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

Lutz Hendricks‡ Todd Schoellman‡
Very preliminary and incomplete.

November 6, 2008

Abstract

Since 1950, U.S. educational attainment has increased substantially. While the median student in 1950 dropped out of high school, the median student today attends some college. In an environment with ability heterogeneity and positive sorting between ability and school tenure, the expansion of education implies a decrease in the average ability of students conditional on school attainment. Using a calibrated model of school choice under ability heterogeneity, we investigate the quantitative impact of rising attainment on ability and measured wages. Preliminary findings suggest that the decline in average ability depressed wages conditional on schooling by 24-45 percentage points. We also find that the entire rise in the college skill premia since 1950 can be explained by the fact that the ability of college graduates has declined much slower than the ability of other groups.

JEL: I2, J24.

Key words: Education.

*For helpful comments we thank seminar participants at the Cleveland Fed and the Clemson University Bag Lunch. The usual disclaimer applies.
‡University of North Carolina, Chapel Hill, CESifo, Munich, and CFS, Frankfurt; lhendricks@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. 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.¹

This dramatic increase in enrollment raises the possibility that the average ability of students achieving a given level of schooling may have declined. In the 1920’s an average student dropped out of high school, while the typical college student belonged to a small elite. Today an average student attends college, while a minority drops out of high school. 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?

Background. Existing measures of human capital and wages commonly assume that student abilities remain constant over time. Based on this assumption, studies typically find large increases in human capital and in the relative wages earned by college graduates since around 1950 (see Goldin & Katz 2008 for a recent compilation of the evidence). We assess how these findings change when we account for the changing ability composition of schooling groups that goes hand in hand with rising educational attainment.

A common measure of human capital is based on Mincerian wage equations of the form \( h = e^{\phi s} \) where \( h \) is human capital and \( s \) denotes years of schooling (Mincer 1974). The growth rate of human capital between two dates is then proportional to the change in schooling: \( \Delta \log(h) = \phi \Delta s \). If the ability composition of education groups declines over time, this approach overstates human capital growth. One purpose of this paper is to measure the magnitude of this bias.

A common measure of relative wages by skill is the ratio of mean earnings of college educated to high school educated persons. Studies of relative wages typically condition on demographic characteristics (such as age and race), but not on ability. They find that the college wage premium exhibits a strong upward trend since 1950 (Goldin & Katz 2008). If the ability composition of college graduates rises relative to high school graduates, this calculation overstates the increases in relative college wages. The second purpose of this paper is to measure the magnitude of this bias.

¹See section 2 for details on how these statistics are constructed.
**Approach.** Changing student abilities will be important if ability is highly dispersed, and if students are strongly sorted by ability into education groups. Unfortunately, comparable, high quality measures of student abilities by education type 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 calibrated 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. 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. However, given that ability is not perfectly observed, the sorting will be 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 1940 to 2000. Specifically, using U.S. Census data, we estimate the educational attainment and wages earned by the cohorts born between 1906 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.

**Preliminary findings.** At all education levels, student abilities decline by large amounts as enrollment increases. As a result, conventionally measured wage growth substantially understates the growth of skill prices. For high school graduates, our model implies that skill prices grow nearly twice as fast as do 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.

1.1 Related Literature

A small number of previous studies have addressed the question we pose. Taubman & Wales (1972) collect data from several studies, including Finch (1946), that report aptitude or achievement test scores for high school graduates and for college students. They find that the test scores of college students rose over time relative to those of high school graduates. Our model’s implications are consistent with their findings. Juhn et al. (2005) have questioned the comparability of studies that use different aptitude tests and cover 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. For students attending grades 8, 9, and 12, his data show test scores rising until about 1965 before they start to decline. 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 Empirical Results

In this section, we document a number of features of U.S. data that both motivate our analysis and are useful for calibrating our model. Using Decennial Census data, we document (i) a dramatic increase in educational attainment since 1940 and (ii) a steep increase in the average wage earned by college graduates relative to high school graduates since 1950. Both observations have been described in the literature (see Goldin & Katz 2008 and the references therein). Using NLSY79 data, we characterize the joint distribution of schooling, aptitude test scores, and wages for the 1960 birth cohort. Consistent with Heckman & Vytlacil (2001), we find strong evidence of educational sorting by AFQT score.

2.1 Educational Attainment and Wages

Our database pools 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 attending school, who do not live in group quarters, and who report positive wage and salary income.

**Individual variables.** 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.
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 \((s, \tau, t)\). 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 \(\tau\) is defined as follows. Denote the mass in a given cell by \(\phi (s, \tau, t)\). Then the fraction of cohort \(\tau\) 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 \(0.1 \sum_\tau \bar{w} (s, \tau, t)\).

As a summary of our empirical work, we document the two facts outlined above: the rise in educational attainment and in the college wage premium.
Cohort educational attainment. Figure 1 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.

The main point we take away from Figure 1 is the following. Among those born around 1900, only a select few finished high school, let alone college. By 1965, more than half of the cohort attained at least some college. If the most able persons attain the highest degrees, the data suggest that high school graduates in 1900 represented the right tail of the ability distribution. By 1965, this is clearly no longer the case. This raises the possibility that the ability levels of students in high school and above may have declined substantially. One goal of this paper is to measure this change in ability.

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.

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.

The rise in relative college wages is often interpreted as an increase in the relative price of skilled labor, possible driven by skill biased technical change (Bound & Johnson 1992). One question we address is: To what extent is this rise due to the changing ability levels of college and high school graduates?

2.2 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.

Aptitudes. Our measure of 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.

**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 delete wage observations prior to the last year of school enrollment or with hours worked
Table 1: Summary statistics: NLSY79 data

<table>
<thead>
<tr>
<th>School class</th>
<th>Dropout</th>
<th>High school</th>
<th>Some college</th>
<th>College+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg.school</td>
<td>9.9</td>
<td>12.0</td>
<td>13.8</td>
<td>17.0</td>
</tr>
<tr>
<td>Real wage at age 35</td>
<td>12.5</td>
<td>15.2</td>
<td>19.0</td>
<td>26.3</td>
</tr>
<tr>
<td>Adj. wage at age 35</td>
<td>11.9</td>
<td>13.9</td>
<td>17.7</td>
<td>22.3</td>
</tr>
<tr>
<td>AFQT percentile</td>
<td>0.25</td>
<td>0.43</td>
<td>0.58</td>
<td>0.77</td>
</tr>
<tr>
<td>N</td>
<td>1404</td>
<td>2034</td>
<td>888</td>
<td>882</td>
</tr>
</tbody>
</table>

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 1 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.

2.2.1 Results: NLSY79 Data

Schooling and ability. Table 2 characterizes educational sorting by ability. For each school class, the table shows the fraction of persons falling into each ability quintile.

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.
Table 2: Schooling and Ability

<table>
<thead>
<tr>
<th>AFQT quintile</th>
<th>Dropout</th>
<th>High school</th>
<th>Some college</th>
<th>College+</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.499</td>
<td>0.200</td>
<td>0.091</td>
<td>0.014</td>
</tr>
<tr>
<td>2</td>
<td>0.278</td>
<td>0.273</td>
<td>0.167</td>
<td>0.036</td>
</tr>
<tr>
<td>3</td>
<td>0.140</td>
<td>0.255</td>
<td>0.240</td>
<td>0.133</td>
</tr>
<tr>
<td>4</td>
<td>0.067</td>
<td>0.192</td>
<td>0.287</td>
<td>0.266</td>
</tr>
<tr>
<td>5</td>
<td>0.015</td>
<td>0.080</td>
<td>0.215</td>
<td>0.550</td>
</tr>
<tr>
<td>Fraction</td>
<td>0.208</td>
<td>0.385</td>
<td>0.173</td>
<td>0.234</td>
</tr>
<tr>
<td>N</td>
<td>1404</td>
<td>2034</td>
<td>888</td>
<td>881</td>
</tr>
</tbody>
</table>

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.

Table 3: Mean wages by schooling and AFQT quintile: NLSY79 data

<table>
<thead>
<tr>
<th>AFQT quintile</th>
<th>Dropout</th>
<th>High school</th>
<th>Some college</th>
<th>College+</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10.47</td>
<td>11.09</td>
<td>13.08</td>
<td>16.41</td>
</tr>
<tr>
<td>2</td>
<td>11.01</td>
<td>12.20</td>
<td>15.32</td>
<td>20.29</td>
</tr>
<tr>
<td>3</td>
<td>11.42</td>
<td>13.00</td>
<td>16.37</td>
<td>18.12</td>
</tr>
<tr>
<td>4</td>
<td>12.16</td>
<td>13.17</td>
<td>16.68</td>
<td>19.70</td>
</tr>
<tr>
<td>5</td>
<td>–</td>
<td>14.24</td>
<td>15.91</td>
<td>20.73</td>
</tr>
</tbody>
</table>

Wages and ability. Table 3 shows the mean log wage at age 35 for each schooling class and AFQT quintile. The numbers are reported as exponentials (in year 2000 dollars) for ease of interpretation. In each school class, the highest ability quintile earns between 20% and 30% more than the lowest quintile.

For completeness, Table 4 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 (see Bowles et al. 2002 for a survey).
Table 4: Wage regressions: NLSY79 data

<table>
<thead>
<tr>
<th></th>
<th>Dropout</th>
<th>High school</th>
<th>Some college</th>
<th>College +</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.061</td>
<td>0.091</td>
<td>0.056</td>
<td>0.103</td>
</tr>
<tr>
<td>$\sigma_\beta$</td>
<td>0.026</td>
<td>0.016</td>
<td>0.028</td>
<td>0.040</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.01</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>$N$</td>
<td>595</td>
<td>1097</td>
<td>390</td>
<td>353</td>
</tr>
</tbody>
</table>

Note: The table regresses log wages at age 35 on AFQT score separately for each schooling group. $\beta$ is the estimated return to schooling, $\sigma_\beta$ is its standard error. $N$ is the number of observations.

3 A Model of School Choice

We develop a model of school choice to measure the changing ability composition of workers with different educational attainment since 1950.

Outline. 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 schooling 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 productivity of 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 $\tau$ and their age $a$, with period $t$ then given by $t = \tau + a - 1$.

Timing. The timing of events within a period is as follows. At birth, each worker is endowed with ability $A^L$. We use the superscript $L$ to denote that this ability is specifically
the ability to learn, or generate human capital from time spent in school. \( A^L \) is not directly observed. Instead, the worker draws a noisy signal of \( A^L \), denoted \( A^* \). Based on this signal, the worker chooses one of \( N \) 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.

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 explains why our analysis focuses on 35-44 year old workers.

We assume that \( \log(A_L) \sim N(\mu_L, \sigma_L) \). Conditional on \( A^L \), the distribution of the signal is \( \log(A^*) \sim N(\log(A^L), \sigma_{A^*}) \). We think of uncertainty as the inability of students to gauge how much they are really learning, and in particular the impact of current classes on future wages. We refer to the special case \( \sigma_{A^*} = 0 \) as the perfect information case. The dispersion of ability \( \sigma_L \) and noise \( \sigma_{A^*} \) are key parameters of the model.

**Schooling.** Upon completing school level \( s \), a worker with ability \( A^L \) is endowed with

\[ h(s, A^L) = (A^L)^{\eta_s} \]  

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 log-wages are normally distributed in our model, which is roughly consistent with the data. Second, if \( \eta_s \) 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 3.2), consistent with the positive correlation between aptitude test scores and schooling observed in U.S. data (see section 2.2).
**Wages.** At ages 1 through $T_s$, students are in school and do not work. After graduation, the worker supplies $h(s, A^L) e_{s,a} l_{s,a}$ units of type $s$ labor. $e_{s,a}$ is an exogenous age-efficiency profile. $l_{s,a}$ is an exogenous age hours profile. $l_{s,a} = 0$ while in school. $1 - l_{s,a}$ is consumed as leisure. In our calibration it is important that we match wages and lifetime earnings by school type, so it is important that we account for variation in hours and experience profiles.

A worker with school type $s$ earns a wage of $x_{t,z,s,t}$ per 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 per unit of work time. $x_t$ 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).

### 3.1 Worker’s Problem

At the stage of school choice, the worker’s state vector contains her ability signal and cohort: $z = (A^*, \tau)$. Upon graduation, the worker’s ability is revealed and the state vector is $z' = (A^L, A^*, \tau)$. We denote by $f(z'|z) = \Pr(A^L|A^*)$ the probability that an individual of type $z$ transitions into type $z'$.

Individuals order paths of consumption and leisure according to

$$
\sum_{a=1}^{T} \beta^a [\log(c_{s,z,a}) + \xi \log(1 - l_{s,a})] - \chi_{s,\tau}
$$

(2)

$\beta > 0$ is the discount factor and $\xi > 0$ determines the relative weight of leisure in the period utility function. Since workers operate in complete markets, they face no consumption risk. We may therefore write consumption $c_{s,z,a}$ as a function of schooling, the ability signal (contained in $z$) and age.

$\chi_{s,\tau}$ 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.

The budget constraint equates the present value of consumption spending with the
expected present value of lifetime earnings.

\[
\sum_{a=1}^{T} \frac{c_{s,z,a}}{R^a} = \int_{z'} f(z'|z) Y(s, z') \, dz'
\]  

where

\[
Y(s, z') = \sum_{a=T_{r+1}}^{T} \frac{x_{r+a-1} z_{s,r+a-1} l_{s,a} e_{s,a} h(s, A^L)}{R^a}
\]

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 $A^L$ contingent bond equals $f(z'|z)$. $R$ is the exogenous gross interest rate.

Denote by $V(s, z)$ the value function of a worker of type $z$ who chooses schooling level $s$; $V(s, z)$ is the solution to the problem of maximizing (2) subject to (3).

### 3.2 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: find the optimal allocation of consumption over time given school choice; then find the school choice that maximizes lifetime utility.

The lifetime consumption profile obeys the standard Euler equation

\[
c_{s,z,a+1} = \beta R c_{s,z,a}
\]

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

\[
c_{s,z,1} = \Lambda^{-1} \int_{z'} f(z'|z) Y(s, z') \, dz'
\]
Lifetime utility is then given by

$$V(s, z) = \sum_{a=1}^{T} \beta^a \left[ \log(c_{s,z,1}) + (a - 1) \log(\beta R) + \xi \log(1 - l_{s,a}) \right] - \chi_{s,\tau}$$  \hspace{1cm} (7)

$$= R \Lambda \log \left( \Lambda^{-1} \int_{z'} f(z'|z) Y(s, z') \, dz' \right) - \hat{\chi}_{s,\tau}$$  \hspace{1cm} (8)

where

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

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

Optimal school choice satisfies

$$s = \arg \max V(s, z)$$  \hspace{1cm} (10)

**Educational sorting.** We derive conditions under which the model implies positive sorting. Each worker’s school choice is determined by the value gap \(V(s + 1, z) - V(s, z)\). If this gap is positive, the worker prefers \(s + 1\) over \(s\). How does the gap change with the ability signal?

Note that \(Y(s, z')\) equals \(h(s, A^L)\) times a constant that does not depend on ability. Therefore

$$\frac{\partial V(s, z)}{\partial \log(A^*)} = R \Lambda \frac{\partial \log \left( \int_{z'} f(z'|z) Y(s, z') \, dz' \right)}{\partial \log(A^*)}$$  \hspace{1cm} (11)

$$= R \Lambda \frac{\partial \log \left( E(A^L)^{\eta_s} | A^* \right)}{\partial \log(A^*)}$$  \hspace{1cm} (12)

Consider first the case of *perfect information* where \(A^L = A^*\). In this case

$$\frac{\partial V(s, z)}{\partial \log(A^*)} = R \Lambda \eta_s$$  \hspace{1cm} (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 \(A^L\) 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 $\eta_s < \eta_{s+1}$, then students segment on expected ability, which in this model is driven by their ability signal. Given a student with expected ability $E(A^L)$ 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 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.

**AFQT.** When we take our model to the data, we interpret observed AFQT scores as a noisy signal of the worker’s ability signal $A^*$. In particular, we assume that the distribution of AFQT obeys $AFQT \sim N(\log(A^*), \sigma_{AFQT})$. Thus, workers are assumed to have better ability information than the econometrician.

### 3.3 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 ($z_{s,\tau}$).

To gain insight into the determinants of educational attainment, consider the indifference condition

$$V(s + 1, z) - V(s, z) = R \Lambda [\log(E[Y(s + 1, z') | z]) - \log(E[Y(s, z') | z])] + \hat{\chi}_{s,\tau} - \hat{\chi}_{s+1,\tau} = 0$$

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

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

where

$$M_s = \sum_{a=T_{s+1}}^{T} \frac{x_{s,\tau+a-1} z_{s,\tau+a-1} e_{s,a}}{R^a x_{s+34} z_{s,\tau+34} e_{s,35}}$$

17
Lifetime earnings has three components: human capital \((A^L)^{\eta_L}\), the wage earned per hour at age 35 (an arbitrary, fixed age), and the time invariant factor \(M_s\). The model therefore implies that schooling is time invariant if (i) schooling costs grow at the same rate for all levels, so that \(\hat{\chi}_{s+1,t} - \hat{\chi}_{s,t}\) 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 \(g_{X_s} > g_{X_{s+1}}\) then attainment level \(s + 1\) will become relatively less costly over time 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 \(\chi_s\) and \(z_s\).

Two other possible changes that we do not consider explicitly can also be accounted for in this model. First, an increase in the average ability of students \((\mu_L)\) is isomorphic to a particular parameterization of skill biased wage growth in our model. That is, for any change in \(\mu_L\), there exists a vector of \(z_{s,0}\) changes that leaves hourly wages and optimal education choices unchanged for all cohorts. Growth in \(\mu_L\) therefore does not affect our conclusions about the changes in relative wages and relative human capital growth rates that occurred in the U.S. since 1950.

Increases in education quality \((\eta_L)\) act similar 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.

4 Calibration of the Model

Qualitatively, our model indicates that the expansion of education has led to a decline in the average ability of students for each schooling level. The quantitative importance of this channel depends on the dispersion of ability of workers \((\sigma_L)\) and the degree of sorting between ability and schooling, controlled by the noise of the signal \((\sigma_{A^*})\). To evaluate these parameters and the size of biases in measured wages, we calibrate our model.
Our calibration simulates the outcomes that would have been viewed in the 1950-2000 Censuses for men aged 35-44 (the 1906-1965 birth cohorts). We let a model period correspond to one year, and assume workers live for $T = 70$ years. Our strategy is to allow the $\chi_{s, \tau}$ to vary freely so that we match the fraction of each cohort with each school attainment exactly. Since the focus of our model is how changing attainment affects average ability, this assumption ensures that we fit the main mechanism exactly. We then choose the rest of the parameters to maximize our fit to the data for wages, AFQT scores, and attainment by AFQT scores.

### 4.1 Parameters

The model has 20 parameters that need to be pinned down: the intertemporal parameters $R, \beta$; the productivity parameters, $x, g_x, z_x, g_z$; the education-ability complementarity parameters, $\eta_s$; and the moments of the distributions, $\mu_L, \sigma_L, \sigma_{A^*}, \sigma_{AFQT}$. For all level parameters that grow over time, we take the base year to be 2000.

Some of these parameters are taken outside the model to be consistent with outside evidence. We set $R = 1.05$ for an exogenous interest rate of 5%, and fix $\beta = 1/R$. Some of the parameters cannot be identified separately and must be normalized. For instance it is not possible to identify a general TFP term $x$ and a school-specific productivity term for each school group. We normalize $z_1 = 1$ and $g_{z_1} = 0$. In all the expressions in the model it is $\eta_s \sigma_L$ that matter, so we cannot identify $\sigma_L$ and a full set of $\eta_s$ parameters separately. We normalize $\eta_1 = 1$. Finally, the mean of the ability distribution cannot be separately identified from the level of TFP and school-specific productivity, so we normalize $\mu_L = 1$.

The remaining 14 parameters are calibrated in the model: $(x_{2000}, g_x, z_{2000,2}, z_{2000,3}, z_{2000,4}, g_{z2}, g_{z3}, g_{z4}, \eta_2, \eta_3, \eta_4, \sigma_L, \sigma_{A^*}, \sigma_{AFQT})$.

### 4.2 Moments

Most of the data moments that we target were discussed in Section 2; we review them here, and discuss what parameters they help us pin down. First, we fit the educational attainment of each cohort of men exactly. This moment corresponds to the data in Figure 1. To pin down the level and changes of TFP and school-specific productivity, we target average wages for men aged 35 for each school group in the 6 Censuses, corresponding to Figure 2.
We use AFQT scores to give us some additional information on sorting and wage differences by ability. However, AFQT is a noisy measure of the agent’s signal (which is itself a noisy measure of true ability). To evaluate the amount of noise in AFQT relative to the signal, we use the degree of degree of sorting in education by AFQT quintiles, as given in Table 2. Students sort perfectly based on their school choices; the imperfect sorting seen in the data informs us about the magnitude of noise in AFQT scores.

The key parameters are the measures of effective dispersion, $\eta_2, \eta_3, \eta_4, \sigma_L$, and the degree of sorting by true ability, determined by the noise in students’ signals, $\sigma_{A^*}$. We use two pieces of information to pin down these numbers. First, we use the returns to AFQT by school group, from Table 4. In our model, returns to true ability by school group are $\eta_s$. By substituting AFQT scores, we introduce selection and measurement error. Our model uses the calibrated parameters $\sigma_{AFQT}$ and $\sigma_{A^*}$ to internally simulate these forces in the model as well.

Second, we use information on the dispersion of earnings. In our model, the dispersion of wages and lifetime earnings conditional on schooling are both proportional to $\eta_s^2 \sigma_L^2$, so dispersion is an informative way to pin down the key parameters. Our model abstracts from luck and other transitory shocks to earnings, so it is important that we construct data moments that fit with this idea. Following Guvenen (2007) and others, we do so by estimating the variance of “permanent earnings”, the predictable, non-transitory component of earnings.

We think of log-wages of individual $j$ at time $t$ as being generated by an autoregressive earnings process with an individual-specific fixed component:

$$\log(w_{j,t}) = \alpha_j + \rho \log(w_{j,t-1}) + X_{j,t}\beta + \epsilon_{j,t}$$

where $X$ is the vector of the individual’s characteristics, $\beta$ is the impact of those characteristics on wages, and $\epsilon$ is the shock to wages which is propagates through the AR(1) process. The moment of interest here is the conditional variance of $\alpha_j$. We estimate this income process using PSID data. Results are given in Table 5. Details are available in the Appendix A1.

In using this approach, we assume that variation in $\alpha_j$ 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.
### Table 5: PSID wage regressions

<table>
<thead>
<tr>
<th>Schooling</th>
<th>Std PV LTY</th>
<th>$\sigma_\alpha$</th>
<th>$\rho$</th>
<th>$\sigma_\eta$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dropout</td>
<td>0.389</td>
<td>0.329</td>
<td>0.891</td>
<td>0.174</td>
<td>0.286</td>
</tr>
<tr>
<td>High school</td>
<td>0.390</td>
<td>0.270</td>
<td>0.973</td>
<td>0.110</td>
<td>0.327</td>
</tr>
<tr>
<td>Some college</td>
<td>0.359</td>
<td>0.285</td>
<td>0.881</td>
<td>0.192</td>
<td>0.268</td>
</tr>
<tr>
<td>College</td>
<td>0.444</td>
<td>0.242</td>
<td>0.969</td>
<td>0.154</td>
<td>0.335</td>
</tr>
</tbody>
</table>

Note: The table shows the estimated coefficients obtained from wage regressions using PSID data.

### Table 6: Model parameters

<table>
<thead>
<tr>
<th>Moment</th>
<th>Value</th>
<th>Moment</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_{2000}$</td>
<td>4.138</td>
<td>$g_x$</td>
<td>1.06%</td>
</tr>
<tr>
<td>$z_{2000,2}$</td>
<td>0.855</td>
<td>$g_{z_2}$</td>
<td>0.40%</td>
</tr>
<tr>
<td>$z_{2000,3}$</td>
<td>0.661</td>
<td>$g_{z_3}$</td>
<td>0.14%</td>
</tr>
<tr>
<td>$z_{2000,4}$</td>
<td>0.559</td>
<td>$g_{z_4}$</td>
<td>0.28%</td>
</tr>
<tr>
<td>$\eta_1$</td>
<td>1.000</td>
<td>$\sigma_L$</td>
<td>0.462</td>
</tr>
<tr>
<td>$\eta_2$</td>
<td>1.012</td>
<td>$\sigma_{A^*}$</td>
<td>0.316</td>
</tr>
<tr>
<td>$\eta_3$</td>
<td>1.019</td>
<td>$\sigma_{AFQT}$</td>
<td>0.795</td>
</tr>
<tr>
<td>$\eta_4$</td>
<td>1.021</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 4.3 Calibration Procedure and Parameter Values

We use a two-step procedure to calibrate the model. In the first step we calibrate $\eta_2, \eta_3, \eta_4, \sigma_L$, $\sigma_{A^*}$, and $\sigma_{AFQT}$, targeting the moments on education sorting by AFQT, returns to AFQT, and variability of earnings. This step informs us about the dispersion of and sorting by true ability. In the second step we take the ability of agents by cohort-school type as given and choose the $x_{2000}, g_x, z_{2000,2}, z_{2000,3}, z_{2000,4}, g_{z_2}, g_{z_3}, g_{z_4}$ that best fit average wage data. In practice there is little interaction between parameters, so fitting all parameters and moments jointly produces a very similar calibration result. Our loss function between the model and the data is standard least squares. The resulting parameters are given in Table 6.

The key model parameters are $\eta_\kappa$ and the standard deviations of the various ability measures. We find that AFQT is a very noisy signal of ability. The implied degree of ability dispersion is large, although the effective dispersion $\eta_\kappa \sigma_L$ is similar across school groups.

To understand the magnitude of ability dispersion, consider a world with perfect edu-
cational sorting by ability. The cohort aged 35 in 1950 contains 53% high school dropouts and 30% high school graduates. With perfect sorting, the latter are drawn from the 53rd to the 83rd percentile of the ability distribution. Their mean log ability then equals 0.22; a direct calculation of the mean of a truncated Normal distribution with standard deviation \( \sigma_L \). Moving forward to the year 2000, the cohort aged 35 is drawn from the 10th to the 43rd ability percentiles, with a mean log ability of \(-0.30\). The expansion of schooling leads to a 52% reduction in mean log ability and a slightly larger reduction in mean high school wages. By comparison, the actual change in high school wages between 1950 and 2000 is near 30%. This example illustrates that the calibrated dispersion of abilities is large enough to potentially imply very large movements in wages.

Of course, the actual ability changes implied by the model differ from these simple calculations because households observe noisy signal of their abilities and therefore cannot sort perfectly into school groups. The degree to which students choose ex post suboptimal education levels is governed by the dispersion of noisy in the ability signal, \( \sigma_A \). Its value is about 60% of \( \sigma_L \). This means that about 40% of the household's ability signal is noise. More noise means less sorting, which dampens the effects of enrollment changes on abilities.

A set of parameters that is closely related to our findings is the set of wage growth rates, \( g(x) + g(z_s) \). They range from 1.0% for dropouts to 1.4% for high school. These are substantially higher than the wage growth rates we estimate from the Census data. Our model attributes the differences to changes in worker abilities over time. This foreshadows the finding, discussed in section 5, that measured wage growth substantially underestimates the growth of skill prices.

### 4.4 Model Fit

In this section, we evaluate the model’s ability to replicate the calibration targets. Figure 3 plots the mean log wages of persons aged 35-44 in each Census year relative to the wages earned by high school graduates. The model accounts well for the long-run trends in relative wages. The main feature the model misses is 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 section 2.1.

Figures 4a through 4d shows the density of AFQT scores by schooling level. The model’s implications are compared with the NLSY79 data discussed in section 2.2. Overall, 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 5 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, but especially so in the middle school groups. To understand this, note that the returns in the model are \( \eta_s \frac{d \log A_L}{dAFQT} \). \( \eta_s \) 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.

Figure 6 compares the variance of permanent earnings in each school group with the data described in section 4.2. The main deviation from the data occurs among high school dropouts where the data show a much larger variance than for the other groups. In part, this may be a result of aggregating persons with very different years of schooling. The model also overstates the dispersion among college graduates.

In the model, the variance is roughly the same for all \( s \). Aside from high school dropouts, this is consistent with the data. To see why, note that variance comes from two sources: variance in true ability from workers who choose the ex-post optimal schooling, and variance from workers with unusually high or low ability who choose schooling that is not ex-post
optimal. The two outer categories have higher contributions from the former source, since they include the tails of the ability distribution, but lower contributions from the latter, since workers are less likely to mistakenly attain school levels in the tail. The net contribution of the two sources results in reasonably constant variances.

5 Results

The main question we address in this paper is: To what extent are changing skill premia due to changing skill prices or to changing ability compositions of the different education groups? Figure 7 provides an answer to this question. For each school group, figure 7 shows...
two paths of relative wages. The solid lines are the paths of hourly earnings predicted by the calibrated model. These were already shown in figure 3. The dashed lines show the evolution of relative skill prices, $z_s - z_2$. All wages are in logarithms and expressed relative to high school graduates.

The main message of figure 7 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 relative wage movements. In our model, the average ability of workers in all school groups declines as education expands over time. Moreover, the average ability relative to high school graduates declines 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 23% since 1950, our model implies that the relative price of college educated labor declined by 6%. 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 7 summarizes the changes in wages relative to high school graduates over the period 1950-2000. Measured relative wages declined for high school dropouts and some college, but rose for college graduates (column “data”). The “model” column shows the corresponding changes in relative wages implied by the model. The “fixed ability” column
Figure 6: Dispersion of permanent earnings

![Graph showing dispersion of permanent earnings]

Table 7: Counterfactual Premia Increases

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Baseline Model</th>
<th>Fixed Ability</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;HS-HS</td>
<td>-0.14</td>
<td>-0.15</td>
<td>-0.2</td>
<td>-0.05</td>
</tr>
<tr>
<td>SC-HS</td>
<td>-0.03</td>
<td>-0.05</td>
<td>-0.13</td>
<td>-0.08</td>
</tr>
<tr>
<td>C+-HS</td>
<td>0.23</td>
<td>0.14</td>
<td>-0.06</td>
<td>-0.2</td>
</tr>
</tbody>
</table>

Note: The table shows changes in mean wages relative to high school graduates, 1950-2000. “Data” refers to Census data. “Fixed ability” denotes the changes in model skill prices, $z_s - z_2$.

displays the change in relative skill prices, $z_s - z_2$. In each case, skill prices rise more slowly that measured wages, but the discrepancy is particularly large for college graduates.

Table 8 shows the growth rates of wages for all education groups. Its layout is analogous to that of table 7. The main message is that large drops in ability mask substantial wage growth over the period 1950-2000. A large literature has pointed out that real wages have barely increased since about 1960 (Katz & Autor 1999). Our findings suggest that skill prices may have grown substantially, but measured wages are pushed down by the declining abilities of the students attending each school level.

To illustrate, the measured mean wage of a high school graduate rose by 35% between 1950 and 2000. Our model implies that the rental price of high school labor grew nearly twice as fast (by 72%). However, the expansion of schooling led to a drop in the average ability of high school graduates of 45%, which erodes much of the growth in wages.
Figure 7: Model and Counterfactual (Fixed Ability) Wage Premia

Table 8: Counterfactual Wage Increases

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Baseline Model</th>
<th>Fixed Ability</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;HS</td>
<td>0.2</td>
<td>0.12</td>
<td>0.52</td>
<td>0.4</td>
</tr>
<tr>
<td>HS</td>
<td>0.35</td>
<td>0.28</td>
<td>0.72</td>
<td>0.45</td>
</tr>
<tr>
<td>SC</td>
<td>0.31</td>
<td>0.23</td>
<td>0.59</td>
<td>0.37</td>
</tr>
<tr>
<td>C+</td>
<td>0.58</td>
<td>0.42</td>
<td>0.66</td>
<td>0.24</td>
</tr>
</tbody>
</table>

5.1 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.

Given the calibrated parameters, the model implies strong sorting. Figures 8a through 8b show 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 40% of the signal’s standard deviation is noise.

Another way of assessing the degree of sorting examines how many persons ex post regret their schooling choices. Table 9 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.
Table 9: Ex ante and ex post optimal schooling

<table>
<thead>
<tr>
<th>Ex-Post Optimal</th>
<th>Actual School Choice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;HS</td>
</tr>
<tr>
<td>&lt;HS</td>
<td>75.5%</td>
</tr>
<tr>
<td>HS</td>
<td>23.5%</td>
</tr>
<tr>
<td>SC</td>
<td>1.0%</td>
</tr>
<tr>
<td>C+</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

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.

An example with perfect sorting: To gain further intuition, consider a version of our model with perfect information about ability. It is easy to solve this model, assuming that all calibrated parameters are unchanged. Since its implications are quite similar to the model with imperfect sorting, it is useful for understanding how our findings come about.

Figure 9 shows the schooling decisions and abilities of the cohort aged 35 in 1950. The bell shaped curve is the density of $\log(A_L)$. The vertical lines represent the ability cutoffs that delineate the schooling levels. For example, the least able 53% of the population choose not to complete high school. Their mean $\log(A_L)$ equals $-0.35$.

Now forward to the year 2000. Figure 10 shows the school choices for the cohort aged 35 in that year. Only 10% of the population fail to graduate from high school. Losing the right tail of the ability distribution reduces the mean $\log(A_L)$ for high school dropouts to $-0.81$. The decline in mean ability is 0.46, compared with 0.4 in the calibrated model.

Other education groups yield similar results. For high school graduates, the perfect sorting model implies a decline in mean $\log(A_L)$ of 0.52, compared with 0.45 for the imperfect sorting model. The message is that the perfect information model yields similar ability changes as does the imperfect information model. Note that the resulting wage changes are similar to the ability changes because all $\eta_s$ are close to 1.

The perfect sorting example shows why the relative ability of college graduates rises over time. The expansion of schooling adds lower ability students to the college population. Since college is the highest attainable degree, the right tail is not affected. By contrast, the high school group loses its highest ability members and gains lower ability students who previously would have dropped out. Moreover, the expansion of the some college group pushes the high school group further to the left. This reasoning suggest that the increase in relative college abilities is a robust feature of our model.
6 Robustness
In progress.

7 Conclusion
In progress.
References


Table 10: Summary statistics: Census data

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>106.9</td>
<td>442.1</td>
<td>539.4</td>
<td>3398.3</td>
<td>3939.3</td>
<td>4432.0</td>
</tr>
<tr>
<td>Avg.school</td>
<td>9.9</td>
<td>10.5</td>
<td>11.5</td>
<td>12.6</td>
<td>13.4</td>
<td>13.8</td>
</tr>
<tr>
<td>Real wage</td>
<td>9.7</td>
<td>13.0</td>
<td>15.7</td>
<td>14.3</td>
<td>15.6</td>
<td>16.6</td>
</tr>
<tr>
<td>$w_{1,35}$</td>
<td>3.7</td>
<td>6.5</td>
<td>8.1</td>
<td>9.2</td>
<td>8.3</td>
<td>7.5</td>
</tr>
<tr>
<td>$w_{2,35}$</td>
<td>4.6</td>
<td>7.4</td>
<td>9.7</td>
<td>11.4</td>
<td>10.4</td>
<td>9.5</td>
</tr>
<tr>
<td>$w_{3,35}$</td>
<td>5.9</td>
<td>7.2</td>
<td>10.0</td>
<td>11.6</td>
<td>10.4</td>
<td>10.0</td>
</tr>
<tr>
<td>$w_{4,35}$</td>
<td>6.6</td>
<td>8.1</td>
<td>11.5</td>
<td>13.4</td>
<td>11.0</td>
<td>11.7</td>
</tr>
</tbody>
</table>

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

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.

Summary statistics. Table 10 shows descriptive statistics for each Census year.