} and σ_a. We also experimented with simultaneous variation in σ_{a^*} and σ_a. The results are essentially the sum of varying each parameter separately. We do not report the details in order to conserve space. The case that comes closest to overturning our results combines large noise and small ability dispersion. Some parameter combinations come close to accounting for the estimated amount of permanent wage dispersion while at the same time dramatically reducing the contribution of abilities to measured wage growth. However, these parameters imply that wage variation is mostly due to luck rather than ability. The wage returns to AFQT are very close to zero and lifetime earnings are largely unpredictable by the agent. These implications are at variance with a large literature that emphasizes the role of cognitive skills for wages (Hanushek & Woessman 2008) and with a smaller literature which shows that lifetime earnings are highly predictable (Cunha et al. 2005; Navarro 2008).

4.3 Increased Educational Sorting

A number of studies suggest that educational sorting by ability has increased between the 1920s and the 1960s. Taubman & Wales (1972) compile data from several previous studies

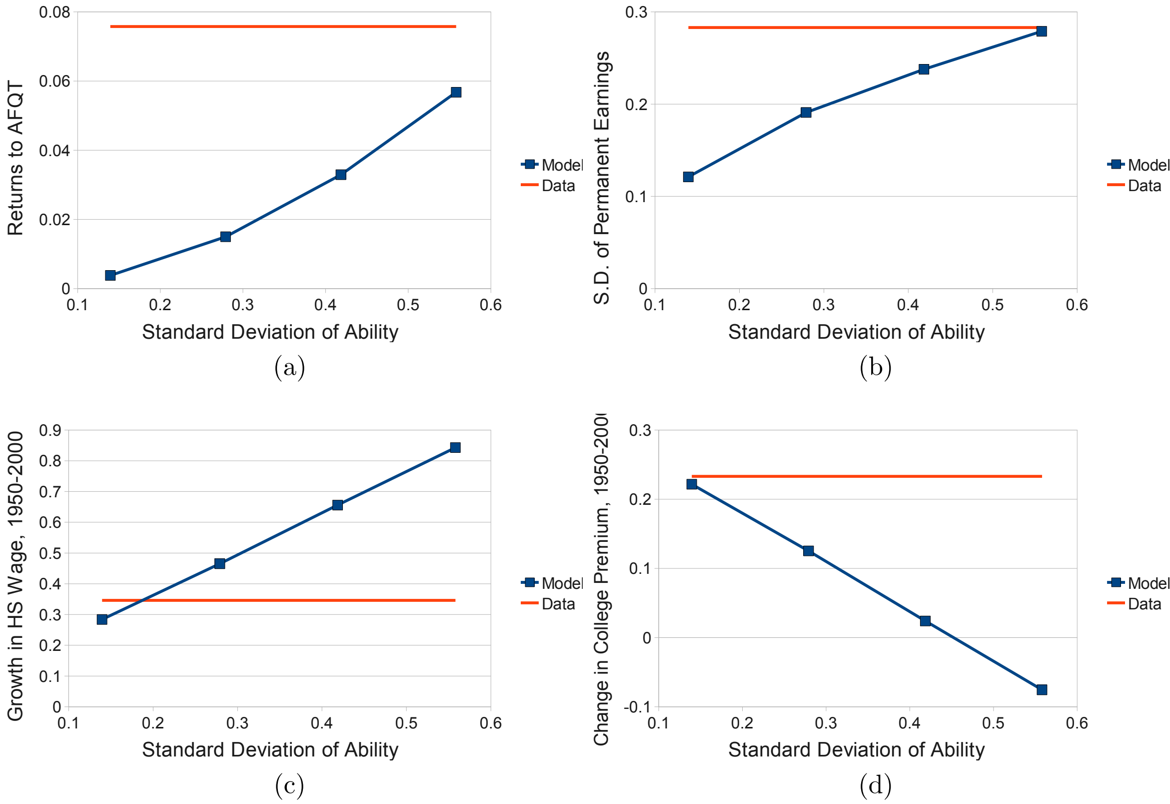


Figure 13: Changing Variance of Ability

that characterize the cognitive abilities of students at different education levels between 1925 and 1963. Their data suggest that the probability of attending college increased disproportionately among the most able students. Similar data are presented by Herrnstein & Murray (1994, chapter 1).

While the comparability of the test scores used and of the student populations covered is an issue, there are reasons to think that educational sorting may have increased. Among the contributing factors may be the declining cost of long distance travel, the relaxation of parental borrowing constraints, and the spreading of standardized testing (Herrnstein & Murray 1994, chapter 1).

We explore the implications of increased sorting by ability in our model. We modify the baseline calibration of Section 3 by assuming that the standard deviation of ability noise (σ_{a^*}) declines by half between 1906 and 1965. Otherwise, all parameters are calibrated using the same targets as in the baseline case.

The main effect of lower quality signals in earlier cohorts is more (ex-post) mistakes and

Table 8: Decomposing Changing Wage Levels With Increased Sorting

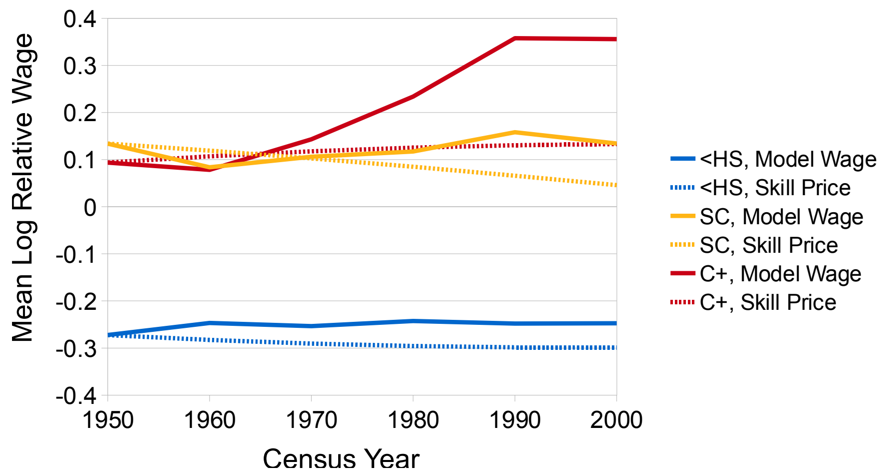
	Data	Model	Skill Price	Ability
<HS	0.21	0.23	0.66	-0.43
HS	0.27	0.21	0.69	-0.48
SC	0.28	0.21	0.6	-0.39
C+	0.59	0.47	0.72	-0.26

Note: The first two columns compare the Census data and model-predicted changes in average wages from 1950-2000. The last two columns decompose the model-predicted changes into the underlying sources: changing skill prices and changing ability.

less sorting by ability compared to Figure 11. Increased sorting over time is an additional effect that raises the average ability of college graduates, since everyone who mistakenly attends college has too low of ability. By similar logic, increased sorting lowers the average ability of high school dropouts, but has ambiguous effects for the other two groups. The calibration algorithm compensates for this by changing the skill price growth rates to continue to fit the wage series. The key parameters (σ_a, σ_{a^*}) do not change much: σ_a is 0.46 instead of 0.55 in the baseline case. Accordingly, the ability bias in measured wage levels is slightly smaller in this experiment. This fact is shown in Table 8, which corresponds to Table 5 of the baseline case. The model now implies that time series wages understate skill price growth by 26-49 percentage points instead of 31-58 percentage points.

Theoretically, increased sorting has an ambiguous effect on the college wage premium: it raises the ability of college graduates and has an ambiguous effect on the ability of high school graduates. Our calibration predicts that the ability of high school graduates also rises, with with a small net positive net effect on the college wage premium. This result is displayed in Figure 14. The intuition is simple: in the 1906 cohort, a high school graduate has above average ability. Hence, most mistakes were made by those with low ability and high signals rather than high ability and low signals, and improved sorting generally raises the ability of high school graduates over time. Again, the quantitative adjustment is small. The baseline model predicts relative skill prices fall by 0.05 and changing relative ability explains the entire rise of the college wage premium; the model with improved sorting predicts relative skill prices rise by 0.04, so changing relative ability explains only 85% of the rise in the college wage premium.

Figure 14: Wage Premia and Relative Skill Prices With Increased Sorting



5 Conclusion

The U.S. experienced a dramatic expansion of education during the post-war period. Our findings suggest that this expansion was accompanied by a substantial decline in the abilities of students. As a result, measured wage growth substantially underestimates true wage growth. Changing abilities of different groups also imply large movements in relative wages. In particular, the entire increase in the college wage premium since 1950 can be attributed to the relative increase in the ability of college graduates compared with high school graduates.

Our findings have implications for the causes of the U.S. educational expansion which we plan to explore in future work. Previous research attributes the increase in schooling to skill biased wage growth (Restuccia & Vandenbroucke 2008). In our estimates, the relative wages of college graduates do not trend upwards over the period 1950 to 2000. This finding suggests that changes in the supply of educated labor may be important for the expansion of U.S. education.

References

[1] Bishop, John H. (1989). "Is the test score decline responsible for the productivity growth decline?" *American Economic Review* 79(1): 178-197.

[2] Bound, John; George Johnson (1992). "Changes in the structure of wages in the 1980s: an evaluation of alternative explanations." *American Economic Review* 82: 371-92.

[3] Bowles, Samuel; Herbert Gintis; Melissa Osborne (2002). "The Determinants of Individual Earnings: Skills, Preferences, and Schooling." *Journal of Economic Literature* 39(4): 1137–176.

[4] Card, David; Thomas Lemieux (2001). "Can falling supply explain the rising return to college for younger men? A cohort-based analysis." *Quarterly Journal of Economics* 116(2): 705-46.

[5] Cunha, Flavio; James Heckman; Salvador Navarro (2005). "Separating uncertainty from heterogeneity in life cycle earnings." *Oxford Economic Papers* 57: 191-261.

[6] Finch, F. H. (1946). *Enrollment Increases and Changes in the Mental Level.* Stanford University Press.

[7] Flynn, James R. (1984). "The mean IQ of Americans: Massive gains 1932 to 1978." *Psychological Bulletin* 95: 29-51.

[8] Flynn, James, R. (1999). "Searching for justice: The discovery of IQ gains over time." *American Psychologist* 54: 5-20.

[9] Geweke, John; Michael Keane (2000). "An empirical analysis of earnings dynamics among men in the PSID: 1968-1989." *Journal of Econometrics* 96: 293-356.

[10] Goldin, Claudia; Lawrence F. Katz (2008). *The Race Between Education and Technology.* Cambridge: Harvard University Press.

[11] Guvenen, Fatih (2007). "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?" *American Economic Review* 97: 687-712.

[12] Hanushek, Eric A.; Ludger Woessman (2008). "The role of cognitive skills in economic development." *Journal of Economic Literature* 46(3): 607-668.

[13] Heckman, James; Edward Vytlacil (2001). "Identifying the Role of Cognitive Ability in Explaining the Level of and Change in the Return to Schooling." *Review of Economics and Statistics* 83(1): 1-12.

[14] Herrnstein, R. J.; C. A. Murray (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press.

[15] Juhn, Chinhui; Kevin M. Murphy; Brooks Pierce (1993). "Wage Inequality and the Rise in Returns to Skill." *Journal of Political Economy* 101(3): 410-442.

[16] Juhn, Chinhui; Dae Il Kim; Francis Vella (2005). "The expansion of college education in the United States: Is there evidence of declining cohort quality?" *Economic Inquiry* 43(2): 303-315.

[17] Katz, Lawrence; David Autor (1999). "Changes in the wage structure and earnings inequality." In: *Handbook of Labor Economics*, Vol. 3, ed. O. Ashenfelter and D. Card. Elsevier.

[18] Laitner, John (2000) "Earnings within education groups and and overall productivity growth." *Journal of Political Economy* 108(4): 807-832.

[19] Mincer, Jacob (1974). *Schooling, Experience, and Earnings*. New York: Columbia University Press.

[20] Navarro, Salvador (2008). "Using Observed Choices to Infer Agent's Information: Reconsidering the Importance of Borrowing Constraints, Uncertainty and Preferences in College Attendance." Mimeo. University of Wisconsin.

[21] NLS User Services (1992). "NLS79 Profiles of American Youth. Addendum to Attachment 106. http://www.nlsinfo.org/ordering/display_db.php3.

[22] Rangazas, Peter (2002). "The Quantity and Quality of Schooling and U.S. Labor Productivity Growth (1870-2000)." *Review of Economic Dynamics* 5: 932-964.

[23] Restuccia, Diego; Guillaume Vandenbroucke (2008). "The evolution of U.S. education: a macroeconomic analysis." Mimeo. University of Toronto.

[24] Ruggles, Stephen; Matthew Sobeck (1997). "Integrated Public Use Microdata Series: Version 2.0." Minneapolis: Historical Census Projects, University of Minnesota.

[25] Taubman, Paul; T. Wales (1972). *Mental ability and higher educational attainment in the 20th Century.* Berkeley: National Bureau of Economic Research and Carnegie Commission on Higher Education.

Table 9: Summary statistics: Census data

	1950	1960	1970	1980	1990	2000
N	106.9	442.1	539.4	3398.3	3939.3	4432.0
Avg.school	9.9	10.5	11.5	12.6	13.4	13.8
Real wage	9.7	13.0	15.7	14.3	15.6	16.6
$w_{1,35}$	3.7	6.5	8.1	9.2	8.3	7.5
$w_{2,35}$	4.6	7.4	9.7	11.4	10.4	9.5
$w_{3,35}$	5.9	7.2	10.0	11.6	10.4	10.0
$w_{4,35}$	6.6	8.1	11.5	13.4	11.0	11.7

Notes: The table shows summary statistics for the Census sample. N is the number of observations (in thousands). Avg.school denotes average years of schooling. Real wage is the average real wage of all persons in the sample. $w_{s,35}$ denotes the average real wage of persons aged 35.

Appendix

A1 Census Data

Samples. We use 1% samples for 1950-1970 and 5% samples thereafter. In 1950, only sample line individuals report wages and hours worked. This reduces the effective sample size to only one quarter of the 1960 sample. As a result, the 1950 estimates of mean log wages by birth cohort are noisy. Table 9 shows descriptive statistics for each Census year.

Educational attainment. Our measure of educational attainment is the IPUMS variable EDUCREC. It distinguishes nine levels of education, which we aggregate into four groups: less than high school, high school, some college, and at least college completed.

Before proceeding, it is useful to discuss a technical detail in the construction of the educational attainment data. Figure 1 shows discrete jumps between adjacent cohorts that are observed in different Census years. One reason is that the wording of the educational attainment question changed between 1980 and 1990. Prior to 1990, HIGRADE recorded years of schooling completed. Since 1990, EDUC99 asks for the highest degree attained. This affects in particular whether people report high school or some college.

We do not see a compelling way of correcting this problem. Goldin & Katz (2008) use Current Population Survey data to estimate the changes in education between 1980 and 1990. Two problems prevent us from adopting their approach: (i) The magnitude of the mismeasurement likely changes from one Census year to the next. The reason is that differences in the educational attainment questions affect only a subset of the population. The size of this population changes with the distribution of educational attainment. (ii) We observe jumps in educational attainment also between 1970 and 1980, even though both years use the HIGRADE version of the attainment question.

The outstanding feature of the data is the large decline in the fraction of high school dropouts. The changes in the attainment questions affect mainly those who are the border between two degrees (e.g., high school vs. some college). Since most of those identified as dropouts in 1940 report less than 11 years of schooling, we are confident that they did not achieve a high school degree. We therefore believe that the decline in high school dropouts is real and not an artifact of the changing data collection.

Wages. We calculate hourly wages as the ratio of wage and salary income (INCWAGE) to annual hours worked. Annual work hours are the product of weeks per year times hours per week. For consistency, we use intervalled weeks and hours for all years. Where available we use usual hours per week. Wages are computed only for persons who report working "for wages" (CLASSWKR) and who work between 520 and 5110 hours per year.

All dollar figures are converted into year 2000 prices using the Bureau of Labor Statistics' consumer price index (CPI) for all wage earners (all items, U.S. city average).

Adjusted wages. Our model abstracts from wage variation due to demographic characteristics such as marital status or place of residence. We remove this variation from our wage data using standard wage regressions. We divide the population into groups according to age and educational attainment. For each group, we regress the logarithm of wages on indicators for marital status, race, region of residence, and urban status as well as age and schooling. The adjusted wage is defined as the measured wage net of effects due to covariates other than age and schooling. Wage variation due to schooling within education groups is also removed. All wage data reported in this paper are based on adjusted wages.

Aggregation. For consistency reasons we calculate all cohort and year aggregates from a matrix of summary statistics that is indexed by school group, birth year, and year (s, τ, t). For each cell, the matrix records mean log wages, aggregate earnings and hours, etc.

The data cover men aged 35-44, so that each cohort born between 1906 and 1965 is observed exactly once. The age range is chosen so that schooling is completed and most men participate full time in the labor market.

The educational attainment of birth cohort τ is defined as follows. Denote the mass in a given cell by $\phi(s,\tau,t)$. Then the fraction of cohort τ in group s is given by $\phi(s,\tau,t) / \sum_s \phi(s,\tau,t)$. Since cohorts are observed at different ages, the educational attainment estimates are not fully comparable. However, data for pseudo-cohorts suggest that educational attainment does not change substantially between the ages of 35 and 44.

The mean log wage of school group s at date t is defined as an equally weighted average of the mean log wage of all cohorts recorded at t. Denote the mean log wage of a cell by $\bar{w}(s,\tau,t)$. Then the mean log wage of group s is defined as $0.1 \sum_\tau \bar{w}(s,\tau,t)$.

Estimation of M_s. One of the calibration targets is M_s: the ratio of lifetime earnings to age 35 wages. M_s is constructed as follows. We estimate longitudinal experience wage and hours worked profiles for each school group. Lifetime earnings are defined as the present value of fitted wages times hours over the age range $T_s + 6$ through 68, discounted to age 18. M_s is given by lifetime earnings divided by the mean fitted wage at age 35.

Age profiles are estimated by regressing log wages (or hours) on an experience quartic and a birth year quadratic. These regressions pool all years and are separately estimated for each school group.

A2 NLSY79 Data

Schooling. We construct each person's highest grade attained and last year in school from annual reports of school enrollment and grade completed. These reports contain numerous inconsistencies. Many persons report single years of school enrollment late in life, often without any change in the highest grade attained. We treat such observations as invalid.

We calculate the last year of school as the last year in which the person reports enrollment and an increase in highest grade attained. Visual inspection of individual schooling histories suggests that this algorithm leads to sensible results. However, there is no unambiguous way of distinguishing valid from erroneous schooling observations.

Since our model does not permit persons to return to school once they started working, we delete 728 individuals who completed schooling at ages greater than years of schooling plus 12.

Wages. We calculate hourly wages as the ratio of labor income to annual hours worked. Labor income includes wages, salaries, bonuses, and two-thirds of business income. We

Table 10: Summary statistics: NLSY79 data

School class	Dropout	High school	Some college	College
Avg.school	9.9	12.0	13.8	17.0
Real wage at age 35	12.5	15.2	19.0	26.3
Adj. wage at age 35	11.1	13.8	17.7	22.3
AFQT percentile	0.25	0.43	0.58	0.77
N	1404	2034	888	882

delete wage observations prior to the last year of school enrollment or with hours worked outside the range $[520, 5110]$. We also delete wage observations outside the range $[0.02, 100]$ times the median wage. Wages are deflated by the CPI.

We remove from the wage data variation that is due to demographic characteristics not captured by our model. This is done by regressing log wages on schooling, experience, race, marital status, and region of residence. Separate regressions are estimated for each year and schooling group (high school dropouts, high school graduates, some college, and college+). As before, we construct adjusted wages, by removing the effects of race, marital status, region, and schooling sub-group within each education group.

For consistency with the Census data, we focus on wages earned at age 35. Since not all persons are interviewed at age 35, we interpolate these wages. For each person with at least 10 valid wage observations, we fit a quadratic experience wage profile using OLS. We use the wage predicted by this regression at age 35.

Summary statistics. Table 10 summarizes the data. For each school class, the table shows average years of schooling, the average AFQT percentile rank, the mean log wage at age 35, and the number of persons in the sample.

A3 PSID Data

This section describes how we estimate the variance of permanent wages. We use the 1968 to 2003 waves of the Panel Study of Income Dynamics (PSID). The sample contains all men who report at least 15 valid wage observations between the ages of 18 and 65. Wage observations are valid if hours worked fall in the interval $[520, 5110]$ and labor income is positive. Wages below 0.02 times the median or above 100 times the median are deleted.

Labor income includes wage and salary income as well as the labor income share of self-employment income. The real wage is defined as total labor income divided by total hours worked, deflated by the Consumer Price Index.

Estimating the stochastic process governing wages. Our estimation strategy follows Guvenen (2007). The first step is to form a residual wage. We pool all observations within a given school group and regress the log real wage on a quartic in experience. We assume that the residual wage is governed by a process of the form

$$\log(w_{j,t}) = \alpha_j + X_{j,t}\beta + \zeta_{j,t} + \varepsilon_{j,t} \tag{19}$$

$$\zeta_{i,t} = \rho\zeta_{i,t-1} + \hat{\varepsilon}_{i,t} \tag{20}$$

where the error terms $\varepsilon_{i,t} \sim N(0, \sigma_\varepsilon)$ and $\hat{\varepsilon}_{i,t} \sim N(0, \sigma_\zeta)$ are independently distributed. $\log(w_{j,t})$ is the log residual wage of person j at date t. It is composed of a fixed effect α_j, a persistent shock $\zeta_{j,t}$, and a transitory shock $\varepsilon_{j,t}$.

We estimate the parameters of the wage process by minimizing the sum of squared deviations between the empirical covariance matrix of wages and the one implied by the model (19). All deviations are equally weighted. Only elements of the empirical covariance matrix with at least 200 contributing individuals are retained.