

The Option Value of Human Capital: Higher Education and Wage Inequality*

Sang Yoon (Tim) Lee[†] Yongseok Shin[‡] Donghoon Lee[§]

June 2017

Abstract

Going to college is a risky investment in human capital. However, two options inherently embedded in college education mitigate the risk: (i) college students can quit without completing four-year degrees after learning about their post-graduation wages and (ii) college graduates can take jobs that do not require four-year degrees (i.e., underemployment). These options reduce the chances of falling in the lower end of the wage distribution as a college graduate, rendering standard mean-variance calculations misleading. We show that the interaction between these options and the rising wage dispersion, especially among college graduates, is key to understanding the muted response of college enrollment and graduation rates to the substantial increase in the college wage premium in the United States since 1980. Furthermore, expanding subsidies to induce more students to attend college has a negligible net benefit: Due to selection at the enrollment stage, marginal students who only enroll thanks to larger subsidies are far more likely to drop out of college or become underemployed even with a four-year degree, implying small wage gains from college education.

Keywords: Wage inequality, educational attainment, underemployment

*The views expressed in this article are those of the authors and do not necessarily reflect the position of the Federal Reserve Banks of New York and St. Louis or the Federal Reserve System.

[†]Toulouse School of Economics and CEPR; E-mail: sylee.tim@tse-fr.eu.

[‡]Washington University in St. Louis, Federal Reserve Bank of St. Louis and NBER; E-mail: yshin@wustl.edu.

[§]Federal Reserve Bank of New York; E-mail: donghoon.lee@ny.frb.org.

1 Introduction

In the early 1980s, American men with at least four years of college education earned about 40 percent more on average than those whose education ended with high school. By 2005, this college wage premium rose to above 90 percent. During the same time period, the fraction of men with a four-year college degree in the working-age population all but remained constant. This masks a slight increase in the fraction of high school graduates who enter college (enrollment rate) being offset by a decrease in the fraction of college enrollees who eventually earn a four-year degree (graduation rate). Also during the same period, overall wage inequality increased substantially, even among men with the same level of education. This rise in within-education-group wage dispersion was sharpest among college graduates, a fact rarely noted in the literature.

This paper develops a quantitative model of educational choice to explain why the enrollment and graduation rates were unresponsive to the rising college wage premium. In the model, individuals differ in their returns to college—the wage gains from completing four-year college education—and make sequential decisions: first, whether or not to enroll in college, and then whether or not to graduate with a four-year degree. The individual returns are not fully known to students until they enter labor markets, making both college enrollment and graduation risky investments.

However, we emphasize two real options inherent in college education that render standard mean-variance trade-off calculations misleading. First, college enrollees may choose not to complete a four-year degree after learning more about their individual returns while in college. In the data, this would include students explicitly dropping out of colleges (more commonly two-year colleges), as well as those who complete a two-year college but do not transfer to a four-year college. We will refer to them collectively as “some-college” or “college dropouts.” The second option is underemployment: college graduates may take jobs that do not require a four-year degree. This operationalizes the idea that the wage distribution faced by a college graduate stochastically dominates the one he would have faced without a four-year degree.

We find that an increase in the dispersion of the returns-to-college distribution and how it interacts with the two options above are important for understanding the changes in education and labor market outcomes. A mean-preserving spread of the returns-to-college distribution does not shift the realized wage distribution of college graduates symmetrically because of the options. Those with higher returns to college do graduate and pull the observed college wage premium higher. However, those who learn that their own returns are lower than expected either quit college (and hence fall out of the calculation for college premium) or become underemployed upon graduation (which truncates the left tail of the college wage distribution). This results in a higher

college wage premium *ex post*.

This explains why the fraction of college graduates (among men) did not increase in 2005, despite the much higher college wage premium. An across-the-board increase in the returns to college would have resulted in a substantial rise in both enrollment and graduation, contrary to the data. Indeed in our calibration, the population mean of the returns-to-college distribution does not rise nearly as much (only by 12 percentage points) as the observed college wage premium (by 57 percentage points) between 1980 and 2005. We find that more than half of the rise in the observed college wage premium is driven by a mean-preserving increase in the *variance* of the returns to college, both in terms of individual heterogeneity and risk—although the two work through different economic mechanisms. In contrast, only one-fifth of the college premium increase is explained by a higher population mean return to college. At the same time, the larger returns-to-college dispersion makes college education a riskier investment *ex ante*, counteracting the positive effect on college enrollment and graduation rates from the 12-percentage-point increase in the mean returns to college.

The distinctive features of our model—sequential college enrollment-graduation decisions and underemployment options—are also consistent with several empirical facts largely neglected in the literature. While the college wage premium increased from 36 percent to 93 percent between 1980 and 2005, the wage premium of those who have some college education but no four-year degree (i.e., some-college wage premium over the high-school-only group) only increased from 6 percent to 21 percent. Over the same period, even among college graduates, the wage premium of those underemployed increased from 11 to 46 percent. In other words, the headline increase in the college wage premium is primarily driven by the wage increase among college graduates with jobs that *do* require four-year degrees (from 44 to 116 percent over the same period).

The underemployment option, in particular, provides two useful economic insights. First, increases in heterogeneity versus *ex-post* residual risk have opposing effects on the underemployment rate, allowing us to separate their relative contributions to the increased wage inequality among college graduates. Second, because underemployment protects college graduates from the left tail of the wage distribution in a probabilistic sense, an increase in the second moment of the returns-to-college or residual wage distribution raises the third moment (skewness) of the resulting college-graduate wage distribution, as in the data.

Our model also provides a new perspective on the debate of whether too few or too many students are going to college.

For marginal high school graduates in our model who are indifferent between enrolling in college or not, the expected wage gains from enrollment are small. If we nudge such marginal students to go to college, using 2005 figures, 95 percent of them would

quit college, consequently raising their wage by only 15 percent on average compared to what they would have earned without enrolling. Even the other five percent who would eventually graduate from colleges only earn a wage premium of 33 percent—roughly one-third of the college premium in the overall population. As a consequence, the marginal students’ wage gains from college enrollment are only 16 percent on average.

The low wage gains are not surprising once we realize that marginal students do not go to college because they have realistic, low expectations of their returns to college in the first place. We note that our result is much smaller than the estimates in the empirical literature, e.g. [Card \(1999\)](#). The main reason for this discrepancy is that we explicitly account for the “selection on gains” that arises when students decide whether or not to graduate from colleges, based on signals of their post-graduation returns. Although much of the literature addresses positive selection into college enrollment following [Heckman et al. \(2006\)](#) and others, typically the next stage of selection—whether a marginal college student will graduate or not—is not explicitly addressed, likely biasing upward the returns-to-college estimates, e.g., [Carneiro et al. \(2011\)](#).

Extending this exercise, we find that college tuition subsidies designed to push more students into colleges will only have a negligible net benefit. If we make the first two years of college free in the 2005 benchmark, the enrollment rate increases by almost 10 percentage points (from 65 to 75 percent). Those whose enrollment decision is changed by the subsidies (“switchers”) are mostly marginal students, and the model predicts that 97 percent of them will quit college before completing a four-year degree. As a consequence, the average wage gains from the subsidized college enrollment are only 13 percent. In addition, while more switchers come from poorer backgrounds, rich switchers see larger wage gains than poor switchers, mainly because initial (family) wealth is positively correlated with the individual returns.

In a related analysis, we find that subsidizing college graduation rather than enrollment generates even smaller benefits. The explanation is that selection on individual returns is sharper at the graduation stage than at the enrollment stage, because students learn about their individual returns while they are in college.

Finally, on the opposite end of the debate, our model shows that the existence of underemployed college graduates does not necessarily mean that too many students are going to (and graduating from) college. Underemployment is a natural outcome of the real option inherent in the risky human capital investment of college education.

Contribution to the Literature This paper analyzes the linkage between the rising within-education-group wage dispersion and the changes in educational attainment over time, which has been hitherto unexplored in the literature. The existing papers on trends in educational attainment typically focus on between-education-group inequal-

ity (i.e., college premium). We find that the second moments of the individual returns to college (in terms of both heterogeneity and risk) are important for understanding the trends in educational attainment and college premium.

Our model departs from the college vs. high school dichotomy, and incorporates college dropouts and underemployment. The assumption that college students decide whether or not to graduate after learning more about their post-graduation wages separates our model from the few existing papers on college dropouts, in which students quit college for reasons that are orthogonal to their wage gains from college education (e.g., preference shocks). Underemployment is a modeling element we are introducing to the literature, although the notion itself is not entirely new.

By explicitly modeling these two real options, we add to the literature in three ways. First, by interacting the options and the rising wage dispersion (most conspicuously among college graduates), we show that standard mean-variance trade-off calculations may be misleading. In particular, a large part of the rise in the observed college premium is a direct result of a mean-preserving spread in individual heterogeneity and risk. Second, we quantify how strong the positive selection into college *graduation* is, and show that this margin is important for computing the hypothetical gains to college education of marginal high-school graduates. Finally, we draw attention to a few trends in the data that are often overlooked: changes over time in some-college premium, underemployment premium, and wage inequality *within* education groups.

The trends in college premium and college enrollment are analyzed in [Heckman et al. \(1998\)](#) and [Lee and Wolpin \(2006\)](#). They focus on the schooling response to an increase in the mean difference between high- and low-skilled workers, but do not distinguish between college enrollment and completion. However, [Bailey and Dynarski \(2011\)](#) and [Bound and Turner \(2011\)](#), for example, emphasize that college graduation rates decreased while enrollment rates increased between 1980 and 2000. This fact defies a simple explanation based on population-wide increases in education premia, and calls for an explicit modeling of the college completion as well as enrollment decisions.

In [Heckman and Urzua \(2009\)](#), [Stange \(2012\)](#), [Stinebrickner and Stinebrickner \(2012\)](#), and [Hendricks and Leukhina \(2017\)](#), students decide whether or not to enroll and graduate from college after obtaining new information. One key difference is that this decision in our model is directly based on the expected post-graduation labor market outcomes rather than, say, students' academic ability that predicts their chance of completing the degree requirements. In addition, these papers do not consider the interaction between this option and the second moment of the payoff distribution.

Another paper that features college dropouts is [Athreya and Eberly \(2013\)](#), which emphasizes “failure shocks” that exogenously oust students from college. This contrasts with the assumption in our model (and also in the above papers) that dropping out is

a voluntary decision negatively correlated with the expected gains from college. Our modeling choice is more in line with the empirical evidence on the predictability and selection in dropout/graduation decisions—e.g., [Altonji \(1993\)](#), [Bowen et al. \(2009\)](#), [Bound and Turner \(2011\)](#), [Stinebrickner and Stinebrickner \(2014\)](#), and [Hendricks and Leukhina \(2017\)](#).

On underemployment, there is only a nascent empirical literature, in which the same phenomenon is typically viewed as skill mismatch and accordingly labeled “overeducation,” e.g., [Leuven and Oosterbeek \(2011\)](#) and [Clark et al. \(2015\)](#).

On the empirical side, [Lemieux \(2010\)](#) documents that much of the rise in wage dispersion coincides with the rise in the college premium and that much of the increasing dispersion originates from more educated workers. [Autor et al. \(2008\)](#) focuses on the difference in residual wage inequality at different income percentiles. While suggestive of the increasing dispersion in the returns to college, these empirically-oriented papers do not consider its effect on the college enrollment and graduation decisions.¹

Finally, our model allows us to decompose the wage distribution into ex ante heterogeneity vs. ex post risk, which is the focus of [Cunha et al. \(2005\)](#) and [Chen \(2008\)](#). Our decomposition results are consistent with their finding that much of the wage dispersion among more educated workers is predictable from individual heterogeneity.

2 Empirical Facts

Our primary data source is the Integrated Public Use Microdata Series of the Current Population Survey March supplement (CPS). One challenge in making consistent comparisons over time is that the CPS has continuously revised variable definitions. To address this problem, we closely follow the procedure of [Autor et al. \(2008\)](#) (AKK hereafter) as described in Appendix A.1. Our data analysis is based only on white, male, full-time, full-year workers aged 26–50, and does not make any meaningful distinction between wages and earnings.

2.1 Educational Attainment

Figure 1(a) shows the fraction of our sample in each educational attainment category from 1965 to 2007. People are sorted into bins based on their highest levels of education: high school dropouts (HSD), high school graduates (HSG), some college but no four-year degree (SMC), four-year college degree (CLG), and at least some graduate school education (GTC for “greater than college”). While there is a clear increasing trend

¹[Brown et al. \(2015\)](#) computes the risk-adjusted college wage premium over time. However, it does not delve into individual education decisions or selection based on return heterogeneity.

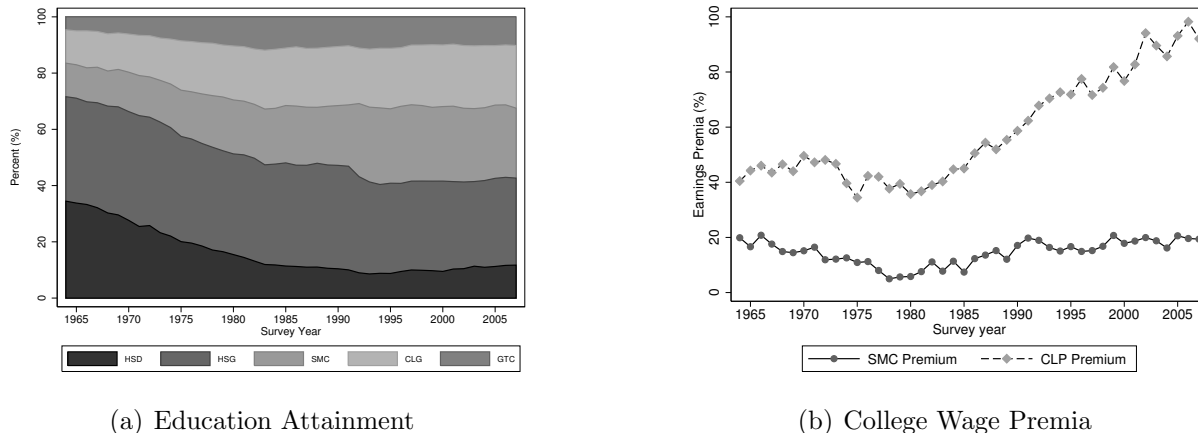


Fig. 1: Educational Attainment and Wage Premia

White males, ages 26-50. HSD: high school dropout, HSG: high school graduate, SMC: some college, CLG: four-year-degree but no graduate school, GTC: graduate school, CLP: college graduates (CLG+GTC). Wage premia are relative to HSG.

for both high school and higher categories until the early 1980s, all categories have remained remarkably stable since the late 1980s.²

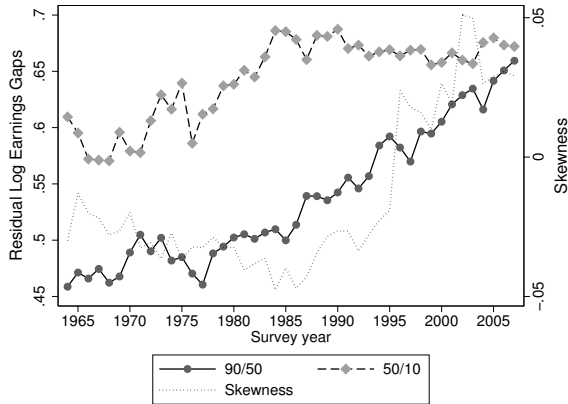
Throughout the observed period, almost half of those who enroll in college do not obtain a four-year degree (SMC divided by the sum of SMC, CLG, and GTC). Prior to 1992, it is not possible to break SMC further down into those with a two-year degree and those who drop out of two- or four-year institutions in the CPS. For recent years, by combining data from CPS, NLSY97, and [National Student Clearinghouse Research Center \(2012\)](#), we find that over 60 percent of those in SMC are dropouts (the majority of them from two-year institutions) throughout the sample period (1992–2007), 10–15 percent have academic degrees, and the rest are vocational degree holders. Another observation is that a significant portion of those earning a four-year degree continue on to graduate school (GTC divided by the sum of CLG and GTC). In summary, those who enroll in colleges finish with very different education outcomes.³

2.2 Education Wage Premia

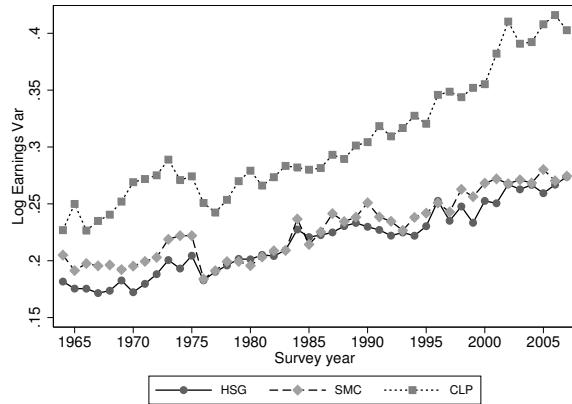
In Figure 1(b), we plot the wage premia for some-college (SMC) and college-graduates (the union of CLG and GTC, which we denote CLP for “college-plus”). The wage premium is defined as the relative difference in the (age-composition-adjusted) average

²The discontinuity at the HSG–SMC margin between 1991 and 1992 reflects a change in the coding convention. Prior to 1992, the CPS only recorded the respondent’s completed years of schooling, with no information of whether a degree was obtained.

³For the female counterpart of 1(a), see Figure 7 in Appendix D. The fractions of women in CLG and GTC categories have steadily increased throughout the observed period. This differential trend suggests gender-specific changes in higher education and in labor markets, which are beyond the scope of this paper.



(a) Residual Earnings Ratios



(b) Log Earnings Variance

Fig. 2: Wage Inequality

White males, ages 26-50. HSG: high school graduate, SMC: some college, CLP: college graduate. 90/50 and 50/10 denote the percentile residual earnings ratios controlling for education and age.

wage between a higher-education group and high school graduates. The two premia show different trends. The some-college group earns a modest premium, about 20 percent in the 1960s and 2000s with a low of 5 percent around 1980. Within the SMC group, dropouts, vocational degree holders, and academic degree holders display similar levels and trends in terms of their premia, which justifies our treating them as a single education category.⁴ The college premium is significantly higher, and increased almost linearly between 1980 and 2005 (from about 40 to 90 percent).

In a simple model in which individuals choose their education level to maximize labor income, the steep rise in the college premium is difficult to reconcile with the constant fraction of college graduates. This has been the subject of much debate, and the earlier literature focused on non-market factors such as the minimum wage (Card and DiNardo, 2002) or supply-demand frameworks (Katz and Murphy, 1992; Autor et al., 2008). We further point out that any proposed explanation would run into another (hitherto neglected) challenge: reconciling the flat some-college premium with the rise in the fraction of the some-college group.

2.3 Residual Wage Dispersion

Recent studies—e.g., Taber (2001), Lemieux (2010), and AKK—focus more on residual wage inequality than wage differences attributable to observed characteristics such

⁴Dropouts and vocational degree holders comprise approximately 65% and 20% of the SMC group throughout 1992–2007, respectively. Both groups have more or less similar premia levels and trends as the group as a whole. Academic degree holders’ premium is about 5 percentage points higher than the SMC average in 1992, and about 10 percentage points higher in 2007. However, not only are they the smallest group, but also the gaps are nowhere near the gaps between SMC and CLP.

as education and demographic variables. In [Taber \(2001\)](#), the rise in residual wage inequality is explained as an increase in the heterogeneity of unobserved skill.⁵ AKK does not explicitly offer an explanation, but shows that the rise is skewed. Specifically, they identify a rising 90-10 ratio (i.e., the 90th percentile residual wage divided by the 10th percentile residual wage) throughout their sample period (more or less same as ours), but further find that most of this is due to a rise in the 90-50 ratio, with the 50-10 ratio all but flat since the mid-1980s. [Figure 2\(a\)](#) shows the 90-50 and 50-10 residual wage ratios in our sample, where residual wage is computed by controlling only for education and age. Despite the fewer number of controls, our findings are similar to AKK’s.

The steep rise of the 90-50 ratio compared to the 50-10 ratio implies that the skewness of the wage distribution must have risen. The dotted line of [Figure 2\(a\)](#) shows that the skewness of the log residual wage distribution rose from -0.04 in 1980 to 0.04 in 2005.

In [Figure 2\(b\)](#), we show the variance of log wages by educational attainment. While the some-college log wage variance closely tracks that of high school graduates, the college-graduate (CLP) log wage variance is larger and increases more rapidly.⁶

We do not provide a structural explanation of the rising residual wage inequality, and our model takes it as given. Instead, we are proposing a mechanism (the sequential option to continue college or not, and the underemployment option) that can turn a symmetric spread in heterogeneity and risk into a more skewed wage distribution, especially among college graduates.

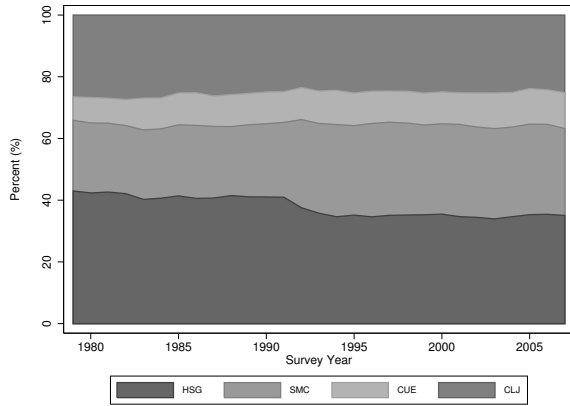
2.4 Underemployment

If a college graduate has difficulty finding a job that requires a college degree or realizes that the available jobs in his field pay poorly, he can instead find a job that has a lower education requirement. Obviously, the opposite is not true. In other words, even if labor markets are segmented by education, the more educated always have access to the less educated’s markets, but not the other way around.

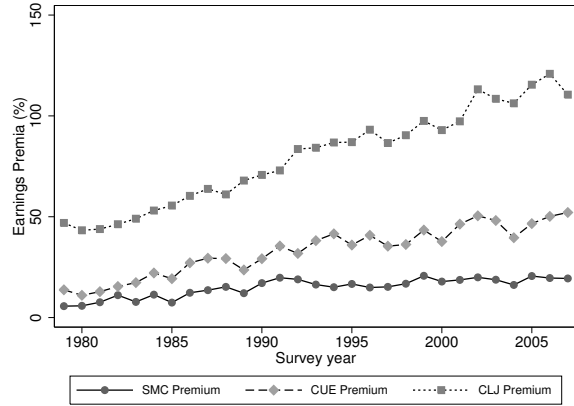
We label the phenomenon of college-educated workers working in jobs with lower education requirements “underemployment.” To determine whether an individual is underemployed, we refer to training and education requirements by detailed occupation, tabulated by the Bureau of Labor Statistics (BLS) as part of their Employment Projection (EP) program. The table is constructed by determining the “typical path”

⁵[Chay and Lee \(2000\)](#) finds that a fraction of the increase in both within- and between-group wage inequality is attributable to the rising importance of unobserved skills.

⁶This is also confirmed by the Theil index decomposition in [Table 12](#) in the [Appendix D](#). Most of inequality is within-group (rather than between-group) inequality. In turn, most of the within-group inequality is explained by the inequality within the college and above groups, and more so in 2005 than in 1980.



(a) Graduation and Underemployment Rates



(b) Underemployed Wage Premia

Fig. 3: Underemployment

White males, ages 26-50. HSG: high school graduate, SMC: some college, CUE: underemployed college graduate, CLJ: college graduate with college job.

of entry for an occupation. Easy examples are physicians and lawyers, for which a professional degree is required. In general, the BLS uses employment shares by occupation and educational attainment from the Census, employer requirements from O*NET, post-secondary program completion data from the National Center for Education Statistics, and qualitative information obtained from educators and employees.⁷

The EP publishes the requirement table biennially at the three-digit occupation level. To address the problem that some occupations may have changed their degree requirement in response to worker supply and not because of the actual job content, we use the 1998 requirement table and apply it to all years. The 1998 table is the earliest available table that requires the least amount of crosswalks to merge CPS and EP occupation codes, thus minimizing the number of occupations for which we do not have degree requirement information.⁸ Also, to minimize the number of instances in which it is unclear whether a worker is underemployed or not, we focus only on college graduates who work in jobs that do not require a four-year degree.⁹

In Figure 3(a), we show educational attainment since 1979, now separating underemployed college graduates (CUE) from those with college jobs (CLJ). The sum of CUE and CLJ are the fraction of college graduates in the economy (also the sum of CLG and GTC in previous figures). To the (small) extent that the fraction of college

⁷See Appendix A.2 for more details.

⁸When different years' tables are used, the resulting overall underemployment trends remain similar.

⁹Associate degree requirement data are noisy at best. Few jobs are categorized as requiring a two-year degree, and even among those that do, they simultaneously state that some work experience can substitute for the degree. Furthermore, many jobs which should require a two-year degree according to EP are filled with high-school-only workers in the CPS.

graduates increased in the economy, most of the increase is actually coming from the underemployed category.

As shown in Figure 3(b), underemployed college graduates do earn more than high school graduates, and increasingly so over time (by 11 percent in 1980 and 46 percent in 2005). However, this underemployment premium is much closer, both in growth and levels, to the some-college premium than to the wage premium commanded by college graduates with college-degree-requiring jobs. This fact partly explains the rising skewness of the residual wage distribution shown in Figure 2(a).

3 Individual Choice Model

We now develop a model of college enrollment and graduation decisions, where the decisions are made before students fully learn their individual returns to college.

Individual i 's return to college is denoted by z_i , which is fixed throughout his lifetime. This may capture the ability and human capital acquired from early childhood through high school, the quality of the college one can be admitted to, the match quality between the college and the student, and so on.

A period in the model is two years. At age 19 (or period $s = 1$), high school graduates start their lives with financial assets a_{1i} and a prior about their z_i . We assume that individuals may borrow up to a natural limit defined by a minimum wage. In our calibration (Section 4.2), it is feasible for even the poorest students to enroll in and graduate from colleges. However, initial wealth is a determinant of college enrollment and graduation, because college is a risky, discrete investment: Poorer students are less likely to take this risk.¹⁰

If they immediately enter the labor market, they draw a wage from a high-school wage distribution. If they enroll in college, they pay a two-year college cost and receive a signal \hat{z}_i about their true z_i at the end of period $s = 1$. If they choose to quit college at the beginning of $s = 2$ (age 21) and enter the labor market, their z_i is revealed and they draw a wage from a some-college wage distribution $G_d(w_i|z_i)$. If they continue and graduate, they pay the college cost for the final two years, observe their true z_i , and draw a wage from a college wage distribution $G_c(w_i|z_i)$ at the beginning of period $s = 3$ (age 23). The underemployment option is built into the distribution functions G_d and G_c , as we explain in Section 3.2.

In our model, college education plays three roles. First, colleges provide students with information on their individual-specific returns to college, consistent with the view that colleges are experience goods. Second, we also assume a human capital

¹⁰Our quantitative analysis shows that the role of wealth in college enrollment and dropout decisions grew more important between 1980 and 2005 (Figure 5 in Section 5.3), because of higher costs of college and larger wage risk.

accumulation aspect: If a student does not attend or complete college, he will be compensated for only a fraction of z_i in the labor market. Third, college education gives students certification or qualification that allows them to look for jobs in the segments of the labor market that require at least some college education.

The first role warrants additional discussion. [Betts \(1996\)](#) and [Arcidiacono et al. \(2010\)](#) find that students learn about their individual labor market returns while in school, especially in their senior year. Recent papers including [Stange \(2012\)](#) and [Stinebrickner and Stinebrickner \(2012\)](#) provide detailed evidence that college students learn about their likely academic performance over time. Another literature, e.g., [Altonji et al. \(2015\)](#), studies how undergraduate students switch majors while in college. If we combine the idea that students learn about which major fits them best with the fact that different majors lead to different wage distributions, they are consistent with our assumption that students learn about their own returns to college.

Another point worth clarifying is that, in our model, we think of those who enroll in college but quit without earning a four-year degree as exercising an exit option. This sounds natural from the point of view of those enrolling in a four-year institution. In the U.S., this is not inaccurate for those starting in two-year institutions either. According to a report from the National Center for Education Statistics for 1994–2009, more than 80 percent of community college freshmen say that their ultimate goal is a bachelor’s or higher degree ([Horn and Skomsvold, 2011](#)). The National Student Clearinghouse Research Center reports that between 2005 and 2008, about 20 percent of two-year college students did transfer prior to their fifth year, and about 60 percent of them obtained a bachelor’s degree within four years of transfer (70 percent if transferring with an associate’s degree). This conditional-on-transfer graduation rate is virtually the same as the graduation rate of students who begin in four-year institutions. In this context, we think of the completion of four-year-degrees as the default outcome, and college quits or not transferring to a four-year institution as the exit option.¹¹

3.1 Heterogeneous Returns and Signals

While an individual’s ex-post return to college is governed by z_i , it is unknown to the individual before labor market entry, and the ex-ante return is governed by the individual-specific priors. Individuals’ returns, z_i , are assumed to be distributed normally in the population:

$$z_i \sim \mathcal{N}\left(\mu_z - \frac{\sigma_z^2}{2}, \sigma_z^2\right). \tag{1}$$

¹¹In addition, as noted in Section 2.2, the different groups within the some-college category display similar levels and trends of premia during the period 1992–2007, when such a breakdown is possible in the data.

The variance parameter captures the degree of return heterogeneity in the population. Since we use z_i to denote the returns in log-point differences, the mean is shifted by one-half the variance to keep the level mean constant when the variance changes.

Individual i is assumed to have a normal prior on z_i at $s = 1$ (age 19): $\mathcal{N}(\mu_{z_{1i}}, \sigma_{z_{1i}}^2)$. We further assume that $\sigma_{z_{1i}}^2$ is identical across individuals—that is, $\sigma_{z_{1i}}^2 = \sigma_{z_1}^2$ for all i .¹² If an individual enrolls in college, he receives a signal \hat{z}_i at the end of $s = 1$:

$$\hat{z}_i = z_i + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2), \quad (2)$$

where ϵ_i is independent of z_i and also i.i.d. across individuals. Those in college use Bayesian updating to form a posterior on z_i which is $\mathcal{N}(\mu_{z_{2i}}, \sigma_{z_2}^2)$, where

$$\mu_{z_{2i}} \equiv \frac{\sigma_\epsilon^2 \mu_{z_{1i}} + \sigma_{z_1}^2 \hat{z}_i}{\sigma_{z_1}^2 + \sigma_\epsilon^2} \quad \text{and} \quad \sigma_{z_2}^2 \equiv \frac{\sigma_{z_1}^2 \sigma_\epsilon^2}{\sigma_{z_1}^2 + \sigma_\epsilon^2}. \quad (3)$$

Because $\sigma_{z_1}^2$ and σ_ϵ^2 are assumed to be the same for everyone, so is $\sigma_{z_2}^2$.

3.2 Constructing Education Specific Wage Distributions

Although we assume that individual wages are drawn from education-specific wage distributions (conditional on z_i), the labor markets are not completely segmented due to the underemployment option. While high school graduates can only access the market for high school graduates, some-college workers can access both the high school and some-college markets, and college graduates can access all markets. This idea is built into the education-specific wage distributions $G_h(w_i)$, $G_d(w_i|z_i)$, and $G_c(w_i|z_i)$. Note that these distributions represent the wage distribution as perceived by an individual with returns to college z_i .

We assume the distribution function G_h , from which high school workers draw a wage is log-normal:

$$\log w_i \sim \mathcal{N}\left(-\frac{\sigma_h^2}{2}, \sigma_h^2\right).$$

This means that wage dispersion among high school graduates is entirely explained by luck or risk.¹³ On the other hand, G_d and G_c depend on z_i , and are constructed from auxiliary distributions F_d and F_c , using the underemployment option as follows.

A some-college worker with returns z_i draws a some-college job wage w_{di} from an auxiliary distribution $F_d(w_{di}|z_i)$, which is log-normal conditional on z_i :

$$\log w_{di} \sim \mathcal{N}\left(z_i - \frac{\sigma_d^2}{2}, \sigma_d^2\right).$$

¹²We also worked out versions of the model in which individuals' prior variances were heterogeneous, but there was virtually no effect on our moments of interest for a wide range of parameterizations.

¹³Cunha and Heckman (2007) finds that most of the increase in wage dispersion among high school graduates comes from the increased variance of the unpredictable component, so ours is not an unreasonable assumption in the context of the rise in overall wage dispersion. We discuss this further in Appendix C.

Since he can also access the market for high school graduates, we assume that he additionally draws a wage w_{hi} from G_h . Then a some-college worker's wage is formed by

$$w_i = m_d \cdot \max \{w_{di}, w_{hi}\},$$

where the factor m_d captures partial returns to his incomplete college education. Hence, the resulting some-college wage distribution G_d is the maximum of two log-normal random variables. It is as if the some-college worker makes two draws, one each from G_h and F_d , and takes the larger of the two. The factor m_d , if less than one, reflects a partial return for the incomplete college education. Figure 8 in Appendix D visualizes how the individual-specific some-college wage distribution G_d is constructed from F_d and G_h , for two different values of z_i .¹⁴ Because the eventual some-college wage distribution for a given z_i is the maximum of two random variables, the $G_d(w_i|z_i)$ distribution is right-skewed, and more so for individuals with higher z_i .

Similarly, we assume that a college graduate with returns z_i can access all three labor markets (high school, some-college, and college graduates), and his wage is formed by

$$w_i = \max \{w_{ci}, m_u w_{di}\} \equiv \max \{w_{ci}, m_u m_d \cdot \max \{w_{di}, w_{hi}\}\},$$

where the college-job wage w_{ci} is drawn from an auxiliary log-normal distribution F_c :

$$\log w_{ci} \sim \mathcal{N} \left(z_i - \frac{\sigma_c^2}{2}, \sigma_c^2 \right). \quad (4)$$

So the resulting college graduate wage distribution G_c is essentially the maximum of three log-normal random variables: A college graduate draws a college-job wage w_{ci} and a some-college wage w_{di} , the latter being the maximum of two independent random variables, and takes the larger of w_{ci} and $m_u w_{di}$. If he takes the latter (a some-college or high-school job), he is underemployed, and $m_u \leq 1$ will capture any wage loss from being underemployed. Figure 9 in Appendix D visualizes the individual-specific college graduate wage distribution G_c for two different values of z_i . Because the eventual college-wage distribution for a given z_i is the maximum of three random variables, $G_c(w_i|z_i)$ is even more right-skewed than the some-college wage distribution $G_d(w_i|z_i)$, and again more so for higher z_i .

The G distributions constructed in this section are the distributions from which workers of a given education level and z_i draw their wages. The resulting education-specific wage distributions in the population will depend on the joint distribution of

¹⁴If a some-college worker chooses the wage from G_h , he is an underemployed some-college worker. However, we ignore calibrating their population shares and premia, since data on such some-college underemployment is too noisy.

z and educational attainment outcomes in the population. Figure 10 in Appendix D shows how educational attainment varies probabilistically with z .

3.3 Individual's Problem

A high-school graduate i at $s = 1$ (age 19) makes decisions based on his state $(a_{1i}, \mu_{z_{1i}})$, his initial assets and the mean of his prior distribution on his own returns to college. While individuals differ in their true z_i , it is unknown and hence their prior on z_i enters the problem instead. Also, one's normal prior has both mean $(\mu_{z_{1i}})$ and variance $(\sigma_{z_{1i}}^2)$, but we assume that the variance of the prior distribution is the same for everybody (Section 3.1). For our formulation, the vector $(a_{1i}, \mu_{z_{1i}})$ suffices as the state variables at age 19. We suppress the index i below unless necessary.

The high-school graduate chooses whether or not to enroll in college. If he does not, he immediately enters the labor market as a high-school graduate, drawing a wage from $G_h(w)$. If he enrolls, he pays the expenses x_1 (potentially subsidized) for the first period (two years) of college, and chooses consumption and next period assets (a_2) . He may borrow up to a natural debt limit defined by a minimum wage, and in our calibration (Section 4.2) college is in the budget set of even the poorest students.¹⁵ Still, initial wealth is a determinant of college enrollment, because college is a risky, discrete investment. We express his value at $s = 1$ recursively:

$$V_1(a_1, \mu_{z_1}) = \max_{\text{work, school}} \left\{ V_h(a_1), \right. \tag{5}$$

$$\left. \max_{a_2} \left[u((1+r)a_1 - a_2 - x_1(1-v(a_1))) + \beta \int V_2(v(a_1), a_2, \mu_{z_2}) dF_1(\mu_{z_2} | \mu_{z_1}) \right] \right\}.$$

The mean of the updated prior distribution in period $s = 2$, μ_{z_2} , evolves according to the Bayesian updating formula (3), and $F_1(\cdot | \mu_{z_1})$ is the c.d.f. of μ_{z_2} conditional on μ_{z_1} , which is determined by the distribution of the signal \hat{z} . More explicitly, F_1 is a normal distribution with mean μ_{z_1} and variance $\sigma_{z_1}^4 / (\sigma_\epsilon^2 + \sigma_{z_1}^2)$. As can be seen, the in-college signal variance σ_ϵ^2 directly affects the enrollment decision.

The terminal value $V_h(a_1)$ is the expected value for a high school graduate who begins working with assets a_1 , which we characterize in the next section. The continuation value $V_2(v(a_1), a_2, \mu_{z_2})$ is the value at the beginning of the next period if he enrolls. We assume that a fraction of tuition costs will be paid for by grants and subsidies, and the rate subsidized declines with the student's initial wealth, a_1 . This

¹⁵Many studies, in particular [Lochner and Monge-Naranjo \(2011\)](#), find little evidence of binding borrowing constraints for college education during our sample period.

grant function $v(a_1)$ is a step function with three possible values,

$$v(a_1) = \begin{cases} v_1 & \text{if } a_1 \leq \bar{a}_1 \\ v_2 & \text{if } \bar{a}_1 < a_1 \leq \bar{a}_2 \\ v_3 & \text{if } \bar{a}_2 < a_1, \end{cases}$$

where $0 < v_3 < v_2 < v_1 < 1$, implying that in fact all students receive a positive amount of grants. Note that $v(\cdot)$ only depends on the student's initial ($s = 1$) level of assets, and is fixed throughout college—if he stays in college. This is why $v(a_1)$ enters the continuation value function V_2 below. Problem (5) induces the optimal policies

$$\chi_E(a_1, \mu_{z_1}) \quad \text{and} \quad a_2^*(a_1, \mu_{z_1}), \quad (6)$$

where $\chi_E = 1$ if the individual decides to enroll in college and 0 otherwise, and $a_2^*(\cdot)$ is the optimal savings function if he chooses to enroll. Once in college, he receives the signal \hat{z} and begins the second period ($s = 2$) with assets $a_2 = a_2^*(a_1, \mu_{z_1})$. Based on the updated prior (μ_{z_2}) and assets (a_2), he decides whether to complete college. If he continues, he pays the expenses for the second period x_2 —with the grants covering a fraction $v(a_1)$ —and chooses next period assets a_3 . His value is

$$V_2(v, a_2, \mu_{z_2}) = \max_{\text{work, school}} \left\{ \int V_d(a_2, z) dF_2(z|\mu_{z_2}), \right. \\ \left. \max_{a_3} \left[u((1+r)a_2 - a_3 - x_2(1-v)) + \beta \int V_c(a_3, z) dF_2(z|\mu_{z_2}) \right] \right\}, \quad (7)$$

where the distribution function $F_2(\cdot|\mu_{z_2})$ is the posterior c.d.f. of an individual's z formed from (μ_{z_1}, \hat{z}) according to the Bayesian updating in (3). The values $V_d(a_2, z)$ and $V_c(a_3, z)$ are the expected values for a some-college worker and a college graduate who joins the labor market with assets a_2 and a_3 , respectively. We assume that his true return z is revealed upon his entry into the labor market. These two value functions are characterized in the next section.

3.4 Terminal Values

The values of entering the labor market at $s = 1$ (as a high school graduate), $s = 2$ (as a some-college worker), and $s = 3$ (as a college graduate) with assets a are, respectively,

$$V_h(a) = \int V(s = 1, a, w) dG_h(w) \quad (8a)$$

$$V_d(a, z) = \int V(s = 2, a, w) dG_d(w|z) \quad (8b)$$

$$V_c(a, z) = \int V(s = 3, a, w) dG_c(w|z). \quad (8c)$$

When an individual starts working for the first time, he draws a wage w from his education-specific distribution: $G_h(w)$, $G_d(w|z)$, or $G_c(w|z)$. With the realization of w , all uncertainties are resolved, and he solves a deterministic consumption-saving problem for the rest of his life. He can borrow and save at a given interest rate subject only to the natural lifetime borrowing constraint $a_{T+1} \geq 0$. Given a constant (two-year) interest rate r and discount factor β , we can derive the continuation utility of a worker who starts working at $s \in \{1, 2, 3\}$ (ages 19, 21 or 23), works until period $R = 24$ (age 65) and lives until period $T = 29$ (age 75):

$$V(s, a, w) = \max \left\{ \sum_{j=s}^T \beta^{j-s} u(c_j) \right\}$$

subject to

$$\sum_{j=s}^T \frac{c_j}{(1+r)^{j-s}} = (1+r)a + \max \{w \cdot e_h(s), \underline{w} \cdot e_r(s)\}, \quad (9)$$

where

$$e_h(s) \equiv \sum_{j=s}^R \frac{y_h(j)}{(1+r)^{j-s}} \quad \text{and} \quad e_r(s) \equiv \sum_{j=s}^R \frac{1}{(1+r)^{j-s}} \quad (10)$$

and $y_h(j)$, $j = 1, \dots, 24$, is the average age-earnings profile of a high school graduate. We normalize the first year average earnings of a high-school worker to $y_h(1) = 1$. We assume a minimum wage \underline{w} that applies equally to all workers, putting a lower bound on their lifetime income. The functions $e_h(s)$ and $e_r(s)$ transform the hourly wage w into a present-discounted sum of lifetime earnings, evaluated at period s . Assuming iso-elastic utility $u(c) = c^{1-\gamma}/(1-\gamma)$, we can solve for $V(s, a, w)$ in closed form.

3.5 Discussion on Model Assumptions

No going back to school We are assuming that entering the labor market is an absorbing state. If re-enrollment were allowed, high-school and some-college workers with low wage realizations or revelations of higher-than-expected z 's would return to college. However, for more than 95 percent of the population in the NLSY, their educational attainment is finalized by age 25 (although there are slightly more delayed college enrollment and re-enrollment in NLSY97 than in NLSY79). In addition, if we make an alternative assumption that workers learn their true z slowly over time, rather than instantly upon entering the labor market, this “no re-enrollment” constraint will not be as binding. For these reasons, we do not expect that abstracting from this dimension would have a large quantitative impact on ex-post outcomes.

Once-and-for-all wage shock realization With the assumption that the wage shock realization is once-and-for-all upon labor market entry, we may be exaggerating

the risk in the actual earnings process: Fixing the same ex-ante probabilistic wage distribution, if wages move stochastically over time, workers can partly self-insure through saving and borrowing. However, the magnitude of this overestimate is likely small, since it is known that more of the inequality in lifetime earnings is accounted for by differences in workers’ initial conditions (i.e., earnings in their early to mid twenties) than by differences in idiosyncratic shocks over their working life (Keane and Wolpin, 1997; Huggett et al., 2011; Guvenen et al., 2015).

However, we need more discussion on our assumption that underemployment is a permanent state. In the NLSY, underemployment is a temporary yet persistent phenomenon. For those in the sample who are underemployed for at least one period, almost half their working life is spent in underemployment on average, broken into multiple spells, with each spell on average lasting six years. Furthermore, Clark et al. (2015) finds an enduring negative wage effect from underemployment: Those who exit underemployment after a spell of at least four years experience a wage penalty of 2–5 percent, controlling for observables including occupation. In sum, for college-graduates experiencing underemployment, it is a persistent state, both in duration and in impact.

4 Calibration

We calibrate the model to the 1980 U.S. as a benchmark, and then separately to 2005 as a comparison. We consider the 1980 and 2005 cross-sections as two different steady states. They are sufficiently removed from each other, and no cohort in our data straddle them. (We are only considering 26–50 year olds in each cross-section.) The steady-state assumption implies that all trends are attributed to time effects.¹⁶

4.1 Population Distribution

We first make assumptions on the population distribution over which we aggregate individual choices and outcomes. At the beginning of $s = 1$, each individual i in our model is fully described by the trivariate vector

$$(\log a_{1i}, \mu_{z_{1i}}, z_i) \sim F_0,$$

where $\log a_{1i}$ is log initial wealth, $\mu_{z_{1i}}$ is the mean of individual i ’s prior on his own return to college, z_i is his true return to college (as yet unobserved), and F_0 is the joint population distribution. We assume that F_0 is trivariate-normal. Since we already assumed that the marginal distribution of z_i is normal in (1), this adds seven additional parameters: the population means and variances of $\log a_{1i}$ and $\mu_{z_{1i}}$, and the three

¹⁶We also worked out an alternative calibration exercise, in which we targeted time-averaged moments for 1978–82 and, separately, 2003–07. The quantitative results remained more or less intact.

pairwise correlation coefficients ρ_{az_1} (between $\log a_{1i}$ and μ_{z_1i}), ρ_{az} (between $\log a_{1i}$ and z_i), and ρ_{zz_1} (between z_i and μ_{z_1i}). The marginal distributions of initial assets a_{1i} and individuals' prior means μ_{z_1i} in the population are:

$$\log a_{1i} \sim \mathcal{N}\left(\mu_a - \frac{\sigma_a^2}{2}, \sigma_a^2\right) \quad \text{and} \quad \mu_{z_1i} \sim \mathcal{N}\left(b_z + \mu_z - \frac{\sigma_z^2}{2}, \sigma_{\mu_{z_1}}^2\right),$$

where b_z captures potential biases in beliefs compared to the actual population mean of z_i (which is μ_z). If $b_z > 0$, students are optimistic on average; otherwise they are pessimistic. An alternative interpretation of positive and negative b_z is consumption value and non-pecuniary cost of attending college, respectively. The population variance $\sigma_{\mu_{z_1}}^2$, which describes the dispersion of individual μ_{z_1i} 's in the population, is not to be confused with the variance of the individual priors, $\sigma_{z_1}^2$, which describes how certain individuals are of their beliefs.

Further clarification is warranted. Individual i makes college enrollment decisions based on his normal prior distribution on his z_i . The mean and variance of his normal prior are μ_{z_1i} and $\sigma_{z_1}^2$. The prior mean μ_{z_1i} has a non-degenerate distribution across individuals, and is assumed to be correlated with their true returns z_i according to the correlation coefficient ρ_{zz_1} . Above, we introduced b_z to allow the distributions of z_i and μ_{z_1i} to have different cross-sectional means. We include b_z not because it is necessary for calibration, but because we are interested in measuring the average optimism or pessimism in the data and how it changes between 1980 and 2005. As for the cross-sectional variances of z_i and μ_{z_1i} , respectively σ_z^2 and $\sigma_{\mu_{z_1}}^2$, we impose the assumption that they are the same: $\sigma_z^2 = \sigma_{\mu_{z_1}}^2$.¹⁷

While students do not directly observe their own returns z_i , we assume they know the population distribution F_0 . Then, since they observe their own initial assets a_{1i} , they can utilize the fact that their initial wealth and true z_i are correlated when they form their priors. To capture this, we set the variance of individuals' priors $\sigma_{z_1}^2$ (which is assumed to be the same for everyone) to $(1 - \rho_{az_1}^2)\sigma_z^2$: The uncertainty over one's beliefs is only so great as the population variance of returns z_i conditional on one's assets. That is, the degree of uncertainty of one's belief is less than the actual dispersion of the true returns z_i in the population.¹⁸

In summary, with the two restrictions on the variances of the three distributions, we forgo two degrees of freedom in our calibration. We do this mainly because there is no usable survey data (especially for 1980) that is informative about the dispersion of the prior mean in the population ($\sigma_{\mu_{z_1}}^2$) or the variance of an individual's prior ($\sigma_{z_1}^2$).

¹⁷We have tried alternative quantitative exercises relaxing this assumption, but the main results remain the same, except when $\sigma_{\mu_{z_1}}^2$ is extremely large or small.

¹⁸In reality, the precision of a high school student's information is likely even larger, since he has access to more information than only his (family's) financial wealth. For example, test scores could reveal more information about a student's returns to college.

Parameter	Value		Description
(β, γ, r)	$(0.96^2, 2, (1.04)^2 - 1)$		Standard values
	1980	2005	
$w_{hi} \sim \log \mathcal{N}\left(-\frac{\sigma_h^2}{2}, \sigma_h^2\right)$	$\sigma_h^2 = 0.201$	0.259	Average HSG starting salary μ_h normalized to 1
$a_{1i} \sim \log \mathcal{N}\left(\mu_a - \frac{\sigma_a^2}{2}, \sigma_a^2\right)$	$\mu_a = 0.937$	1.035	Gale and Scholz (1994)
$\sigma_z^2 = \sigma_{\mu_{z_1}}^2 = \sigma_{z_1}^2 / (1 - \rho_{az_1}^2)$	$\sigma_a^2 = 0.804$		NLSY97
x_1	See Section 4.3		Only σ_z^2 calibrated within model
x_2	0.234	0.323	First 2-year college cost
	0.630	0.873	Last 2-year college cost
(\bar{a}_1, \bar{a}_2)	$[f_a(\bar{a}_1), f_a(\bar{a}_2)]$ $= (0.20, 0.55)$		$f_a(a) = \Phi\left(\frac{\log a - (\mu_a - \sigma_a^2/2)}{\sigma_a}\right)$, Φ is standard normal c.d.f.
(v_1, v_2, v_3)	$(0.326, 0.077, 0.017)$		Fraction of college costs covered by public grants
\underline{w}	0.643	0.448	2-year PDV minimum wage

Table 1: Parameters Taken Directly from the Data

4.2 Parameters Taken Directly from the Data

For each steady state, we need to calibrate 26 parameters, of which 15 can be directly pinned down by their data counterparts or the literature. All dollar values are deflated to 2000 dollars using the chain-weighted (implicit) price deflator for personal consumption expenditures published by the BEA.

The discount factor and relative risk aversion coefficient (β, γ) are fixed to standard values of $(0.96^2, 2)$, and the interest rate to $r = 1.04^2 - 1$. The discount factor and interest rate are compounded since one period is two years.

In the model, mean high school graduate wage at $s = 1$ are normalized to 1. Since a period in the model is two years, we use average wages of 19–20 year-olds to normalize all wages in the data. Specifically, the average two-year present discounted value wage for high school graduates at age 19 is

$$\bar{w}_h = \bar{w}_{h_{19}} + \bar{w}_{h_{20}}(1+r)^{-\frac{1}{2}}, \quad (11)$$

where \bar{w}_{h_a} is the mean wage for a -year-old high school graduates. This value is \$19.13 in 1980 and \$19.23 in 2005, both in 2000 dollars. Since these are hourly rates and we assume that all individuals work full time, we not only normalize annual earnings by \bar{w}_h but further divide them by 1,400 hours (35 hours per week times 40 weeks). The log wage variance for high school graduates, σ_h^2 , is 0.201 and 0.259 in 1980 and 2005.

Initial assets in the model mainly capture heterogeneity in economic support from students' parents. We assume a natural borrowing constraint, and front-load all pos-

Type of Institution	1980	2005
Two-year public	789	1,965
Four-year public	1,641	4,925
Four-year private	7,170	19,046

Table 2: College Costs (in 2000 dollars)

sible transfers received from parents into initial assets.

To obtain the mean of log assets μ_a , we refer to [Gale and Scholz \(1994\)](#), which reports that average net worth in the 1986 Survey of Consumer Finances was \$144,393 per household in 1985 dollars, and 63 percent of total net worth comes from intergenerational transfers. Assuming a two-person household with two children, this amounts to \$68,382 of asset transfers per child in 2000 dollars. For 2005, we multiply this value by the ratio of the average present discounted value of lifetime earnings for the entire population in 1980 and 2005, leading to \$75,780. To compute σ_a^2 , we compare the mean and median annual inter-vivos transfers among young adults aged 16-22. These values were \$1,227 and \$486, respectively, in the NLSY97. This implies a positive skewness partially justifying our choice of a log-normal distribution for initial assets, with $\sigma_a^2 = 2 \log(1227/486)$, from the formula for mean and median of normal distributions.

We refer to *Trends in College Pricing*, published annually by the College Board, to construct the college costs parameters (x_1, x_2) . We exclude room and board and only include tuition and fees, since all individuals would incur living costs regardless of college attendance. We make the first two years of college cheaper than the latter two years, by setting x_1 as the average cost of attending a two-year public, four-year public or four-year private institution, and x_2 as the average cost of attending a four-year public or private institution.¹⁹ These costs are shown in [Table 2](#). Clearly, the cost of college has been rising over time: from 1980 to 2005, average annual costs increased from \$3,200 to \$8,645 for the first two years of college, and from \$4,406 to \$11,986 for the latter two years, all in 2000 dollars.

Grants are modeled as a decreasing function of students' initial wealth ([Section 3.3](#)). In particular, we assume that the fraction of college costs subsidized is a declining three-step function of initial wealth, a_{1i} . According to the 2000 *Guide to U.S. Department of Education Programs*, students with family income below \$30,000 received an average of \$2,820 in public grants; above \$30,000 but below \$80,000 received an average of \$668; and above \$80,000 an average of \$143. These family income thresholds correspond to approximately the 20th and 55th percentiles of the 2000 family income distribution.

¹⁹ This makes enrolling financially easier than graduating, and is intended to capture the fact that four-year institutions are more expensive than community colleges. While these costs clearly differ across individuals and institutions, there is evidence that they do not vary much across family income groups, see for example [Johnson \(2013\)](#) and [Abbott et al. \(2014\)](#).

Moment	1980	2005
Enrollment rate	0.577	0.646
Discontinuation rate	0.393	0.451
Underemployment rate	0.237	0.326
College wage premium	0.361	0.934
Some-college premium	0.057	0.214
Underemployment premium	0.110	0.467
High school log wage variance (from Table 1)	0.201	0.259
College graduate log wage variance	0.272	0.405
Some-college log wage variance	0.201	0.262
Log corr. family income and own earnings	0.178	[0.185]
Enrollment by family income quartile*	0.409	0.421
College attainment by earnings quartile**	0.292	0.402

*Difference in enrollment rates of the first and fourth quartiles by parental income.

**Difference in fraction of college graduates of the first and fourth quartiles by own earnings.

Table 3: Target Moments

All moments are for white males, ages 26–50 in 1980 and 2005 CPS, except the correlation among family income, own earnings, and enrollment rates. These correlations are from the NLSY79/97, except for the 2005 correlation between family income and own earnings for which data does not yet exist—we instead show the corresponding simulated moment in brackets.

Moreover, *Trends in Student Financing of Undergraduate Education*, published by the National Center for Education Statistics, does not show any systematic changes across family income quartiles in the fraction of college costs covered by federal or state grants between 1995 and 2008. Hence, we fix the grant function for both 1980 and 2005 in our model to match the 2000 numbers.²⁰ Since we assume that initial wealth is log normal, we choose the thresholds \bar{a}_1 and \bar{a}_2 such that

$$\Phi\left(\frac{\log \bar{a}_1 - (\mu_a - \sigma_a^2/2)}{\sigma_a}\right) = 0.20 \quad \text{and} \quad \Phi\left(\frac{\log \bar{a}_2 - (\mu_a - \sigma_a^2/2)}{\sigma_a}\right) = 0.55,$$

separately for the μ_a and σ_a^2 in 1980 and 2005, where Φ is the standard normal c.d.f. For the three groups divided by the thresholds, the parameters v_1 , v_2 , and v_3 are fixed to the average amount of grants received within each group as a fraction of the average costs in 2005 for the first two years of college: i.e., $(v_1, v_2, v_3) = (2820, 668, 143)/8645$.

The minimum wages in 1980 and 2005 were, respectively, \$6.27 and \$4.39, corresponding to two-year present discounted values of \$12.29 and \$8.62. Our assumption of the natural borrowing limit and minimum wages implies that college attendance and graduation are in even the poorest student’s budget set.

²⁰Neither data dates as far back as 1980—the U.S. Department of Education was only created in 1980.

4.3 Parameters That Are Jointly Calibrated

We determine the remaining 11 parameters through minimum-distance calibration:

$$\hat{\Theta} = \arg \min_{\Theta} [M(\Theta) - M_d] [M(\Theta) - M_d]' \quad (12)$$

where Θ is the vector of parameters, $M(\Theta)$ the simulated moments from the model given Θ , and M_d the empirical moments reported in Table 3, most of which were shown in various figures of Section 2. For this joint calibration, we have as many parameters as there are target moments. For numerical details, refer to Appendix B.

We have a fully parameterized model, and the identification of heterogeneity, beliefs, and risk is through distributional assumptions. Still, the key idea in the joint calibration is that the heterogeneous returns (z) are closely correlated with realized wages (differentially for some-college, underemployed, and non-underemployed college graduates),²¹ and the beliefs μ_{z_1} are correlated with the decision to enroll in and graduate from college.

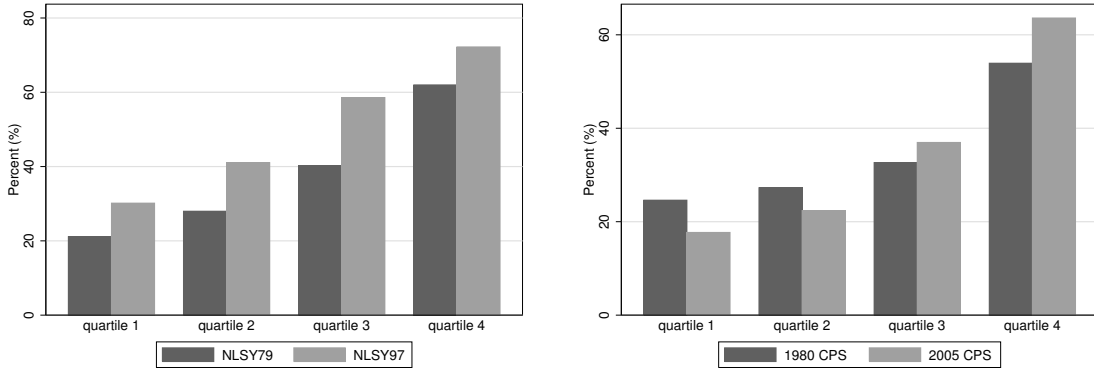
In constructing the target moments, we first compute the relevant moments for each age (from 26 to 50), and then equally weight the age-specific moments. Because our model assumes a uniform age distribution and deterministic age-wage profiles, it is straightforward to construct the model counterparts.

We now explain in more detail how we discipline the correlation parameters of the initial distribution of wealth, prior mean, and true return, F_0 . Of the 12 parameters completely characterizing the trivariate-normal F_0 , nine were set directly by available data in Section 4.2. The remaining three correlation parameters are ρ_{az} , ρ_{zz_1} , and ρ_{az_1} .

Initial assets in the model represent family income in the data. True return z in the model is the key determinant of lifetime earnings, and μ_{z_1} , an individual's prior mean, governs whether he decides to enroll and graduate. Hence we target (i) the correlation between family income and present discounted value of own lifetime earnings in the NLSY79, which is informative about ρ_{az} ; (ii) the difference in enrollment rates of the first and fourth family income quartiles in the NLSY79/97, which is informative about ρ_{az_1} ; and (iii) the difference in the fraction of college graduates of the first and fourth (own) earnings quartiles in the CPS, which informs us on ρ_{zz_1} .²² Despite (ii), it should be noted that ρ_{az_1} is not the main driver of the correlation between family income and enrollment in the model. Initial assets are an important determinant of enrollment because college is a risky discrete investment: All else equal (including μ_{z_1}), richer students are more likely to enroll in college.

²¹Our calibration focuses only on the first two moments of the education-specific wage distributions. In Section C.1, we non-parametrically compare the entire wage distributions obtained from the model with their empirical counterparts.

²²Family background information is not in the CPS, and lifetime earnings realizations are not yet available for the NLSY97 cohort. We keep all multi-racial white individuals in the NLSY in addition to whites.



(a) Enrollment by Family Income Quartiles (b) College Graduation by Earnings Quartiles

Fig. 4: Enrollment and Graduation by Quartiles

Left: white males, NLSY79 and 97. Right: white males, ages 26–50, 1980 and 2005 CPS. Earnings are first controlled for age then ordered into quartiles.

Family (i.e., parental) income in the NLSY79 is taken as the mean of the reported values when the student is 16 and 17 years old. To compute the present discounted value of lifetime earnings, we impute missing wage observations by linearly interpolating log wages in adjacent years, and then use an annual discount rate of four percent. We cannot do the same for the NLSY97, because the surveyed youths were 12–16 years old at the end of 1996 and we lack their wage data for later in life. Consequently, for our calibration of 2005, we fix ρ_{az} to its estimate from the 1980 calibration. The simulated correlation in 2005 between family income and own lifetime earnings (in square brackets in Table 3) is slightly higher than its value in the 1980 data.²³

Lastly, we go back to the CPS to compute what fraction within each earnings quartile has a four-year college degree. This is the converse of the college premium construction that first groups people by education. We form earnings quartiles after first subtracting age-specific wage means for each year. Figure 4(b) shows that a larger part of earnings inequality is attributable to education levels in 2005 than in 1980.

5 Results

We first report the jointly calibrated parameter values in Table 4, for the two separate steady states of 1980 and 2005. Because we have the same number of parameters and target moments in the joint calibration, our model replicates all the target moments

²³As described in Appendix A.2, we follow Lochner and Monge-Naranjo (2011) for the definition of family income and enrollment. That paper stresses that the difference in enrollment rates by family income quantiles has become larger in recent years. While there is some evidence of this in our Figure 4(a), the change from 1980 to 2005 is not large. This is because we are only looking at white males, who are more likely to fall in the higher family income quantiles of the entire NLSY sample.

Parameters	1980	2005
b_z (“optimism”)	0.242	0.290
s_ϵ (signal-noise ratio, $1 + \sigma_{z_1}^2/\sigma_\epsilon^2$)	1.511	1.354
σ_z^2 (pop. variance of z)	0.188	0.308
μ_z (pop. mean college returns)	-0.055	0.066
m_d (partial return, some-college)	0.870	0.898
m_u (partial return, underemp.)	0.577	0.617
σ_c^2 (var. of log residual college wage; F_c)	0.231	0.395
σ_d^2 (var. of log residual some-college wage; F_d)	0.383	0.460
$\rho_{az} \equiv \rho(a, z)$	0.448	[0.448]
$\rho_{az_1} \equiv \rho(a, \mu_{z_1})$	0.402	0.453
$\rho_{zz_1} \equiv \rho(z, \mu_{z_1})$	0.644	0.764

Table 4: Calibrated Parameters, 1980 and 2005

All parameters are calibrated to the moments in Table 3, except for $\rho(a, z)$ for 2005 which we fix to its 1980 value as we lack the corresponding target moment in 2005.

of Table 3. While looking at the parameter values themselves is not as informative as a systematic decomposition exercise, we first describe which parameter values changed more significantly between 1980 and 2005.

Both the mean and the variance of log college returns in the population (μ_z, σ_z^2) are larger in 2005, and the residual wage variance for jobs requiring a college degree, σ_c^2 , almost doubled between the two periods, consistent with the observed rises in college premia and wage dispersion among college graduates.

We also see a slight increase in the bias or optimism parameter b_z , and a small decline in the in-college signal-noise ratio, $s_\epsilon = 1 + \sigma_{z_1}^2/\sigma_\epsilon^2$, implying that the in-college signal in 1980 had marginally more information content than in 2005. These findings call for more thinking. In a simpler framework, one would need more pessimism (i.e., a large drop in b_z) to reconcile the modest rise in enrollment with the large rise in the college premium. In our model, however, enrollment does not rise despite students being more optimistic on average. Similarly, to have an even higher college dropout rate, one would require more precise in-college signals. In our model, relative to 1980, students in 2005 expect the mean college return to have risen even more than it actually did, and do not learn more about themselves in college, but nonetheless do not enroll nor graduate as much. The correlation between one’s true return and prior mean, ρ_{zz_1} , increases from 0.64 in 1980 to 0.76 in 2005, implying that students in 2005 have a better idea of their true returns at the enrollment stage. This is consistent with the findings that more recent cohorts base their education decisions more on expected returns than other factors, such as distance to school (Hoxby, 2009).

Two correlation parameters were calibrated to match the difference in the enrollment rates of the first and fourth family income quartiles (ρ_{az_1}), and the difference

Moment	1980	(1) μ_z	(2) σ_z^2	(3) σ_c^2	2005
Enrollment rate	0.577	0.952	0.424	0.536	0.646
Discontinuation rate	0.393	0.215	0.517	0.598	0.451
Underemployment rate	0.237	0.193	0.210	0.326	0.326
College wage premium	0.361	<u>0.934</u>	0.963	0.557	<u>0.934</u>
Some-college premium	0.057	0.162	0.075	0.099	0.214
Underemployment premium	0.110	0.637	0.573	0.063	0.467
College graduate log wage variance	0.272	0.317	<u>0.405</u>	<u>0.405</u>	<u>0.405</u>
Some-college log wage variance	0.201	0.223	0.223	0.209	0.262
Log corr. Family income and own earnings	0.178	0.286	0.242	0.141	[0.185]
Enrollment by family income quartile*	0.409	0.097	0.412	0.417	0.421
College attainment by earnings quartile**	0.292	0.417	0.380	0.204	0.402

*Difference in enrollment rates of first and fourth family income quartiles.

**Difference in fraction of college graduates of first and fourth quartiles by own earnings.

Table 5: Comparative Statics

Starting from the 1980 economy, we only increase the population mean μ_z to match the 2005 college premium in column (1), and the population variance σ_z^2 and college wage uncertainty σ_c^2 to match the 2005 college wage variance in (2) and (3) respectively. Note that σ_c^2 , matching 2005 college wage variance also leads to matching the 2005 underemployment rate.

in the fraction of college graduates in the first and fourth quartiles by own earnings (ρ_{zz_1}). Figures 11 and 12 in Appendix D show these statistics for all quartiles, so that one can assess the model performance in terms of untargeted moments. Figure 11 shows how larger the enrollment rates are in the second through fourth family income quartiles than in the first quartile, in the NLSY and our model. We plot the inter-quartile differences instead of levels, because the unconditional enrollment rates in the 1980 and 2005 CPS (and hence also in the model) differ from those in NLSY79 and 97, especially for 1980. In Figure 11 we see that the model overstates the enrollment rates of the second and third family income quartiles relative to that of the first quartile for 1980. The model is closer to the data for 2005. Figure 12 shows that the model fits the college graduation rates by own earnings quartiles well in both periods.

5.1 Comparative Statics

In Table 5, we start with the calibrated parameters from 1980 and then change key parameters one at a time. We pair each chosen parameter with the empirical moment that is most directly related to it. We then choose a new value for the parameter, holding all else fixed, so that the new model outcome exactly matches the paired empirical target moment in 2005. (Naturally, the model will not match any other 2005 moments.) We reproduce the empirical moments for 1980 and 2005 from Table 3 in the far left and right columns of Table 5.

We begin with the population mean of returns, μ_z , in column (1), which is raised from its 1980 value of -0.055 to 0.497 to exactly match the observed college premium in 2005 (0.934, underlined). This represents a massive across-the-board increase in returns to college. As a consequence, the enrollment rate reaches an unrealistically high 95 percent, and among those who enroll, only 22 percent quit college before earning a four-year degree. Since we assume that some-college workers and underemployed college graduates also enjoy some partial returns, both the some-college and underemployment premia increase as well.

This exercise shows that a model that relies only on changes in the population mean returns μ_z to explain the rising college premium between 1980 and 2005, even with the some-college category and underemployment, would be amiss on educational attainment trends unless significant pessimism or negative preference shocks are invoked. It would also predict that few students would become underemployed, and even then, earn a large premium. Furthermore, with the across-the-board increase in returns to college, family income is less of a factor for enrollment. This contrasts with the data: The difference in enrollment rates between the top and bottom family income quartiles is larger in 2005 than in 1980. In sum, this exercise suggests that a model abstracting from changes in second (or higher) moments cannot resolve the tension between rising college premia and stagnant college enrollment and graduation rates.

In columns (2) and (3) of Table 5, we respectively increase σ_z^2 (from 0.188 to 0.523) and σ_c^2 (from 0.231 to 0.512) to match the college wage variance in the 2005 data (0.405, underlined), fixing all other parameters at their 1980 values. In the model, college graduates' wages have two parts: the individual-specific return z , about which students have priors and also learn while in college, and pure luck (w_c for college jobs, and w_d or w_h if underemployed). If we only increase σ_z^2 , students' decision rules are not affected—it only changes how individuals are distributed.²⁴ An increase in σ_c^2 on the other hand directly affects students' decision rules, since it alters the riskiness of the labor market for college graduates. In either case, it is a mean-preserving spread, because of the way we parameterized the distributions in equations (1) and (4).

We highlight that a (mean-preserving) increase in either variance results in a large increase in the college premium, from 0.361 to 0.963 (σ_z^2) or to 0.557 (σ_c^2). In both cases, enrollment rates drop and college dropout rates rise. However, the underemployment rates for college graduates move in opposite directions in the two cases. In fact, when σ_c^2 is raised in column (3) to match the college-wage variance in 2005, by sheer chance, the simulated underemployment rate also coincides with the 2005 data (boldfaced). This differential effect on underemployment rates helps us break down the wage dispersion

²⁴When increasing σ_z^2 , we only change the population variance of z but keep constant the variance of individual priors ($\sigma_{z_1}^2$) and the population variance of the mean of individual priors ($\sigma_{\mu_{z_1}}^2$).

among college graduates into heterogeneity in returns (controlled by σ_z^2) versus residual uncertainty (controlled σ_c^2).

College graduates' wage premium increases because of the optionality embedded in college enrollment and graduation. With the options, a mean-preserving spread of the returns-to-college and residual college wage distributions does not shift the realized wage distribution of college graduates symmetrically. Those with higher returns to college and good realizations of college labor market wages pull the observed college wage premium higher. However, those who learn that their own returns to college are lower than expected are more likely to either discontinue college (and hence fall out of the calculation for the college premium) or become underemployed upon graduation (which truncates the left tail of the college wage distribution). As a result, we have a higher college premium ex post. This mechanism cautions against applying standard mean-variance trade-off calculations to college enrollment and graduation decisions.

The decrease in the enrollment and graduation rates in response to the higher σ_z^2 is merely a result of the larger heterogeneity shifting masses across the thresholds for enrollment and graduation decisions in the space of wealth and (updated) prior means, since individual decision rules are not affected.²⁵ The same applies to the lower underemployment rate and higher underemployment premium.

On the other hand, the higher σ_c^2 increases the probability of underemployment, and in turn, raises the thresholds for enrollment and graduation decisions—i.e., a lower enrollment rate and a higher college dropout rate. The thresholds move in a way that makes wealth a more important determinant of the enrollment and graduation decisions.²⁶ With the higher σ_c^2 , first there is the direct effect that it is more likely to draw wages below a given threshold (which will result in underemployment). This direct effect is reinforced by the weaker selection on individual returns—since students' wealth plays a more important role in enrollment and graduation decisions.²⁷ As a result, we have a higher underemployment rate and a lower underemployment premium. Consistent with this last explanation, the enrollment difference between the top and bottom family income quartiles is slightly larger (i.e., wealth is more important in the enrollment decision) and the correlation between realized labor income and college graduation is weaker (i.e., individual z is less important in the graduation decision).

Through the lens of our model, the changes between 1980 and 2005 in education outcomes and education-specific labor market outcomes can be accounted for in the following way. The modest increase in college enrollment can be mostly attributed to

²⁵See Figure 5 in Section 5.3 for a visual representation of the thresholds.

²⁶This shift in the enrollment and college dropout thresholds in response to a higher σ_c^2 is qualitatively similar to the change between 1980 and 2005 shown in Figure 5.

²⁷Consistent with the weaker selection on z at the graduation decision stage, the some-college premium is actually higher with the higher σ_c^2 .

the higher population mean return to college. While it also raises education premia, a larger part of the higher college premium in 2005 comes from a higher variance of the college return distribution, which also has the effect of lowering college enrollment. These two together could potentially lead to a big increase in the underemployment premium as well, but they are counterbalanced by the higher residual college wage variance, which instead leads to a higher underemployment rate among college graduates. The intermediate steps of college education—i.e., discontinuing college education or becoming underemployed—are important not only because they are conspicuous features of the data that aid identification, but also because they generate downside risk that does not bring down the college premia with it.

5.2 Decomposition of the Change in Target Moments

In this section we conduct the converse exercise as an attempt to isolate the contribution of each model element to the changes in education and labor market outcomes between 1980 and 2005. To be more specific, we start from the 2005 economy, and replace a chosen parameter’s 2005 value with its calibrated value in 1980. All other parameters are left at their 2005 values. The resulting model-generated moments, now deviating from the perfectly matched 2005 data in all dimensions, are reported in Table 6.

Column (1) shows the change in moments when only the population mean return to college μ_z in 2005 (0.066) is replaced with its 1980 calibrated value (-0.055). This is exactly the reverse of what was in column (1) of Table 5. Since the 1980 μ_z is smaller, both the college wage premium and enrollment rates decline from the 2005 levels. However, all the education wage premia are still closer to their 2005 values than their 1980 values. We interpret this as evidence that the main role of the increase in μ_z was to encourage more students to enroll, rather than to increase the college premium.

Replacing either variance parameter, σ_z^2 or σ_c^2 , with its lower 1980 value, shown in columns (2) and (3) of Table 6, reduces the college premium—actually by more than in column (1)—but raises the enrollment rate.²⁸ Put differently, without the increase in either variance between 1980 and 2005, the resulting enrollment rate already overshoots its level in the 2005 data (0.718 or 0.659 vs. 0.646), although the corresponding college premium is not as high as the 2005 data (0.739 or 0.791 vs. 0.934). This is another way of highlighting the tension between the higher college premium and the stagnant enrollment/graduation rate in a model that relies only on the population mean—column (1) of Table 5. However, the two variances have different implications on underemployment rates and premia. Columns (2) and (3) together show that, from 1980 to 2005, the larger σ_z^2 is more responsible for the higher underemployment

²⁸When replacing σ_z^2 with its 1980 estimate we leave $\sigma_{z_1}^2$ and $\sigma_{\mu_{z_1}}^2$ at their 2005 values.

Moment	2005	(1) μ_z	(2) σ_z^2	(3) σ_c^2	(4) σ_z^2, σ_c^2
Enrollment rate	0.646	0.542	0.718	0.659	0.730
Discontinuation rate	0.451	0.459	0.447	0.349	0.359
Underemployment rate	0.326	0.336	0.330	0.277	0.279
College wage premium	0.934	0.809	0.739	0.791	0.637
Some-college premium	0.214	0.191	0.219	0.186	0.198
Underemployment premium	0.467	0.367	0.346	0.517	0.418
College graduate log wage variance	0.405	0.391	0.353	0.329	0.276
Some-college log wage variance	0.262	0.257	0.253	0.256	0.250
Log corr. family income and own earnings	[0.185]	0.163	0.164	0.206	0.186
Enrollment by family income quartile*	0.421	0.455	0.373	0.414	0.360
College attainment by earnings quartile**	0.402	0.335	0.346	0.442	0.388

*Difference in enrollment rates of the first and fourth family income quartiles.

**Difference in fraction of college graduates of the first and fourth quartiles by own earnings.

Table 6: Decomposing the Change between 1980 and 2005

Holding all other parameters at their 2005 values, we replace a chosen parameter with its value in the 1980 calibration. The chosen parameters are, respectively: the population mean of z in column (1); the population variance of z , σ_z^2 in (2); log residual college wage variance, σ_c^2 , in (3); and both σ_z^2 and σ_c^2 in (4).

premium and the larger σ_c^2 for the higher underemployment rate. These results are a mirror image of the comparative statics in columns (2) and (3) of Table 5.

In column (4), we replace both σ_z^2 and σ_c^2 with their respective 1980 values. The college premium is then only 0.637. From this number, we calculate that, of the 57 percentage point increase in college premium in the data, 30 (or more than half of the increase) come from the increased dispersion of heterogeneity and risk. By comparison, column (1) shows that only 12 percentage points of the college premium increase (or one-fifth of the increase) come from the increased population mean of the return-to-college distribution.

5.3 Option to Discontinue College Education

In our model, those who enroll in college defer college graduation decisions until the arrival of new information while in college, the signal at $s = 2$. The additional information on individual returns after enrollment implies that the graduation margin is more selective than the enrollment margin.

In Figure 5, we show the enrollment decision and the college dropout probabilities in our calibrated economies of 1980 and 2005. The horizontal and vertical axes show the individual states a_1 (normalized by the average high-school worker's two-year earnings from age 19 to 20) and μ_{z_1} , respectively. The shaded areas represent those who choose to enroll in college. The different shades correspond to different probabilities with which they will discontinue college education before earning a four-year degree. Marginal

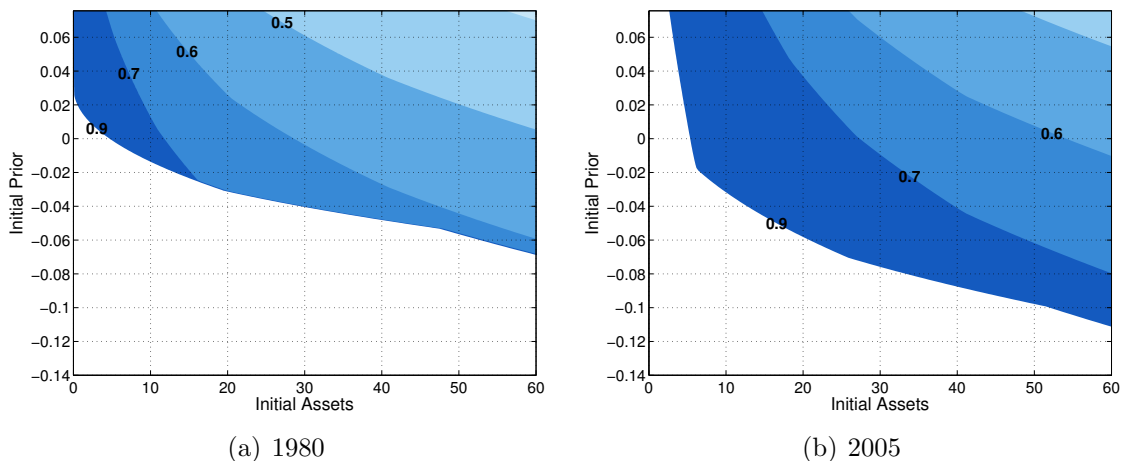


Fig. 5: Probability of Dropping Out of College

Horizontal axis: initial assets (a_1) as multiples of the average high-school worker’s two-year (ages 19 and 20) earnings; vertical axis: mean of initial prior (μ_{z_1}). The shaded region represents college enrollment. For those who enroll, ex-ante probabilities of not completing a four-year degree are shown in contour plots.

enrollees, in the darker areas, are more likely to drop out of college than those who are wealthier or have higher prior means. Wealth plays a more important role in college enrollment and dropout decisions in 2005 than in 1980, which can be explained by the higher costs of college and larger wage risk in 2005.

When we make the education choice a static decision—i.e., one has to decide and commit to being a high-school-only, some-college, or college-graduate worker before observing the in-college signal on his own returns—the enrollment threshold shifts to the northeast. One incentive to enroll in college is to obtain information on individual returns. The loss of this incentive is enough to push those near the enrollment threshold away from colleges. Those near the northeast corner will choose to be a college graduate, and those near the new threshold will choose to be a some-college worker. The mass of the latter depends on the cost of college during the first two years relative to the final two years and also the some-college partial return parameter (m_d).

For this exercise, we make the in-college signal worthless—that is, we set the signal-noise ratio s_ϵ to one by letting $\sigma_\epsilon^2 \rightarrow \infty$. For each year, we hold all other parameters at their respective calibrated values. The resulting moments are shown in columns $1980_{s_\epsilon=1}$ and $2005_{s_\epsilon=1}$ of Table 7.

With the information value of college enrollment gone, those who were close to the original enrollment threshold now do not enroll, lowering the enrollment rate (by 13 and 7 percentage points for 1980 and 2005 respectively). In addition, because we have fewer marginal students and also there is no learning about one’s returns in college (in particular, no correction of over-optimism), the college dropout rate is lower too (by 39 percentage points for 1980 and 19 percentage points for 2005). Overall, the

Moment	1980	1980 _{s_c=1}	2005	2005 _{s_c=1}
Enrollment rate	0.577	0.450	0.646	0.573
Discontinuation rate	0.393	0.001	0.451	0.263
Underemployment rate	0.237	0.263	0.326	0.341
College wage premium	0.361	0.259	0.934	0.823
Some-college premium	0.057	0.104	0.214	0.228
Underemployment premium	0.110	0.011	0.467	0.375

Table 7: Information Value of College

Effects of removing the in-college signal of one’s own returns in 1980 and 2005. The results are in columns 1980_{s_c=1} and 2005_{s_c=1}.

fraction of people with a four-year college degree goes up. This implies more selection at enrollment, but less at graduation. Hence, while it must be the case that the some-college wage premium goes up, whether the college graduates’ wage premium goes up or down depends on the relative magnitudes of the more positive selection at enrollment and the less positive selection at graduation. The quantitative result shows that the latter prevails: The resulting college premium is lower compared to the 1980 and 2005 benchmarks, by 10 and 11 percentage points, respectively. The less positive selection into graduation is also consistent with the drop in the underemployment premium, even though the underemployment rates are only slightly higher than in the benchmark for both years.

5.4 Underemployment Option

To evaluate the quantitative role of the underemployment option, we recalibrate the model without it in this section. To be specific, instead of the max distributions constructed in Section 3.2, we simply assume that the log of some-college and college wages are drawn from the normal distributions

$$\mathcal{N}\left(\log m_d + z - \frac{\sigma_d^2}{2}, \sigma_d^2\right), \quad \mathcal{N}\left(z - \frac{\sigma_c^2}{2}, \sigma_c^2\right),$$

where m_d is the partial return to two years of college education.

The rest of the model remains the same, so this simple model has one fewer parameter (m_u) than the benchmark. We recalibrate this model to the same set of moments in Table 3, minus the two underemployment moments. The resulting parameters are reported in Table 8.

A comparison of Tables 3 and 8 reveals that the model without the underemployment option needs a larger rise in both optimism (b_z) and mean returns (μ_z) between 1980 and 2005 to fit the data: Without the option-like effect of underemployment, a larger increase in μ_z is necessary to generate the increase in college premia, and, to match the rise in the college discontinuation rate at the same time, a higher b_z is called

Parameters	1980	2005
b_z (“optimism”)	0.266	0.359
s_ϵ (signal-noise ratio, $1+\sigma_{z_1}^2/\sigma_\epsilon^2$)	1.423	1.340
σ_z^2 (pop. variance of z)	0.091	0.172
μ_z (pop. mean college returns)	0.083	0.261
m_d (partial return, some-college)	0.997	0.960
σ_c^2 (var. of log residual college wage)	0.239	0.320
σ_d^2 (var. of log residual some-college wage)	0.158	0.206
$\rho_{az} \equiv \rho(a, z)$	0.497	[0.497]
$\rho_{az_1} \equiv \rho(a, \mu_{z_1})$	0.399	0.429
$\rho_{zz_1} \equiv \rho(z, \mu_{z_1})$	0.803	0.800

Table 8: Recalibrated Parameters without Underemployment

All parameters are calibrated to the moments in Table 3, except $\rho(a, z)$ for 2005, which we fix to its 1980 value because we lack the corresponding target moment in 2005. Since this simple model has no underemployment, the underemployment rate and premium are ignored.

for. This is consistent with the explanation given in other models of higher education, such as [Hendricks and Schoellman \(2014\)](#).

In contrast, with underemployment, the same phenomenon is primarily explained by a rise in the second moment of the return and residual wage distributions. In addition, the underemployment moments in the data were informative about the relative contribution of the degree of heterogeneity in returns to college (σ_z^2) versus the residual college wage dispersion (σ_c^2).

Perhaps more important, in the benchmark model the third moment (skewness) of the residual wage distribution increases in response to a larger second moment of the return and residual wage distributions. In the data, as shown in Figure 2(a), the 90-50 ratio of the residual wage distribution rose by almost 10 log points more than the 50-10 ratio, and the skewness increased from -0.04 to 0.04 between 1980 and 2005. In the simple model without underemployment, the rise in the 90-50 and 50-10 residual wage ratios are identical, and the residual log wage distribution has zero skewness in both 1980 and 2005, because of the log-normality assumption. In the benchmark model, with higher second moments, the 90-50 ratio rises by two log points more than the 50-10 ratio, and skewness rises from 0.01 to 0.04, because underemployment is operationalized as choosing the maximum of three log-normal random variables. While this rise in skewness is small compared to the data, we emphasize that our model can qualitatively shift the skewness at all, out of changes in the second moment of the underlying distribution.

6 Counterfactual Analysis

As summarized in [Card \(1999\)](#), there is a large empirical literature using instrumental variables to estimate the average marginal returns to college, or local average treatment effect of sending a marginal student to college. Many find that the marginal returns to college are close to or even higher than the population average. [Heckman et al. \(2006\)](#) emphasizes the difficulty of controlling for “selection on gains” with instruments, and [Carneiro et al. \(2011\)](#) estimates the average marginal returns to college relaxing parametric assumptions. These approaches still find sizable returns to college for marginal students, though smaller than the population average.

What is different in our approach is that we not only explicitly model selection at the enrollment margin but also at the next stage in which college students decide whether or not to complete a four-year degree. Selection is based on students’ (updated) priors on their own returns to college. In particular, selection at the graduation stage implies that returns to college are nonlinear in years in college: Given the selection into graduation based on individuals’ own returns, the average college premium of those who complete a four-year college degree is more than double the average some-college premium of those who quit college after two years, even if they looked identical at the enrollment stage.²⁹ If we had only looked at the wages of marginal enrollees who eventually graduate with a four-year degree, or assumed that returns to college are linear in years in college, or assumed exogenous college dropouts—as is not uncommon—the estimated marginal returns to college would be biased upward.³⁰

6.1 Marginal Enrollees

Using our model, we compute the wage premium that high-school graduates who are indifferent to college enrollment can expect upon college enrollment. We will highlight the roles of return heterogeneity and of the option to discontinue college education. Before proceeding, we note that our calibration is based only on prime-age white males in the CPS.

The marginal enrollees in this exercise are defined as those whose state variables (a_1, μ_{z_1}) lie close to the enrollment threshold in [Figure 5](#). Because the counterfactual exercises that follow focus on financial assistance, we consider those students who are

²⁹Such a nonlinearity is documented especially for more recent cohorts—e.g. [Heckman et al. \(2006\)](#). Previously it was emphasized in the literature on “sheepskin” effects, e.g. [Kane and Rouse \(1995\)](#)—the idea is that college degrees are signals of worker quality to potential employers. In our model, the nonlinearity reflects selection into graduation based on individual returns. [Lange and Topel \(2006\)](#) had a similar view.

³⁰In the estimation of empirical ordered-choice models, one would need as many instruments as there are stages. In our context, one set of instruments are necessary for the college enrollment decision, and then another for the college graduation decision. This is a serious challenge in practice, and many researchers tend to focus only on the enrollment decision.

Moment	1980		2005	
	Population	Marginal	Population	Marginal
Enrollment premium	0.242	0.067	0.609	0.159
College wage premium	0.361	0.232	0.934	0.329
Discontinuation rate	0.393	0.824	0.451	0.952
Some-college premium	0.057	0.031	0.214	0.151
Underemployment rate	0.237	0.303	0.326	0.414
Underemployment premium	0.110	0.046	0.467	0.319

Table 9: Population vs. Marginal Enrollees, 1980 and 2005.

Marginal enrollees are one percent of the population along the enrollment threshold. We force all of them to enroll and compute the moments (the “Marginal” columns).

marginal in terms of initial wealth, a_1 .

Using the enrollment indicator function $\chi_E(a_1, \mu_{z_1})$ in equation (6), we define

$$\chi_\varepsilon(a_1, \mu_{z_1}) = \chi_E(a_1 + \varepsilon, \mu_{z_1}) - \chi_E(a_1 - \varepsilon, \mu_{z_1}), \quad (13)$$

which is also an indicator function: $\chi_\varepsilon(a_1, \mu_{z_1})$ is equal to one if and only if the student with $(a_1 + \varepsilon, \mu_{z_1})$ decides to enroll in college but the one with $(a_1 - \varepsilon, \mu_{z_1})$ decides not to. This is the sense in which the indicator function χ_ε represents marginal students. We set ε such that the mass of marginal students $\int \chi_\varepsilon dF_0$ is one percent of the population. We then force all individuals with $\chi_\varepsilon = 1$ to enroll in college.³¹

The relevant moments for the population and the marginal enrollees are compared in Table 9, separately for 1980 and 2005. We first define “enrollment premium.” After enrollment, some will quit college as a some-college worker and others enter the labor market as a college graduate, with significant wage differentials between the two groups. Enrollment premium is the average of some-college premium and college-graduate wage premium, weighted by the college dropout and graduation rates.

In both years, the enrollment premium for marginal enrollees is far below the population average, and by much more in 2005. The premium is driven down by the fact that around 90 percent of the marginal enrollees will quit college and become some-college workers, as shown in the college dropout probabilities in Figure 5.³² Some-college workers as a group enjoy a much smaller wage premium than college graduates.

³¹Insofar as the reason that students who live farther from colleges do not enroll is because of commuting or moving costs, the marginal students in this exercise can be compared to the marginal students in studies using distance to college as an instrument. For comparison, we ran OLS and IV regressions on our simulated data, using assets as an instrument for college enrollment and graduation. We find that the IV enrollment premium is in fact lower than in an OLS regression, but the graduation premium is larger, consistent with our subsequent analysis.

³²This is consistent with the evidence that most college dropouts are from two-year or lower-tier four-year public institutions (Bound and Turner, 2011). A marginal student will likely attend such institutions.

While the correlation between one’s true return z and prior mean μ_{z_1} (ρ_{zz_1}) is not perfect—0.64 in 1980 and 0.76 in 2005, it is strong enough to generate meaningful positive selection in terms of z at the enrollment margin. Put differently, the fact that a marginal student decided not to enroll based on low μ_{z_1} implies that he likely has a lower true return than those who chose to enroll.³³ A low true return implies a high probability of quitting college midway, even after enrollment. If he were one of the few in the marginal group who turn out to have been too pessimistic about their own z , he will graduate from college. However, he still has a higher probability of being underemployed than other college graduates, because his true return is likely lower than the graduate average. In addition, even conditioning on education attainment or underemployment status, he is expected to earn less than others in the same category, again because his true return is likely lower than the group average.

6.2 Tuition Support for First Two Years of College

Our results indicate that, to a large extent, college enrollment and graduation positively select students in terms of their true returns. In particular, those near the enrollment threshold may have little to gain by going to college, because the realized large returns to college come from those who graduate, but most marginal enrollees would quit college following an information update. Having said that, especially for 2005, initial wealth (captured by family income) is an important determinant of college enrollment (Figure 5). We now ask whether financial intervention, such as college tuition subsidies, may be justified by a simple cost-benefit analysis. In this section, motivated by the proposals by leading politicians including President Obama, we quantitatively assess the impact of a policy that makes the first two years of college education free.

Specifically, starting from our calibrated 2005 economy, we make the first two years of college free for all individuals, i.e., $x_1 = 0$ in the individual’s problem (5). We call those who change their enrollment decision (from no to yes) because of the subsidy “switchers.”³⁴ We re-compute the population moments (“After”) and compare them with the 2005 benchmark (“Before”) in Table 10. We also compute the moments for all switchers (“All”), and then separately look at the switchers from the bottom (“Poor”) and top (“Rich”) quartiles of the initial population wealth distribution.

With the subsidy, the college enrollment rate rises by 11 percentage points. However, as in Section 6.1, the switchers are by construction marginal enrollees, and their true returns are lower on average than the returns of those who would enroll regardless of the subsidy. Most of the switchers (97 percent) quit college before earning a

³³This is consistent with the finding in [Carneiro and Heckman \(2002\)](#) and [Cameron and Taber \(2004\)](#) that students at the enrollment stage are sufficiently well-informed and financially unconstrained.

³⁴This switcher group includes all the marginal enrollees with $\chi_e = 1$ in (13).

Moment	Population		All	Switchers	
	Before	After		Poor	Rich
Enrollment rate	0.646	0.752	[0.106]*	[0.029]*	[0.016]*
Enrollment premium	0.609	0.511	0.130	0.090	0.163
College wage premium	0.934	0.916	0.361	0.340	0.312
Discontinuation rate	0.451	0.511	0.974	0.985	0.943
Some-college wage premium	0.214	0.192	0.124	0.086	0.154
Underemployment rate	0.326	0.328	0.373	0.462	0.370
Underemployment premium	0.467	0.458	0.098	0.152	0.121

* Those who enroll because of the subsidy (fraction of total population)

Table 10: Subsidizing Enrollment: Free College for First Two Years (2005)

First column moments are from the benchmark 2005 calibration. Second column (“After”) is the population moments after the subsidies are implemented. Third column (“All”) shows the moments for those who enroll because of the subsidy, or “switchers.” The last two columns are for two sub-groups of switchers: “Poor” for those in the bottom quartile of the initial population wealth distribution and “Rich” for the top quartile.

four-year degree. Even those who graduate are more likely to be underemployed than other college graduates. In addition, even after conditioning on education and underemployment, they earn less on average than others in the same group. Their average enrollment premium is only 13 percent, a small fraction of the 61-percent enrollment premium enjoyed by those who would enroll regardless of the subsidy. For the economy as a whole, they raise the college dropout rate and the underemployment rate—only slightly, since so few become college graduates—and depress all education premia.

Since the motivation for real-world proposals of this kind is to give poor students opportunities to get ahead, we explore whether poor and rich switchers are affected differently. In terms of enrollment, as expected, the policy positively affects the poor more than the rich. Of the 11-point increase in the enrollment rate, 3 points come from the bottom quartile of initial wealth, while only 1.6 points come from the top quartile.

However, the poor switchers gain on average much less than the rich switchers in terms of education premia. One reason is that they are more likely to quit college than the rich switchers. Although the policy covers the first two years’ college costs (x_1), those who wish to transfer to or continue in a four-year institution now have to pay x_2 (net of existing grants). Similar to the enrollment threshold in Figure 5, although not shown in the paper, the graduation threshold level of the updated prior mean μ_{z_2} in equation (7) declines with wealth: that is, for a given updated prior mean, those with assets below a threshold quit college, and those above complete four-year degrees. Because the some-college premium is much lower than the college premium, the higher dropout probability of the poor switchers implies that their enrollment premium is lower than the rich switchers’.

The other reason is that the poor switchers’ some-college premium is lower than the

rich switchers’. This requires more explanation. At a glance, this sounds contradictory to the first reason: Because the threshold μ_{z_2} decreases in wealth, the poor switchers who quit college must have a higher μ_{z_2} on average than the rich switchers who drop out. If μ_{z_2} is correlated with true returns z , shouldn’t we see a higher some-college premium among the poor switchers than the rich ones? The answer is that we also have to consider the correlation between initial wealth and true returns: ρ_{az} in Table 4. For a given level of a prior mean, the rich are more likely to have a higher true return z than the poor, because initial wealth and true returns are positively correlated. The lower some-college premium of poor switchers implies that the correlation between initial wealth and true returns dominates the correlation between the updated prior mean and true returns for the switcher group.

We now ask whether this subsidy program is justified by a hypothetical cost-benefit analysis: We compare the increase in the present discounted value of lifetime earnings of switchers with the cost of the government paying the two-year college tuition x_1 only for them. This analysis presupposes that the program can perfectly target the switchers. It is an upper bound on the net benefit per switcher of the program, since any tuition subsidy paid to those who would have enrolled even without it will have a much lower net benefit.³⁵ At the same time, this thought experiment is closer to real-world proposals, which will subsidize only the cost of attending community colleges and lower-tier public universities.

On average, a switcher makes \$19,036 more in lifetime earnings, discounted to age 19. The two-year subsidy costs the government \$16,958 per switcher. In addition, while the program only covers the first two years of college, after which most of the switchers quit, the few switchers that decide to complete a four-year degree will claim a fraction of their final two-year college cost, $v(a_1)x_2$, from the pre-existing grants program. This expected additional cost is \$41 per switcher. Subtracting the total cost per switcher of \$16,999 from the benefit of \$19,036, we arrive at a positive but negligible net benefit of \$2,037. To put things in perspective, this would require that 90 percent of the increased earnings be taxed for the program to break even.

6.3 Tuition Support for Final Two Years of College

In the previous section, even among the marginal students who enroll because of the subsidy, the realized college premium for the few who eventually graduate from college is significantly higher than the some-college premium for those who quit college.

³⁵There are two model details that we want to clarify. First, if we make two years of college free for everyone, even the lifetime earnings of those who would have enrolled anyway increase on average: Because they now have more assets after the first two years of college (i.e., a_2), more of them choose to graduate from college and earn higher wages on average. Second, paying such a student’s two years of college tuition costs less than x_1 because he was already paid grants $v(a_1)x_1$. The new subsidy is picking up $(1 - v(a_1))x_1$.

Moment	Population		All	Switchers	
	Before	After		Poor	Rich
Discontinuation rate	0.451	0.306	[0.094]*	[0.023]*	[0.032]*
Some-college wage premium	0.214	0.176	-	-	-
College wage premium	0.934	0.810	0.342	0.360	0.328
Underemployment rate	0.326	0.340	0.393	0.399	0.398
Underemployment premium	0.467	0.369	0.058	0.061	0.044

* Those who graduate from college because of the subsidy (fraction of total population)

Table 11: Subsidizing Graduation: Free Final Two Years of College (2005)

Here “switchers” are those who graduate from college (instead of dropping out) because of the subsidy. The result is for the first cohort to benefit unexpectedly from the new subsidy program—i.e., the subsidy did not affect their enrollment decision. First column moments are from the benchmark 2005 calibration. Second column (“After”) is the outcome for this cohort as a whole. Third column (“All”) is for all switchers in the cohort. The last two columns are for two sub-groups of switchers: “Poor” for those in the bottom quartile of the initial population wealth distribution and “Rich” for the top quartile.

This observation, combined with the fact that the final two years of college are more expensive than the first two (i.e., $x_2 > x_1$), leads us to another policy experiment: making the last two years of college free. Here we analyze the “short-run” effect by focusing on the first cohort to benefit unexpectedly from this policy—that is, their enrollment decisions are not affected by the anticipation of receiving such subsidies.³⁶

The results are shown in Table 11. In the context of this backloaded subsidy, “switchers” are those who graduate from college instead of dropping out because of the subsidy. In the first column are the moments from the benchmark 2005 calibration. The second column is the outcome for this cohort as a whole, after the policy implementation. The third column (“All”) shows the moments for all switchers in the cohort. The remaining two columns are for two sub-groups of switchers: “Poor” for those in the bottom quartile of the initial population wealth distribution and “Rich” for those in the top quartile. By definition, none of the switchers quits college.

The policy has a strong positive impact on the fraction of college graduates in the cohort—it rises from 35 percent in the 2005 benchmark to 45 percent, as the college dropout rate falls from 45 to 31 percent. However, the switchers’ college wage premium is well below the other college graduates’: as college graduates, switchers earn on average 34 percent more than high-school workers, while the figure is 93 percent for the other graduates. Part of the reason is that the switchers are more likely to be underemployed (39 vs. 33 percent). However, whether they are underemployed or not, they earn much less than the other college graduates. The real reason is similar to what we found when subsidizing the first two years: The decision to quit or complete

³⁶We also worked out the effect on younger cohorts who know of the subsidy when making enrollment decisions. We find only minor quantitative effects—for example, enrollment rates increase by less than one percentage point—and accordingly omit this result from the paper.

college is largely based on students' updated beliefs on their true returns to college, and the beliefs at this stage are even more consistent with the true returns than at the enrollment stage. The switchers on average have low true returns than those who would graduate from college regardless of the subsidy, as is demonstrated by their post-graduation wages. We find that the results are broadly similar across switchers of different initial wealth levels.

We can calculate that the switchers on average would have earned a some-college wage premium of 29 percent had they not graduated from college. With the subsidy, their average wage gain is 5 percent of the average high-school wage (34 minus 29 as percent of the average high-school wage), which is even smaller than the average wage gain from subsidizing the first two years. To the extent that students learn about their individual returns while in college, selection on returns should be sharper at the graduation stage than at the enrollment stage, and it makes sense that the wage gains from subsidized graduation are smaller.³⁷

With the smaller wage gain and the larger cost of subsidizing the last two years (i.e., $x_2 > x_1$), a cost-benefit analysis as in the previous section finds a *negative* net benefit for this subsidy program.³⁸

7 Concluding Remarks

Our model broadens the standard education choice model by considering the college completion decision and the possibility of underemployment as a college graduate. These options interact with rising returns-to-college heterogeneity and risk in a way that reconciles the stagnant college enrollment and graduation rates with the rising college wage premium.

Extending the analysis, we find that policies encouraging college enrollment of marginal students would only lead to small wage gains: Such students are highly likely to quit college without earning a four-year degree, and some-college wage premium is only a fraction of college-graduate premium. That is, one should not point to the rising headline college wage premium and jump to the conclusion that too few are going to college. On the other hand, some commentators think of underemployment as proof that too many people are going to (and finishing) college. In our model, underemployment is a natural outcome of the real option inherent in the risky human

³⁷This is much smaller than the regression discontinuity estimate in [Zimmerman \(2014\)](#), which finds that the marginal students admitted to the flagship public university in Florida typically graduate and earn 20 percent more than those comparable but not admitted, many of whom become some-college workers.

³⁸In relation to the literature—e.g., [Bowen et al. \(2009\)](#) and many others reviewed in [Bound and Turner \(2011\)](#)—that emphasizes policy interventions at the college completion margin rather than the enrollment margin, our result suggests that unconditional tuition subsidies may not be the most effective policy tool.

capital investment of college education. In fact, it is part of what makes a college degree valuable ex ante, protecting graduates from the left-tail risk in the college-graduate labor market.

The one important question this paper sidestepped is what determines the individual return to college (or z_i). We took its distribution in the population as given, and conjectured that it may capture one's innate ability, human capital acquired from early childhood through high school, the quality of colleges one can be admitted to, and so on. When we have a better understanding of what determines z_i and also through what mechanism, we will have productive debates about what policy interventions will be most effective in terms of wage gains. Still, in relating it to educational attainment and education wage premia, it will be important to explicitly take into account the endogenous selection at the college enrollment and completion stages.

References

- Abbott, B., G. Gallipoli, C. Meghir, and G. L. Violante (2014). Education policy and intergenerational transfers in equilibrium. Working Paper 18782, NBER.
- Altonji, J. G. (1993). The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics* 11(1), 48–83.
- Altonji, J. G., P. Arcidiacono, and A. Maurel (2015). The analysis of field choice in college and graduate school: Determinants and wage effects. Manuscript.
- Arcidiacono, P., P. Bayer, and A. Hizmo (2010). Beyond Signaling and Human Capital: Education and the Revelation of Ability. *American Economic Journal: Applied Economics* 2(4), 76–104.
- Athreya, K. and J. C. Eberly (2013). The supply of college-educated workers: The roles of college premia, college costs, and risk. Working Paper 13-02, Federal Reserve Bank of Richmond.
- Autor, D. H., L. F. Katz, and M. S. Kearney (2008). Trends in U.S. wage inequality: Revising the revisionists. *Review of Economics and Statistics* 90(2), 300–323.
- Bailey, M. J. and S. M. Dynarski (2011). Gains and gaps: Changing inequality in u.s. college entry and completion. Working Paper 17633, NBER.
- Betts, J. R. (1996). What do students know about wages? evidence from a survey of undergraduates. *Journal of Human Resources* 31(1), 27–56.
- Bound, J. and S. Turner (2011). Dropouts and diplomas: The divergence in collegiate outcomes. In E. A. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education*, Volume 4, pp. 573–613. Amsterdam: Elsevier.
- Bowen, W. G., M. M. Chingos, and M. S. McPherson (2009). *Crossing the Finish Line: Completing College at America's Public Universities*. Princeton University Press.
- Brown, J., C. Fang, and F. J. Gomes (2015). Risk and returns to education over time. Discussion Paper 10416, Centre for Economic Policy Research.
- Cameron, S. V. and C. Taber (2004). Estimation of educational borrowing constraints using returns to schooling. *Journal of Political Economy* 112(1), 132–182.
- Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3A, pp. 1801–1863. Amsterdam: Elsevier.
- Card, D. and J. E. DiNardo (2002). Skill-biased technological change and rising wage inequality: Some problems and puzzles. *Journal of Labor Economics* 20(4), 733–783.
- Carneiro, P. and J. J. Heckman (2002). The evidence on credit constraints in post-secondary schooling. *Economic Journal* 112(482), 705–734.

- Carneiro, P., J. J. Heckman, and E. J. Vytlačil (2011). Estimating marginal returns to education. *American Economic Review* 101(6), 2754–81.
- Chay, K. Y. and D. S. Lee (2000). Changes in relative wages in the 1980s: Returns to observed and unobserved skills and black-white wage differentials. *Journal of Econometrics* 99(1), 1–38.
- Chen, S. H. (2008). Estimating the variance of wages in the presence of selection and unobserved heterogeneity. *Review of Economics and Statistics* 90(2), 275–289.
- Clark, B., C. Joubert, and A. Maurel (2015). The career prospects of overeducated Americans. Manuscript.
- Cunha, F. and J. J. Heckman (2007). The evolution of inequality, heterogeneity and uncertainty in labor earnings in the U.S. economy. Working Paper 13526, NBER.
- Cunha, F., J. J. Heckman, and S. Navarro (2005). Separating uncertainty from heterogeneity in life cycle earnings. *Oxford Economic Papers* 57(2), 191–261.
- Gale, W. G. and J. K. Scholz (1994). Intergenerational transfers and the accumulation of wealth. *Journal of Economic Perspectives* 8(4), 145–160.
- Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2015). What do data on millions of U.S. workers reveal about life-cycle earnings risk? Manuscript.
- Heckman, J., L. Lochner, and C. Taber (1998). Explaining rising wage inequality: Explanations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of Economic Dynamics* 1(1), 1–58.
- Heckman, J. J., L. J. Lochner, and P. E. Todd (2006). Earnings functions, rates of return and treatment effects: The Mincer equation and beyond. In E. A. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Volume 1, Chapter 7, pp. 307–458. Amsterdam: North Holland.
- Heckman, J. J. and S. Urzua (2009). The option value of educational choices and the rate of return to educational choices. Manuscript.
- Heckman, J. J., S. Urzua, and E. Vytlačil (2006). Understanding instrumental variables in models with essential heterogeneity. *Review of Economics and Statistics* 88(3), 389–432.
- Hendricks, L. and O. Leukhina (2017, October). How Risky is College Investment? *Review of Economic Dynamics* 26, 140–163.
- Hendricks, L. and T. Schoellman (2014). Student abilities during the expansion of US education. *Journal of Monetary Economics* 63(1), 19–36.
- Horn, L. and P. Skomsvold (2011). Community college student outcomes: 1994-2009. Technical Report 2012-253, National Center for Education Statistics.

- Hoxby, C. M. (2009). The changing selectivity of American colleges. *Journal of Economic Perspectives* 23(4), 95–118.
- Huggett, M., G. Ventura, and A. Yaron (2011). Sources of lifetime inequality. *American Economic Review* 101(7), 2923–2954.
- Johnson, M. T. (2013). Borrowing constraints, college enrollment, and delayed entry. *Journal of Labor Economics* 31(4), 669 – 725.
- Kane, T. J. and C. E. Rouse (1995). Labor-market returns to two- and four-year college. *American Economic Review* 85(3), 600–614.
- Katz, L. F. and K. M. Murphy (1992). Changes in relative wages, 1963-1987: Supply and demand factors. *Quarterly Journal of Economics* 107(1), 35–78.
- Keane, M. P. and K. I. Wolpin (1997). The career decisions of young men. *Journal of Political Economy* 105(3), 473–522.
- Lange, F. and R. Topel (2006). The social value of education and human capital. In E. A. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Volume 1, pp. 459–509. Amsterdam: North Holland.
- Lee, D. and K. I. Wolpin (2006). Intersectoral labor mobility and the growth of the service sector. *Econometrica* 74(1), 1–46.
- Lemieux, T. (2010). What do we really know about changes in wage inequality? In *Labor in the New Economy*, pp. 17–59. NBER.
- Leuven, E. and H. Oosterbeek (2011). Overeducation and mismatch in the labor market. Discussion Paper 5523, Institute for the Study of Labor.
- Lochner, L. J. and A. Monge-Naranjo (2011). The nature of credit constraints and human capital. *American Economic Review* 101(6), 2487–2529.
- National Student Clearinghouse Research Center (2012). Snapshot report on degree attainment.
- Stange, K. M. (2012). An empirical examination of the option value of college enrollment. *American Economic Journal: Applied Economics* 4(1), 49–84.
- Stinebrickner, T. and R. Stinebrickner (2012). Learning about academic ability and the college dropout decision. *Journal of Labor Economics* 30(4), 707 – 748.
- Stinebrickner, T. and R. Stinebrickner (2014). Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model. *Journal of Labor Economics* 32(3), 601 – 644.
- Taber, C. R. (2001). The rising college premium in the eighties: Return to college or return to unobserved ability? *Review of Economic Studies* 68(3), 665–691.
- Zimmerman, S. D. (2014). The returns to college admission for academically marginal students. *Journal of Labor Economics* 32(4), 711–754.

A Data Appendix

A.1 March CPS

Most of the data moments used in the analysis are computed from the IPUMS CPS, generally following the data-cleaning procedure of [Autor et al. \(2008\)](#). One difference is that we do not exclude the self-employed or farmers from the sample, because we do not think that education decisions are made only in anticipation of becoming a wage-worker.

Earnings We closely follow AKK, especially when handling top-coding.

1. First, we select all race categories containing “white,” and focus only on individuals in the private labor force.
2. Select only full-time workers, defined as individuals who report they worked 40+ weeks last year, and 35+ usual hours worked per week.
3. Follow AKK to adjust for top-coding by income source (“wages and salary,” “self-employed,” and “farm”). From 1988 onward, each income source is further divided into primary and secondary sources of income. All income sources and categories are separately top-coded. AKK’s procedure is to multiply top-coded incomes by 1.5 of the top-code value, available from the BLS.
4. (*Top trimming*) Compute implied average weekly and hourly earnings. Drop individuals with implied average weekly earnings larger than the top-code value times 1.5/40. Repeat with implied average hourly earnings.
5. Define total earnings as the top-coding adjusted sum of all three (or six for 1988 onward) income categories, also adding primary and secondary sources when both present. Earnings are reported in annual terms. Deflate all earnings to 2000 USD, using the chain-weighted (implicit) price deflator for personal consumption expenditures.
6. (*Bottom trimming*) Drop individuals whose average weekly or hourly earnings are less than one-half the real minimum wage in 1982 (\$112/week in 2000 dollars).

Educational attainment Prior to 1992, the CPS only collected respondent’s highest grade completed. From 1992 onward, they additionally collected data on the highest degree or diploma attained.

1. HSD: less than 12 years of schooling, or report not having a high school diploma or equivalent
2. HSG: at least 12 years of schooling, or report having a high school diploma or equivalent

3. SMC: more than high school but less than 4 years of college, or report having a two-year degree
4. CLG: at least 4 years of college, or report having a four-year degree
5. GTC: at least 6 years of college or report having a post-college graduate degree

To compute statistics, we drop missing observations and use the CPS provided sampling weights, except when computing the earnings premia. To control for age-demographic effects, we use the sampling weights when computing the mean earnings for each age, but simply take the average across the mean earnings for each age from 26 to 50 when computing premia. To compute lifetime PDV earnings values, we discount mean earnings by age by an annual interest rate of 4 percent. For more details, see AKK.

A.2 BLS Education and Training Assignments and NLSY

BLS Education Projections (EP) The BLS publishes an education and training requirement table by detailed occupation in their *Occupation Outlook Quarterly* as part of its EP program (<http://www.bls.gov/emp/home.htm>). Published biennially, the table is based on the BLS's own qualitative analysis, employer information in the O*NET, and post-secondary education information from the NCES, combined with a quantitative analysis of the distribution of educational attainment in each detailed occupation category from the Census and the American Community Survey. We use the 1998 table as a benchmark and test robustness using the 2008 table. We use the 1998 table because the occupation coding changes frequently in both the CPS and the EP, and the 1998 table requires the minimum occupational coding crosswalks between 1980 to 2005.

The CPS uses the OCC1970 convention until 1982, OCC1980 until 1991, OCC1990 until 2002, and then OCC2000. Occupations in the 1998 requirement table are coded according to SOC2000. We use BLS crosswalks to change all occupation codes to OCC1990. Since crosswalks do not provide a perfect matching across the different coding conventions across the CPS and EP, we are forced to drop several occupations—that is, for some CPS occupations in a given year, we do not know the degree requirement.

In our empirical analysis, we define underemployment only for college graduates. To be specific, the underemployed are only those college graduates in the CPS who work in an occupation for which a (four-year) college degree is not required according to the requirements table. For approximately 25 percent of college graduates, we are unable to determine whether they are underemployed. In these cases, we simply assume they work in college-level jobs, so that we remain conservative on the prevalence of underemployment.

NLSY79 and 97 We compute age-earnings profiles in the NLSY by applying the same criteria we applied to the CPS, but dropping observations below half the 1982

minimum wage. We impute missing wage observations by linearly interpolating log wages in adjacent years, and compute lifetime PDV earnings by discounting at an annual interest rate of 4 percent. Missing observations are dropped when computing the correlation between family income and children’s lifetime PDV earnings.

To compute enrollment rates and family income in the NLSY79 and 97, we follow [Lochner and Monge-Naranjo \(2011\)](#), except that we drop high school dropouts. Then we drop children who were not living with an adult for both ages 16 and 17, and take family income as the mean reported values for those ages. We define college enrollment as having 13 or more years of schooling at age 21. If data is missing at age 21, we apply the same criterion at age 22. We drop missing observations to compute enrollment rates.

B Numerical Appendix

In an outer loop, we calibrate Θ (in Table 4) to match target moments M_d (in Table 3), using a downhill simplex method to solve equation (12). Note that all moments are in terms of percentages; the resulting RMSE is less than one percentage point.

B.1 Initialization

Grids All interpolations will be linear. Let n_a denote the size of the grids for a_i , $i \in \{1, 2, 3\}$, and n_z the size of the grids for the true returns z , prior μ_{z_1} , and posterior μ_{z_2} . The grid points for assets are set such that lower points are closer to each other, and the returns and posteriors grid points are set at equi-distant intervals. These grids are used as individual states.

Quadratures and terminal values We use k -dimensional Gaussian-Hermite quadratures whenever numerical integration is needed. Before we solve the individual’s problem, we can set all terminal values as follows:

1. Set quadratures over wages (w_h, w_d, w_c) . Separately from the z -grid, set k -dimensional quadratures over z for each node on the μ_{z_2} grid. Also set separate k -dimensional quadratures over μ_{z_2} for all nodes on the μ_{z_1} grid. All numerical integration is done over these quadratures.
2. For each node on the $n_a \times n_z$ grid for (a_i, z) , compute $V(s, a_i, w)$ and its first derivative w.r.t. a_i in equation (8) for all nodes on the (w_h, w_d, w_c) quadratures.
3. Compute $V_h(a_1), V_d(a_2, z), V_c(a_3, z)$ and their derivatives w.r.t. a_i by integrating over the w -quadratures.
4. For each node on the $n_a \times n_z$ grid for (a_2, μ_{z_2}) , compute the discontinuation option in equation (7) by integrating over the z -quadrature for μ_{z_2} .

5. Similarly, compute the college value for each node on the $n_a \times n_z$ grid for (a_3, μ_{z_2}) by integrating over the z -quadrature. Also obtain the derivative of the college option w.r.t. a_3 by integration.

This gives us all the terminal values and their derivatives.

B.2 Individual Decisions

Policy functions for V_2 Given the derivative of $\int V_c dF_{z_2}$, derive the savings policy a_3^* for the “school” option for each value of (v, a_2, μ_{z_2}) on the $3 \times n_a \times n_z$ grid. Once done, use the policy function to compute the value of the school option and the envelope theorem to compute $\partial V_2(v, a_1, \mu_{z_2})/\partial a_1$ if the school option is chosen. Compare with the work option to derive the binary policy function $\chi_G \in \{0, 1\}$ for 0 =work, 1 =school.

Policy functions for V_1 Given the derivative of V_2 , derive the savings policy a_2^* for the school option for each value of (a_1, μ_{z_1}) on the $n_a \times n_z$ grid, while using the μ_{z_2} -quadrature on each node of μ_{z_1} grid to integrate over V_2 and its derivative w.r.t. a_2 . Once done, use the policy function to compute the value of the school option. Compare with the work option to derive the binary policy function $\chi_E \in \{0, 1\}$ for 0 =work, 1 =school.

B.3 Approximating Distributions

Initial distribution For $s = 1$, linearly approximate a trivariate normal distribution over $(\log a_1, \mu_{z_1}, z)$ given a guess on the parameters.

Enroll? Solve for individual decisions on each point of the approximated distribution at $s = 1$, and compute the mass of individuals who do not enroll. For the rest, compute the masses that fall on an approximated quadivariate distribution over (a_2, μ_{z_2}, a_1, z) , which forms the approximated distribution for $s = 2$.

Discontinue? Solve for individual decisions on each point of the approximated distribution for $s = 2$, and compute the mass of individuals who discontinue. For the rest, compute the masses that fall on an approximated bivariate distribution over (a_3, z) , which forms the approximated distribution for $s = 3$.

Underemployed? The high school wage draw w_h is independent of z . For each z on the grid, we can compute the z -specific wage distributions for college non-graduates and graduates, according to the bivariate and trivariate max distributions for log normal random variables. The college graduates whose college-job wage is not the maximum wage draw are those that we dub underemployed.

All statistics are computed using these approximated distributions.

C Robustness and Alternative Specifications

C.1 Observed Wage Distributions

For all our analysis, we focused only on the first two moments of the education-specific wage distributions. In Figure 6, we compare the entire ex post lifetime wage distributions obtained from the model to the distributions we recover from the CPS data using a Q-Q plot. Specifically, we first demean log wages both in the data and the model, and plot the simulated quantiles of residual log wages against the empirical quantiles for each education category, separately for 1980 and 2005. A perfect match would lie along the 45-degree line. While the simulated distributions are not perfect toward the extreme points, overall the fit is reasonable. The fit will get significantly worse without the college dropout option. Without it, as we observed from Table 7, there is more selection into enrollment and less into graduation. This would push up the lower quantiles of the simulated some-college wage distribution. However, it is still the case that our simulated wage distributions cannot capture the right tail of the high-school wage distribution and the left tail of the college-graduate wage distribution. We attribute such shortcomings primarily to several abstractions we made in the construction of the model, which we discuss below.

C.2 Alternative Model Specifications

Heterogeneous skills specific to high-school jobs If we introduce another dimension of ex ante heterogeneity in the form of skills specific to high school jobs, students with a comparative advantage in those jobs would self-select into them. Unless the correlation between the high school job skill and return to college is strong, this new dimension would push up the right tail of the simulated high school wage distribution. It would also push down the simulated left tail of the college wage distribution, because college graduates who become underemployed (i.e., end up with high school jobs) would have a comparative disadvantage in high school jobs. Overall, it would improve the model fit of the education-specific wage distribution.

However, incorporating this new dimension of heterogeneity would not alter the importance of heterogeneous returns to college and option values in explaining the time trend in educational attainment and wage premia. The main mechanism in our model is the rising dispersion of returns to college interacting with the options embedded in college education. From 1980 to 2005, the wage inequality among high-school workers did not increase nearly as much as that among college graduates. It is not conceptually clear what the comparable options for high-school workers are either. This extension would only have modest consequences on aggregate moments.

Wage offers while in college While students in our model only make wage draws *after* they enter the labor market, other papers, e.g. [Stange \(2012\)](#), assume

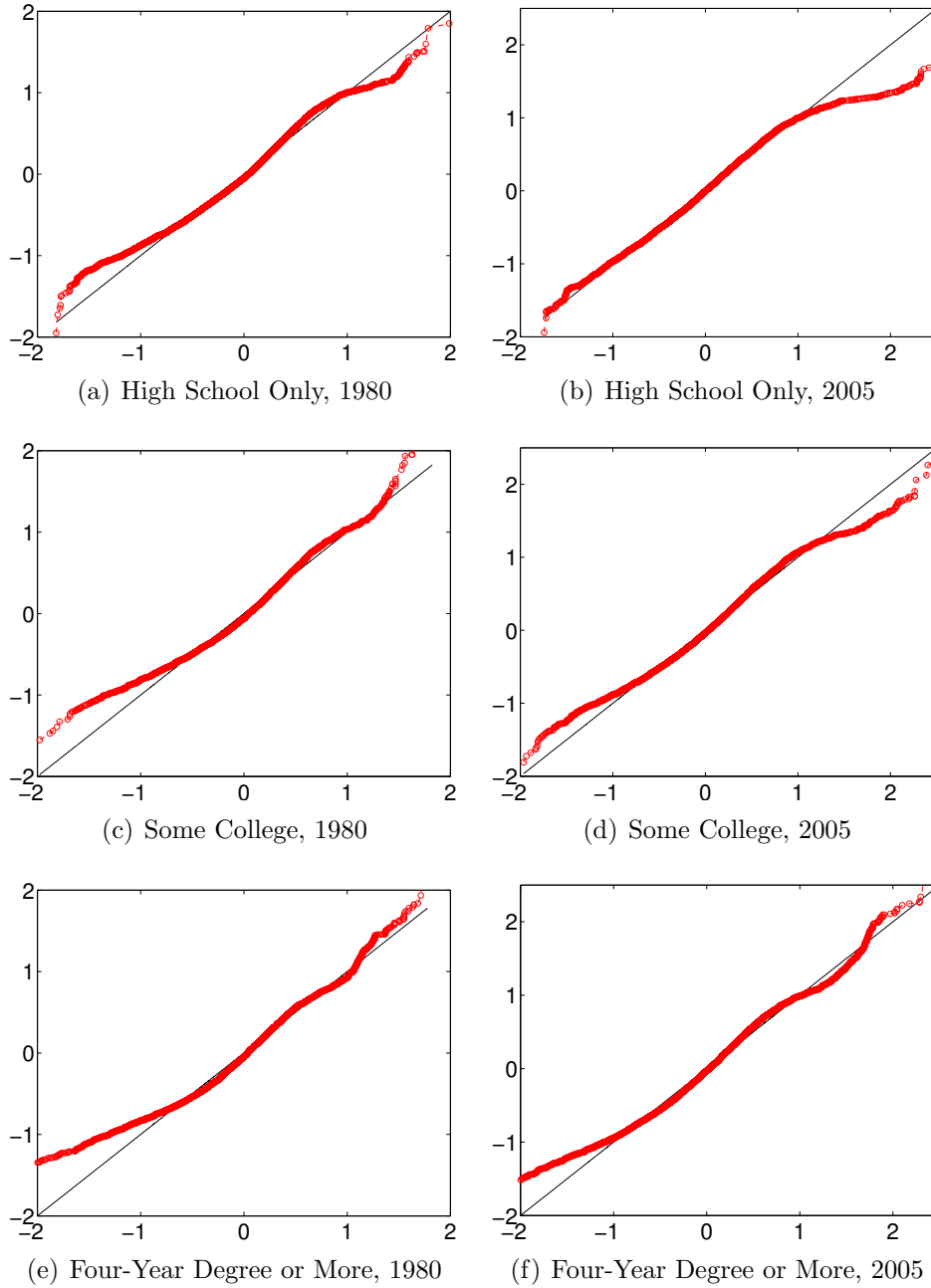


Fig. 6: Quantile-Quantile Plots

Each graph plots the model quantiles of the education-specific log wage distribution (demeaned) on the vertical-axis against the empirical quantiles on the horizontal-axis.

the opposite: wage draws are only made *while in* school. In-college wage draws would introduce a few different effects into our model. First, those who get lucky and draw a wage exceeding their reservation wage in college would quit education, shoring up the right tail of the some-college wage distribution. The flip side is that those who are not as lucky, having stayed in school in the hopes of drawing a better wage, will graduate college with a low wage, showing up in the left tail of the college wage distribution. On the other hand, if the in-college wage offers have information on individual-specific returns to college, it would strengthen the positive selection at the graduation stage. Finally, the wage draws in college increase the value of enrolling in college, implying less positive selection at the enrollment stage. The exact quantitative importance of these considerations in explaining the education and wage trends is left for future research.

D Additional Tables and Figures

	1980	2005
Total	0.7111	1.2243
Between	0.0785	0.2237
Within	0.6324	1.0001
HSG	0.5547	0.8730
SMC	0.5932	0.9589
CLP	0.8251	1.2595
Between CLP	0.0129	0.0430
Within CLP	0.8121	1.2165
CLG	0.7928	1.1212
GTC	0.9267	1.3047

Table 12: Theil Index Decomposition, White Males, Ages 26–50

HSG: high school graduate, SMC: some college, CLP: all college graduates, CLG: only college graduates, GTC: greater than college. To make the indices comparable across years and education groups, we normalize them by $\log N$, where N is the sum of the sampling weights, and multiply by 100. Total, between and within indices are normalized by the size of the entire sample in each year, times 100. Education specific indices are normalized by the size of each group in each year, times 100.

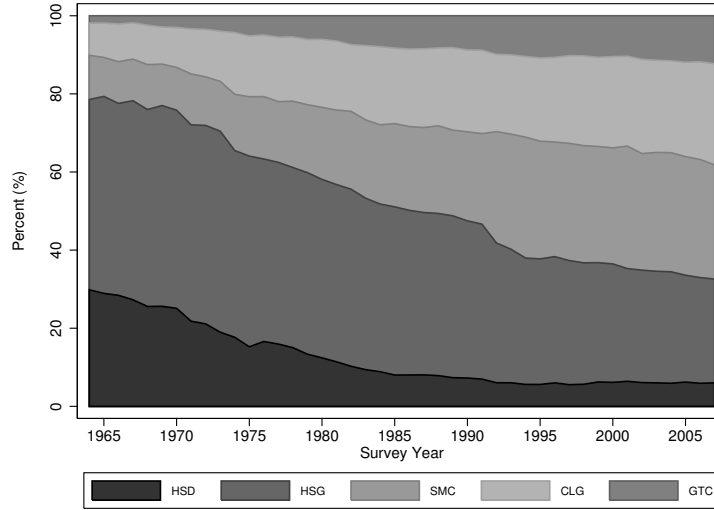


Fig. 7: Education Attainment, White Females, Ages 26-50

HSD: high school dropout, HSG: high school graduate, SMC: some college, CLG: college graduate, GTC: greater than college.

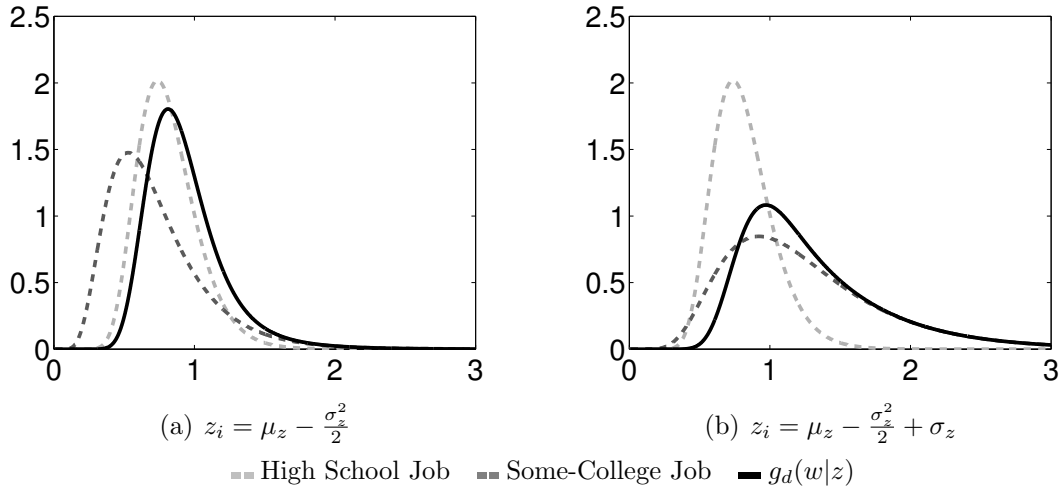


Fig. 8: Some College Wage Distribution

Some-college wage distribution for an individual whose return z_i is exactly at the population mean (left panel), and one standard deviation above the mean (right panel). The horizontal axis is multiples of mean high school wage. The G distribution is determined by the larger wage of a high school job, which would lead to underemployment, and a some-college job.

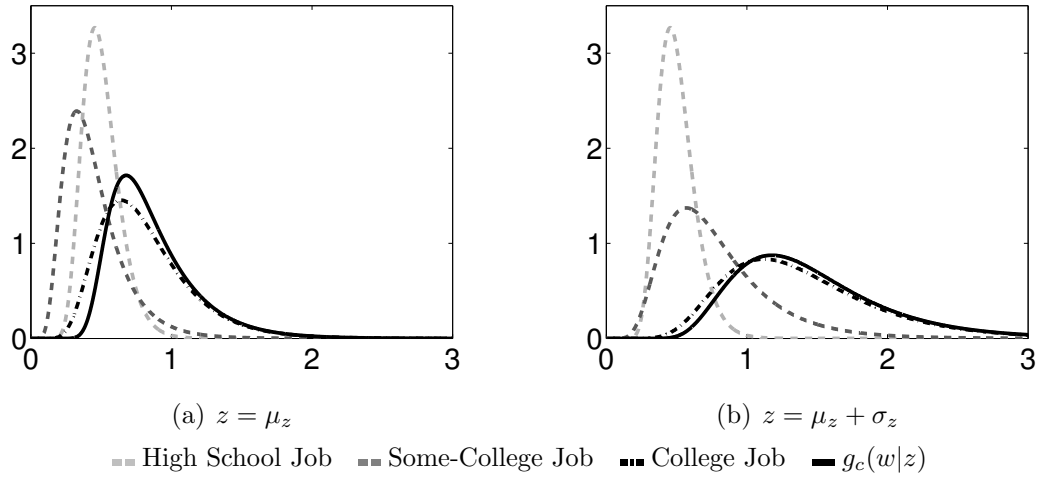


Fig. 9: College Graduate Wage Distribution

College graduate wage distribution for an individual whose returns are exactly at the population mean (left panel), and one standard deviation above the mean (right panel). The horizontal axis is multiples of mean high school wage. The distributions are determined by the largest wage of a high school job, a some-college job—both of which would lead to underemployment—and a college job.

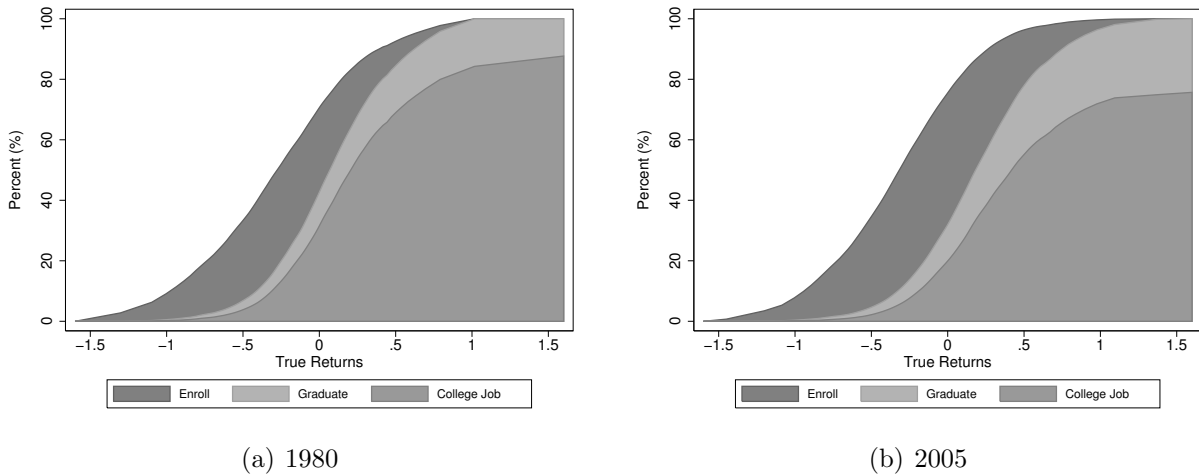


Fig. 10: Distribution of Education Choices

Implied educational attainment choices implied by calibrated parameters. For each level of $\log z$, unknown to the individual, we plot the fraction of individuals who enroll, graduate, and obtain a college job.

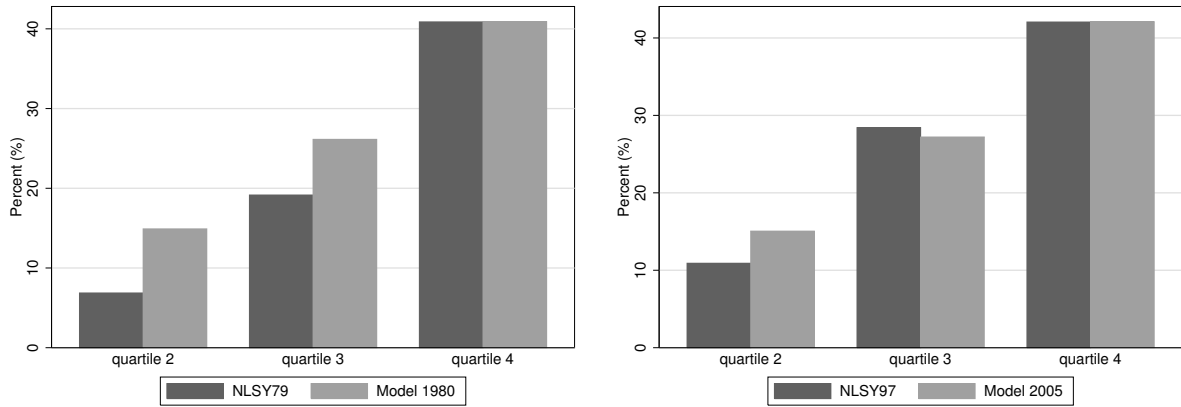


Fig. 11: Enrollment Rates by Family Income, Data and Model

Enrollment rates in the NLSY and the model by family income quartiles. Family income in the model is initial assets a_1 . The model targets the 1980 and 2005 CPS, and the enrollment rates differ between the CPS and NLSY. For this reason, we show the enrollment rates of the second to fourth quartiles minus that of the first quartile.

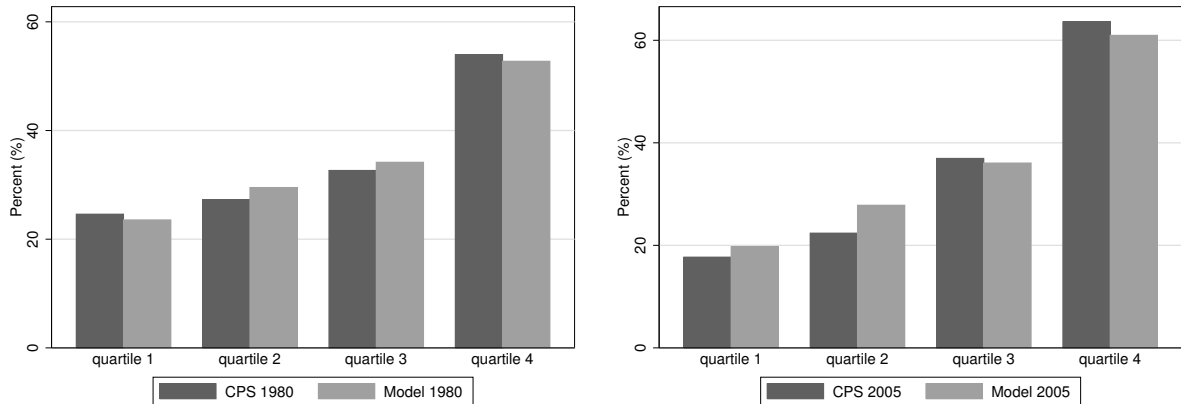


Fig. 12: College Attainment by Own Earnings, Data and Model

Fraction of college graduates in the CPS and the model by own earnings quartiles. Earnings in the CPS are controlled for age. Earnings in the model are lifetime earnings.