Review of Economic Dynamics 31 (2019) 224–244

Contents lists available at ScienceDirect

Review of Economic Dynamics

www.elsevier.com/locate/red

Factors affecting college attainment and student ability in the U.S. since 1900 [☆]

Kevin Donovan [a,*], Christopher Herrington [b,1]

[a] Yale University, United States of America
[b] Virginia Commonwealth University, United States of America

ARTICLE INFO

Article history:
Received 11 January 2017
Received in revised form 6 July 2018
Available online 25 July 2018

JEL classification:
I2
J24
N31

Keywords:
College attainment
Student ability
Borrowing constraints
College costs
Education premia

ABSTRACT

We develop a dynamic lifecycle model to study long-run changes in college completion and the relative ability of college versus non-college students in the early twentieth century. The model is disciplined in part by constructing a historical time series on real college costs from printed government documents dating to 1916. The model captures nearly all of the increase in attainment and ability sorting between college and non-college individuals between the 1900 to 1950 birth cohorts. Time variation in college costs, the college earnings premium, and the precision of ability signals all play a critical role for explaining different data moments and time periods, primarily through their interaction with binding borrowing constraints. Our quantitative results imply that attainment is broadly driven by the interaction of changing real college costs and the rising earnings premium, while ability sorting is driven by the earnings premium and increasing precision of ability signals.

1. Introduction

During the twentieth century, higher education expanded dramatically in the United States. As shown in Fig. 1a, bachelor's degree completion as a share of the 23-year-old population increased from less than four percent to more than 30 percent between the 1900 and 1972 birth cohorts. Concurrent with the increase in college attainment, the gap in measured cognitive ability widened substantially between individuals who completed at least some college and those whose formal education ended with high school (i.e., "non-college" individuals). This pattern is shown in Fig. 1b, which plots the difference between the average ability percentile of college and non-college individuals.[2] The average college student born

[☆] The authors are grateful to Diego Restuccia and two anonymous referees for numerous insightful comments and suggestions. Thanks also to Eleanor Dillon, Todd Schoellman, Gustavo Ventura, and participants at Arizona State, the University of Richmond, Midwest Macro Meetings, and the AEA Annual Meeting for helpful comments. The usual disclaimers apply.

[*] Corresponding author at: Yale School of Management, 165 Whitney Ave., New Haven, CT 06511, United States of America.
E-mail addresses: kevin.donovan@yale.edu (K. Donovan), herringtoncm@vcu.edu (C. Herrington).

[1] 301 W. Main St, Richmond, VA 23284, United States of America.

[2] Fig. 1b is the percentage point difference in average ability between college and non-college individuals. Following Hendricks and Schoellman (2014) we define non-college individuals as those who have graduated high school but not completed one year of college. That is, they either never attended college or left during the first year. We maintain this definition throughout the paper.

https://doi.org/10.1016/j.red.2018.07.003
1094-2025/

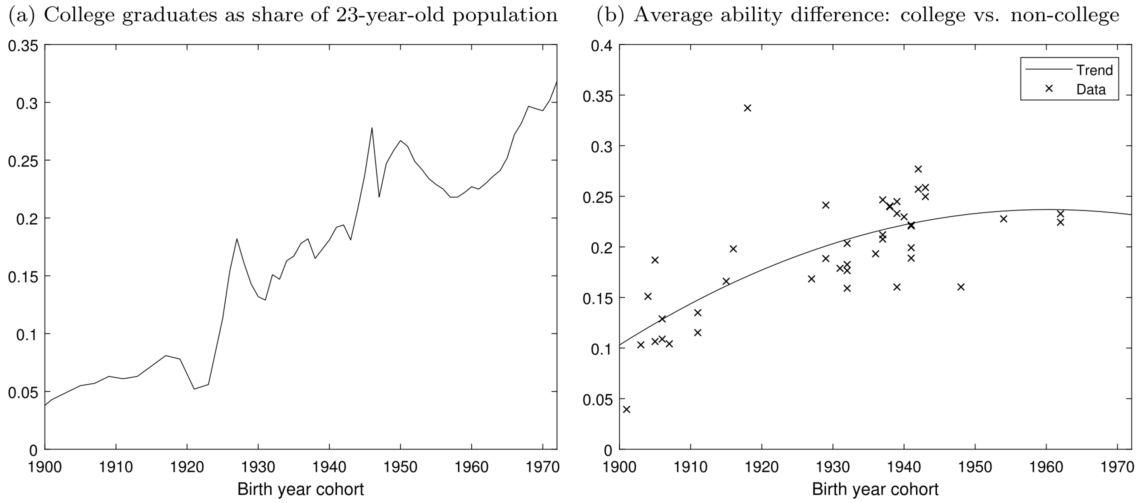

Fig. 1. College completion and ability sorting among U.S. birth cohorts 1900–1972.

around 1900 had measured cognitive ability about 10 percentage points above the average non-college individual, and this gap more than doubled by the 1940s birth cohorts.[3]

We think it important that theories of educational attainment be able to jointly confront both trends for two reasons. First, both result from the decisions of high school graduates making college attendance decisions. Thus, underlying economic factors affecting one margin potentially impact the other. Second, the ability of students potentially affects college completion through its impact on dropout risk. While much research has focused on post-World War II policy in accounting for these trends, less is known about the factors driving attainment and sorting among pre-World War II cohorts.[4] This is in spite of the fact that these cohorts account for the vast majority of attainment growth and ability sorting during the 1900s (Fig. 1).

The goal of this paper is to evaluate the causes of these empirical trends in a historical context and study what – if any – relationship there is between the factors driving college attainment and selection by ability. This task is complicated by the various changes in the aggregate economy and education sector over the early 20th century. We combat this issue by developing a quantitative theory of educational attainment disciplined using newly digitized historical data. Specifically, we develop a lifecycle model populated by an exogenous number of male and female high school graduates in each birth cohort. Endowed with financial assets and an ability to complete college – both of which are heterogeneous – these individuals immediately decide whether to enter college or go to work. Moreover, ability is not perfectly known. Consistent with evidence in Cunha et al. (2005), individuals observe only a noisy signal about true ability, which introduces a measure of uncertainty when combined with the limited borrowing opportunities for college. These features allow us to quantitatively compare a number of potential explanations within the framework of the model.

Naturally, our model predictions rely heavily on the costs and benefits associated with attending college, and we therefore take care to measure them properly. On the cost side, we construct out-of-pocket costs for tuition and fees by combining data from a series of printed government documents dating back to 1916. This data work allows us to feed in realistic time series data for college costs, instead of relying on assumptions such as a constant tuition growth rate. Furthermore, we construct the opportunity cost of attending college, along with the benefit of attaining a degree, by computing detailed wage profiles and dropout risk by ability. Wage profiles are estimated from the 1940 to 2000 U.S. Censuses and 2006–2010 American Community Surveys. The wage profiles are allowed to depend on sex, age, and education to accurately capture the changing education earnings premia in the United States. These time series are exogenously fed into the model to capture the relevant tradeoffs faced by high school graduates when choosing to enter college.

In our framework students can pay for college using a combination of endowed financial assets (i.e., parental transfers) and borrowing. Yet despite the historical data we do have, estimating how these features vary over time is beyond what we have available. We therefore set parameters for these as follows. First, we use more recent and reliable micro data to estimate a flexible joint distribution of ability and parental transfers by combining data from the 1979 National Longitudinal

[3] Throughout the paper we refer to birth cohorts whenever we use the term “cohort” unless otherwise specified.

[4] For example, Lochner and Monge-Naranjo (2011) emphasize student loan policies as the driving force behind the post-World War II increase in both ability sorting and college attainment. A notable exception to this post-World War II focus is Kaboski (2005), who studies the impact of earnings premia and college costs on college attainment during the early twentieth century. That paper lacks the detailed college cost measures we utilize here, among other differences.

Survey of Youth (NLSY) and High School & Beyond Survey (HSB).[5] Second, we assume that the borrowing constraint is a constant fraction of expected lifetime income. We choose this fraction to match attainment in the initial 1900 birth cohort and hold it fixed across all cohorts; however, the amount of borrowing available changes with both the level of wages and the premia afforded to those who complete more than high school education. Moreover, because ability determines the likelihood of completing college and thus expected earnings, individuals of different ability will have different borrowing limits as well. In addition, we model the introduction of Federal student loans for later cohorts, which can offer expanded borrowing opportunities for some individuals.

With the calibrated model in hand, we assess its capacity to replicate college attainment and ability changes across cohorts. Our benchmark quantitative results show that it can. Between the 1900 and 1950 birth cohorts, the model predicts annualized growth in college attainment of 3.77 percent, compared to 3.98 percent in the data. Similar results hold when we consider other sub-periods. In terms of ability sorting, the model predicts that the average college attendee in the 1900 birth cohort has IQ 10 percentage points higher than the average non-college individual. This gap more than doubles by the 1950 cohort, just as in the data. Thus, the model predicts almost the entire increase in college attainment from the 1900 to 1950 cohorts while also matching the increased ability sorting over the same period. When we consider post-1950 cohorts, we find that it matches the overall trend between cohorts born 1950 to 1972, predicting an annualized growth rate of 1.2 percent, compared to 0.80 percent in the data. Indeed, over the entire 1900 to 1972 birth cohort, the model matches the annual growth in college attainment almost exactly, predicting a 2.98 percent annual growth rate, compared to 3.00 in the data, even though the calibration only targets college attainment for the initial cohort. We note, however, that as in much of the literature, the model struggles to match the time series variation within the 1950 to 1972 birth cohorts (Card and Lemieux, 2001a), so these post-1950 results should be interpreted with more caution.[6]

Since the model matches the aggregate trends well, we next investigate the underlying causes of our results. We find that time series variation in earnings premia, college costs, and ability signal variance interact critically with binding borrowing constraints to generate our results. To start, note that purely from a lifetime income perspective, college costs are sufficiently low and earnings premia sufficiently high that *all* high school graduates would like to attend college. Put differently, borrowing constraints are critical for keeping some high school graduates out of college.

These binding constraints then interact with time varying college costs and earnings premia to generate our results. First, the low earnings premium we observe for pre-1930 cohorts endogenously tightens the borrowing constraint through its effect on expected income, implying fewer individuals can fund college education. The rising premium for post-1930 cohorts begins to loosen borrowing restrictions, thus increasing college attainment. A similar result emerges for costs – relative to income, college costs spike for 1910 to 1920 cohorts, drop sharply between the 1920 and 1930 cohorts, then rise steadily after that. Declining costs allow more people to fund college education and help generate the large increase in attainment among the 1920 to 1930 cohorts. The combination of the two features are critical. Without these cost fluctuations, we would first overshoot attainment for 1910–1920 cohorts and then undershoot post-1920. Without the rise in earnings premia we would undershoot attainment for the 1930–1950 cohorts. The combination of rising costs and rising earnings premia jointly generate the steady increase in attainment seen in the data for post-1930 cohorts.

Interestingly, however, costs play relatively little role in understanding the rising ability gap over time. To see why, consider the type of student who changes her attendance decision in response to a decline in college costs. Intuitively, only marginal ability individuals switch because high ability students can already borrow sufficiently against their own future income to fund college. That is, they borrow against the fact that they are more likely to graduate and reap higher wages. The marginal student is below average compared to the existing set of college students, but above average compared to the set of non-college individuals. When these marginal students switch from non-college to college, the ability gap is little affected because the average ability of both college and non-college individuals decreases. The same logic applies if ability is imperfectly observed, as in our model, but additional noise in the college entry decision further tempers the quantitative impact of costs on ability sorting.

The college earnings premium, on the other hand, plays an important quantitative role in shrinking the ability gap among the earliest cohorts. Our counterfactual exercise holding the earnings premium constant across cohorts increases the ability gap by 65 percent for the 1900 cohort and 33 percent for the 1920 cohort, when the actual earnings premia were substantially lower. Because the low earnings premia generate tight borrowing constraints, a number of high ability students with low assets do not attend college. As the earnings premia rise, a larger fraction of all individuals attend college, implying that the group of non-college individuals becomes increasingly comprised of first-year college dropouts. These dropouts are negatively selected on ability, thus inducing a larger ability gap as the returns to college increase.

Finally, the increasing precision of ability signals over time plays a critical role in understanding sorting. We attribute this to rapid growth of standardized testing during the first half of the century, which gave students more precise information

[5] We allow transfers to grow over time with income. The distribution around average transfers, however, is fixed by our procedure. See the calibration in Section 4 for more details.

[6] Specifically, attainment declines among 1950s cohorts before a substantial rebound among 1960s birth cohorts. Our model cannot match this decline due to the strong growth in the college earnings premia faced by these cohorts, a well-known issue in the literature (Card and Lemieux, 2001a). We highlight this later when assessing the importance of earnings premia for understanding our results. We also perform two additional experiments in Appendix D which can close the gap between model and data by more than half. Bound et al. (2010) and Castro and Coen-Pirani (2016) also investigate college attainment for these cohorts in more detail.

about their own ability relative to peers in their cohort. If early cohorts had ability signals as precise as later cohorts, our model would instead generate a decline in the ability gap over time. Again, the result rests critically on individuals' ability to finance college. With high signal variance, the set of people with perceived high ability are not necessarily those with truly high ability. These misinformed individuals internalize their expected loose borrowing constraints due to (perceived) high ability, and thus choose to enter college. As the signal gets more precise, true ability becomes more aligned with perceived ability, and the truly high ability students begin to sort into college. Thus, the precision of the signal shifts only the *type* of people who go to college, with little change in the overall number.

Taken together, our results show the intricate balance of several times varying forces are required to match the college attainment and ability sorting of cohorts in the early twentieth century.

1.1. Related literature

This paper is primarily related to the literature on college attainment in the United States over time, which includes recent work by Garriga and Keightley (2007), Restuccia and Vandenbroucke (2013), Hendricks and Leukhina (2018), and Hendricks et al. (2018). These papers are complemented by other work, including Bound and Turner (2002) and Card and Lemieux (2001b), which examines the effects of policy changes like the G.I. Bill and Vietnam war draft during particular points in history. More closely related to our work, Castro and Coen-Pirani (2016) ask whether educational attainment of the 1932–1972 cohorts can be explained by changes in the college earnings premia, tuition, education quality, and cohort average learning ability. Their complete markets model underpredicts college attainment for pre-1950 cohorts, while the combination of limited borrowing with our college tuition time series allows us to match it quite well. Hendricks and Schoellman (2014) also study early 1900 college attainment and ability sorting, but take college completion and student ability as given in order to understand the changes in the college earnings premium in a complete markets model. By contrast, we seek to understand the economic factors that affected college completion and average student ability.

As it relates to post-1950 cohorts, the aforementioned Castro and Coen-Pirani (2016) show that cross-cohort variation in learning ability can alleviate the over-prediction of college attainment. Keller (2014), on the other hand, develops a model of college attendance and quality choice, and points to the slowdown of human capital rental rates as the cause of the post-1950 slowdown in attainment. Lochner and Monge-Naranjo (2011) emphasize student loan policies with limited commitment as the driving force behind post-World War II student ability sorting.

A closely related strand of literature has also focused on the co-related roles of family background and borrowing constraints for college decisions. Keane and Wolpin (2001) finds that altering parental transfers can have significant effects on the overall schooling attainment of children when borrowing constraints are present; however, extending loans to ease borrowing constraints has little impact of college attendance decisions. Ionescu (2009) emphasizes initial human capital and learning ability of students as the primary determinants of college entry rather than family wealth or income during college. Belley and Lochner (2007) show that the role of family background has become increasingly important in recent decades, and that borrowing constraints are a crucial component for explaining this change. Finally, Athreya and Eberly (2016) show that college completion risk can dampen the expected return to college, and that this effect is magnified for low-wealth students who must borrow to finance the costs of college. Collectively these results inform our decision to include borrowing constraints, heterogeneous parental transfers, and college completion risk.

2. Accounting for college attainment

We begin with an accounting exercise to identify the key channels that affect college completion across cohorts. Our measure of college attainment for cohort t is the share of twenty-three year olds with a college degree.[7] This measure can be decomposed into three components

$$\frac{N_t^{grad}}{N_t^{23}} = \underbrace{\left(\frac{N_t^{HS}}{N_t^{23}}\right)}_{\text{HS graduation rate}} \times \underbrace{\left(\frac{N_t^{enroll}}{N_t^{HS}}\right)}_{\text{college enrollment rate}} \times \underbrace{\left(\frac{N_t^{grad}}{N_t^{enroll}}\right)}_{\text{college graduation rate}} . \tag{2.1}$$

N^{HS}, N^{enroll}, and N^{grad} are the number of individuals born in year t who complete high school, enroll in college, and graduate college. The first ratio on the right hand side – the HS graduation rate – is the fraction of all twenty-three year olds who have graduated high school.[8] The latter two are the fraction of high school graduates who enroll in college (the college enrollment rate) and the fraction of college attendees who graduate college (the college graduation rate). We will

[7] While educational attainment is often measured later in life to capture those who complete college at older ages, we prefer this series for two reasons. First, to our knowledge it is the only measure of college completion with consistent time series data for birth cohorts back to 1900. Second, our model is not constructed to evaluate college enrollment decisions of older students who: (i) are generally less financially-dependent upon parents when paying for education; (ii) face different opportunity costs of school after having been in the workforce for some time; and (iii) may anticipate different return on investment in education due to later-life completion.

[8] We also include GED recipients for high school graduating classes starting in 1971, using data from NCES.

Table 1
Growth rate decomposition.

Birth years	Log differences			Annualized percent change		
	γ^{col}	γ^{hs}	γ^{endog}	γ^{col}	γ^{hs}	γ^{endog}
1900–1972	2.13	1.59	0.54	3.00%	2.23%	0.76%
1900–1920	0.54	1.13	-0.59	2.72%	5.80%	-3.08%
1920–1950	1.41	0.50	0.91	4.82%	1.70%	3.13%
1950–1972	0.18	-0.04	0.22	0.80%	-0.19%	0.99%

refer to these ratios as such throughout the paper. Using identity (2.1), the growth rate of college attainment between cohorts born at t and t' is approximated by

$$\gamma^{col}_{t',t} := \log\left(\frac{N^{grad}_{t'}}{N^{23}_{t'}}\right) - \log\left(\frac{N^{grad}_{t}}{N^{23}_{t}}\right) = \gamma^{hs}_{t',t} + \gamma^{enroll}_{t',t} + \gamma^{grad}_{t',t}, \tag{2.2}$$

where γ^{hs}, γ^{enroll}, and γ^{grad} are the log differences of the three ratios on the right hand side of (2.1). Unfortunately we do not have sufficient data back to 1900 to separately compute γ^{enroll} and γ^{grad}. However, anticipating the model somewhat, both γ^{enroll} and γ^{grad} will be endogenous objects in our model, while γ^{hs} will be exogenous. We therefore rewrite equation (2.2) as

$$\gamma^{col}_{t',t} = \gamma^{hs}_{t',t} + \gamma^{endog}_{t,t'} \tag{2.3}$$

where $\gamma^{endog} := \gamma^{enroll} + \gamma^{grad}$ is the growth rate of factors endogenous to our model, and therefore potentially affected by time series variation we consider.[9] Since we have data on college attainment and high school graduation, the endogenous factors are defined as the residual that allows identity (2.3) to hold with equality. Table 1 decomposes γ^{col} for the entire time period (1900–1972 birth cohorts) and three sub-samples. For ease of interpretation, we also convert the log differences into annualized percent changes over each sample period.

Over the entire 1900 to 1972 time period, college attainment increased at an annual rate of three percent. From a growth accounting perspective, nearly three-fourths of this would be attributable to the increase in high school completion, γ^{hs}, consistent with Goldin and Katz (2010). This result, however, masks large differences across sub-samples of the overall time period. For cohorts born 1900 to 1920 college attainment increased at 2.72 percent annually, but the sources of this growth were very different. These cohorts experienced high school graduation rate increases of 5.8 percent annually, but a substantial decrease in endogenous factors (-3.08 percent annually). Yet for the 1920 to 1950 birth cohorts, which account for the bulk of college attainment growth, almost two-thirds of the increase in attainment is accounted for by endogenous factors (i.e., 3.13 of the 4.82 percent annualized increase). Finally, we note that this decomposition also captures the well-documented slowdowns in both high school and college attainment for cohorts born after 1950. High school graduation rates fell by 0.19% annually, and growth in endogenous factors was only one percent per year. Combined, these resulted in college attainment growth of only 0.8 percent annually for the 1950 to 1972 cohorts.

These results are consistent with other direct evidence over this time period as well. Snyder (1993) includes a measure of college completion conditional on high school graduation, defined as

$$\frac{\text{Number of college degrees awarded at year } t+4}{\text{Number of high school graduates at year } t}.$$

This captures roughly the same idea as the ratio N^{grad}/N^{HS} (i.e. the endogenous factors), though it requires assumptions on time-to-completion for the two to be exactly equivalent. Nevertheless, Fig. 2 shows that this series follows the same path uncovered in Table 1, with a substantial decline for cohorts born in the early 1900s and then sustained growth afterward.

Taken together, this evidence suggest that increases in high school graduation rates cannot be relied upon as the sole factor driving subsequent increases in college completion. Rather, it is vital to understand the factors affecting college enrollment decisions and completion rates for a full understanding of college attainment trends during the twentieth century. The remainder of this paper will develop and calibrate a model to quantitatively assess the main contributors to this growth.

[9] That is, our time series data allows us only to measure

$$\frac{N^{grad}_t}{N^{HS}_t} = \frac{N^{enroll}_t}{N^{HS}_t} \times \frac{N^{grad}_t}{N^{enroll}_t},$$

but not the two right-hand side ratios separately.

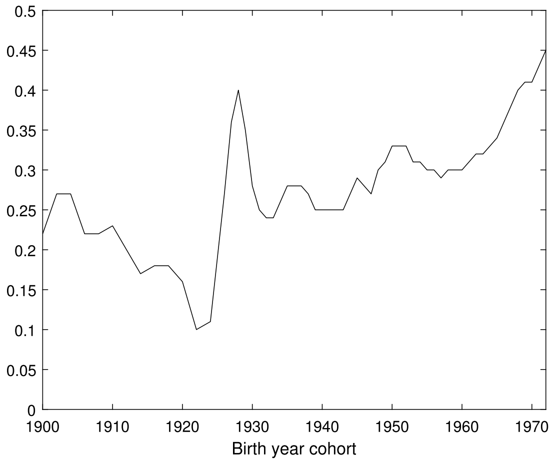

Fig. 2. College completion rate conditional on HS graduation.

3. Model

We now develop a lifecycle model to investigate the causes of increased college attainment and increased ability sorting over time. The relevant features include borrowing limits, uncertain ability, risky completion of college education, and education premia that vary across cohorts. The notation introduced in this section is summarized in Table 6 of the Appendix.

Demographics and preferences Time in the model is discrete, and a model period is one year. Each period, N_{mt} males and N_{ft} females are born, each of whom lives for a total of T periods. Let $a = 1, 2, \ldots, T$ denote age. Each individual maximizes expected lifetime consumption, given by

$$\mathbb{E}_0 \sum_{a=1}^{T} \beta^{a-1} \left(\frac{c_a^{1-\sigma} - 1}{1-\sigma} \right).$$

Endowments and signals Individuals are ex-ante heterogeneous along three dimensions: their sex, m or f, initial asset endowment k_0, and ability to complete college, denoted α. The probability that an individual completes his or her current year of college is given by $\pi(\alpha)$, where $\pi'(\cdot) > 0$. Ability and initial asset holdings are drawn from a joint distribution with cumulative distribution function $F(\alpha, k_0)$.

While sex and asset endowments are perfectly observable, ability α is not. Instead, each individual receives a signal $\theta = \alpha + \varepsilon$ at the beginning of life. The error term is $\varepsilon \sim N(0, \sigma^2_{\varepsilon,t})$, so signal precision may vary across cohorts. Note that both the signal θ and asset endowment k_0 provide information about underlying ability, as assets and ability are potentially correlated. We therefore denote $\nu = (k_0, \theta)$ as the information an individual has about his or her true ability. After the initial college enrollment decision, ability α becomes publicly observable.

Education decisions The population we consider is high school graduates, so that birth in this model translates to a high school graduation in the real world. At birth, every individual decides whether or not to enroll in college, given sex, asset endowment k_0, and ability signal θ. This is the only time this decision can be made. Once enrolled in college, individuals can exit college by either graduating or by failing out with annual probability $1 - \pi(\alpha)$. After failure, individuals enter the labor force and may not re-enroll, consistent with the finality of dropout decisions discussed in Card and Lemieux (2001a). Graduating college requires C years of full-time education at a cost of λ_t per year. We assume $\pi(\alpha) = 0$ for at any age $a > C$, so that no one spends more than C years in college.[10] If an individual decides to not enter college, he or she immediately enters the labor market and begins to work.

Labor market We adopt the common assumption that individuals of different ages, a, sex s, and education, e, are different inputs into a constant returns to scale production function that requires only labor. Therefore, wages depend on age a, year

[10] Technically, this implies that continuation likelihood π should be written as a function of age and ability. With a slight abuse of notation, that function is

$$\pi(a, \alpha) = \begin{cases} \pi(\alpha) & \text{if } a \le C \\ 0 & \text{if } a > C. \end{cases}$$

We write $\pi(\alpha)$ throughout to economize on some notation, with the understanding that everyone has either graduated or dropped out by age C.

t, sex $s \in \{f, m\}$, education level $e \in \{0, 1, \dots, C\}$, denoted by the function $w(a, t, e, s)$.[11] Individuals have perfect foresight and can forecast their time series of earnings under different education choices.

Savings Each individual can borrow and save at an exogenous interest rate r. Individuals cannot die in debt, so $k_{T+1} = 0$. During life, we utilize two forms of borrowing limits in the model. First, all cohorts share a common version of the natural borrowing limit which allows individuals to borrow a fraction $\gamma \in [0, 1]$ of expected discounted future earnings. Second, cohorts who make college decisions after the introduction of Federal student loans may borrow up to the Federal limit, $\tilde{k}_t$, if it exceeds the natural limit. Formally, individuals may not borrow beyond the threshold $\bar{k}$, where

$$\bar{k} = \begin{cases} -\max\{\tilde{k}_t, \gamma \cdot \mathbb{E} \sum_{n=a+1}^{n=T} \frac{w(n,t,e,s)}{(1+r)^{n-a}}\} & \text{if } a < T \\ 0 & \text{if } a = T \end{cases}$$

The last term is required to guarantee that individuals do not die with any debt. Note that both the expectations operator and wage can depend on a number of factors, including ability α, age a, year t, education e, and sex s. Therefore, the borrowing constraint will be written as the function $\bar{k}(\alpha, a, t, e, s)$. In a slight abuse of notation, we will write $\bar{k}(a, t, e, s)$ when the borrowing constraint does not depend on ability α, as is the case once an individual finishes college.[12]

3.1. Timing and recursive problem

At the beginning of year t, N_{mt} men and N_{ft} women are born at age $a = 1$. Again, each individual is initially endowed with assets k_0, sex s, unobserved ability α, and a signal θ of true ability. Immediately, each individual decides whether or not to enroll in college. If she enrolls in college, true ability is realized, and the individual proceeds through college. In the case of leaving college due to failure or graduation, she proceeds to the labor market and works for the remainder of her life.[13] Individuals who do not enroll in college proceed directly to the labor market, where they receive the wage associated with age a, education $e = 0$, and sex s.

Recursive problem for worker For individuals currently not enrolled in college, their ability is irrelevant for their decision problem. Therefore, the value of entering year t at age a with assets k, years of college education e, and sex $s \in \{f, m\}$ is:

$$\begin{aligned} v^w_{a,t}(k, e, s) &= \max_{k'} u(c) + \beta v^w_{a+1,t+1}(k', e, s) \\ s.t. \quad c + k' &= (1 + r)k + w(a, t, e, s) \\ k' &\geq \bar{k}(a, t, e, s) \end{aligned}$$

Recursive problem for college student If instead an individual is currently enrolled in college, she has already completed e years of his education and must pay λ_t in college costs for the current year. The probability that she passes and remains enrolled the next year, however, depends on her ability α.

The value of being enrolled in college at year t at age a, with assets k, ability α, e years of education completed, and sex $s \in \{f, m\}$ is:

$$\begin{aligned} v^c_{a,t}(k, \alpha, e, s) &= \max_{k'} u(c) + \beta \Big[\pi(\alpha) v^c_{a+1,t+1}(k', \alpha, e+1, s) + (1 - \pi(\alpha)) v^w_{a+1,t+1}(k', e, s)\Big] \\ s.t. \quad c + k' - \lambda_t &= (1 + r)k \\ k' &\geq \bar{k}(\alpha, a, t, e, s) \end{aligned}$$

The college enrollment decision Given the value of being enrolled in college and working, we can define the educational decision rule at the beginning of life. Recall that at this point, α is unknown, but each individual receives information $\nu = (k_0, \theta)$. Each individual then constructs beliefs over possible ability levels by using Bayes' Rule. Let $G(\alpha; k_0, \theta)$ be the cumulative distribution function of beliefs over ability levels. Given all this, an individual of sex s born in year t with assets k_0 and signal θ enters college if and only if the expected value of entering college is higher than the value of entering the workforce. The college decision is therefore represented by the value function

[11] Note that while ability α has no direct effect on realized wages, it does affect expected wages because higher ability students are more likely to graduate college and earn higher wages. We discuss this further and relax the assumption in an experiment in Appendix D.

[12] The assumption that ability uncertainty is immediately resolved after the college choice implies that the expectation over the borrowing constraint does not directly include uncertainty over α, but only uncertainty about graduating college.

[13] We adopt "she" and "her" for simplicity, with the reminder that both sexes exist in the model.

$$v_t(k_0,\theta,s)=\max\left\{\int\limits_{\alpha} v^c_{1,t}(k_0,\alpha,1,s)G(d\alpha;k_0,\theta),\, v^w_{1,t}(k_0,0,s)\right\}$$

with associated decision rule $\phi_t(k_0,\theta,s)=1$ if the individual decides to enroll, and $\phi_t(k_0,\theta,s)=0$ otherwise.

We next turn to quantifying this model to assess the channels underlying changes in college attainment over time.

4. Calibration

Our goal is to use the model to quantitatively assess how several economic features have affected college attainment and ability sorting over time. Toward that end we first set some model parameters to standard values. We set the length of working life at $T=48$, implying that individuals born into the model at age 18 would retire at age 65. The number of periods required to complete college is $C=4$, so individuals in the model have post-secondary education $e\in\{0,1,2,3,4\}$. The real interest rate is set to $r=0.04$ in all periods, and the discount rate is $\beta=0.96$, a standard value in models with annual periods.

We calibrate the rest of the model in a series of steps, which are discussed in detail below. First, we utilize historical evidence to construct time series for high school graduates N_{mt} and N_{ft}, college costs λ_t, life-cycle wage profiles $w(a,t,e,s)$, and Federal student loan limits, $\tilde{k}_t$. These are fed exogenously into the model. Next, we utilize modern micro data to estimate the probability of passing college $\pi(\alpha)$ and the joint distribution of initial assets and ability $F(\alpha,k_0)$. These data cover cohorts born during the early 1960s, which are near the end of the cohorts we consider. Lacking earlier data, we hold these parameters constant across all cohorts. Finally, we are left to set the borrowing constraint and ability signal variance. We choose these parameters while solving the model in order to match important moments in the data.

4.1. Exogenous time series inputs to the model

4.1.1. Number of high school graduates

As previously mentioned, N_{mt} males and N_{ft} females are "born" into the model each year, meaning they graduate high school and enter the model eligible to make college enrollment decisions. We take high school completion, and thus the population of potential college enrollees, as exogenous. The data on high school graduates are taken from the Historical Statistics of the United States Millennial Edition Online, and we use linear interpolation to supply missing values. The *Digest of Education Statistics*, published by the National Center for Education Statistics, also provides data on total GED recipients starting in 1971, and by age starting in 1974. Because these individuals are also eligible for college enrollment, we add the number of GED recipients under age 19 the population of high school graduates in each year. This adjustment affects cohorts from 1953 onward (i.e., the high school graduating classes from 1971 onward).[14]

4.1.2. College costs

Annual college costs per student, λ_t, are measured as the annual total tuition and fees paid by students to public universities divided by the number of full-time equivalent students attending those universities. The complete time series is constructed by splicing together data from historical print sources including the *Biennial Surveys of Education* (1916–1918 to 1956–1958 editions) and the *Digests of Education Statistics* (annually since 1962). Further details on the data are found in Appendix A.

Fig. 3 plots the series for real college costs, both in real dollars (Fig. 3a) and as a share of average income (Fig. 3b).[15] Note the dramatic fluctuations as a share of average income. As a share of average income, the first year of college for the 1920 cohort is more than double than of the 1900 cohort. Moreover, it falls from 35 percent to below 10 percent. This is a function of both a decline in the real cost of college combined with the economic rebound from the Great Depression. Cohorts born post-1930 experience a steady increase in college costs relative to income. Anticipating our results somewhat, these fluctuations turn out to play a critical role in our model's ability to match attainment patterns. In Section 6 we provide further discussion of the policies and economic changes that generate the fluctuations in relative costs.

Several aspects of this series are worth noting. First, by measuring revenue that comes directly from students we account for any discounts or aid that would not be reflected in the sticker price of tuition. This includes scholarships or aid provided by the university, as well as local, state, or federal subsidies, including the World War II G.I. Bill, which we discuss in more detail later. Second, we focus on public universities because the important decision margin for our purposes is *whether* to attend, not *where* to attend. The relevant cost for the marginal student choosing whether to attend is that for the public university. The question of *where* to attend, which may include higher cost or higher quality options like private universities, is beyond the scope of our analysis. Third, we intentionally exclude costs for room and board because we view these as consumption expenditures.[16]

[14] To impute the number of GED recipients under age 19 for 1971–1973, we assume the age distribution is the same as in 1974. We also assume that GED recipients are evenly split between male and female since the data are not disaggregated by sex.

[15] The cost of a four-year college degree is computed as the present discounted value of annual costs at age 18–21. Fig. 3a show this series adjusted to year 2000 dollars using the overall Consumer Price Index (CPI).

[16] The main results are robust to the inclusion of room and board, which we can also measure.

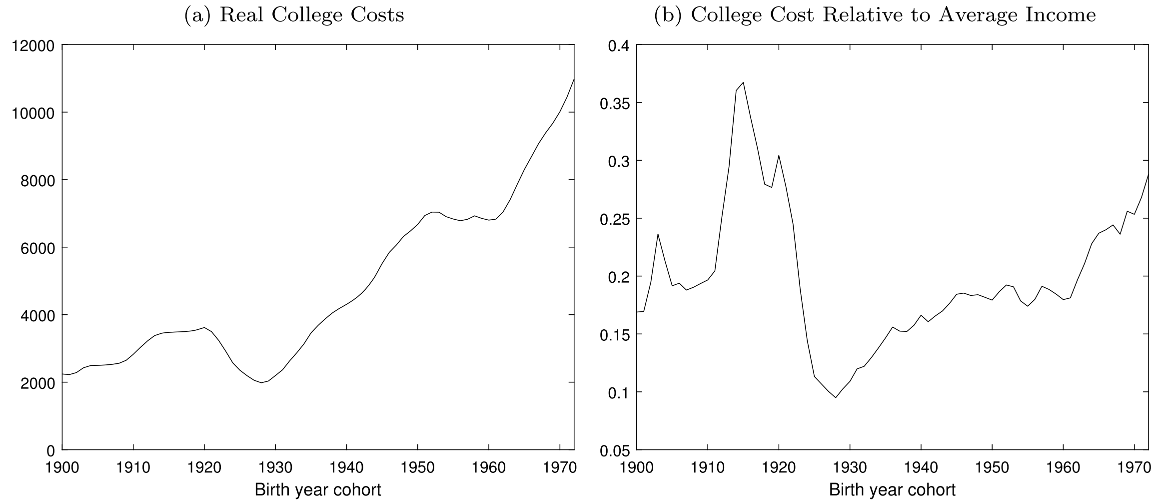

Fig. 3. College costs.

4.1.3. Life-cycle wage profiles

Life-cycle wage profiles $w(a,t,e,s)$ are estimated using decennial U.S. Census data from 1940 through 2000, along with American Community Survey (ACS) data from 2006–2010. Each ACS data set is a 1 percent sample of the U.S. population, so that when combined they constitute a 5 percent sample of the U.S. population, similar to a decennial census. The data are collected from the Integrated Public-Use Microdata Series (IPUMS) (Ruggles et al., 2010), and include wage and salary income, educational attainment, age, and sex. From age and education data we compute potential labor market experience, x, as age minus years of education minus six. We assume that wages can be drawn from one of three education categories corresponding to $e = 0$, $e \in \{1, \ldots, C-1\}$ and $e = C$ in the model. That is, individuals with zero years of completed college get the "high school" wage profile, those with at least one year of completed college but no degree get the "some college" wage profile, and those with a college degree get the "college" wage profile. We restrict the data to include the non-institutionalized population between ages 17 and 65 who report being in the labor force and estimate the following regression:

$$\log(w_i(a,t,e,s)) = \beta^0 + \beta^s + \beta^e + \sum_{j=1}^{4} \delta_j^e x^j + \sum_{j=1}^{2} \theta_j^e c^j \tag{4.1}$$

where $w_i(a,t,e,s)$ is the wage for individual i at age a in year t with education e, and sex s. In this formulation β^0 is a constant, β^s and β^e are sex and education fixed effects, $\sum_{j=1}^{4} \delta_j^e x^j$ is an education-specific experience quartic, and $\sum_{j=1}^{2} \theta_j^e c^j$ is an education-specific cohort quadratic. Note that in model notation, age $a = 1$ corresponds to 18 years old, so experience is $x \equiv (18 + a) - (e + 12) - 6$, and birth cohort is $c \equiv t - a - 18$. Our estimation captures the U-shaped pattern in the college premium found in previous papers, and as we will show, this plays a critical role in our quantitative results.[17]

Fig. 4 plots the college premium we estimate. For cohorts born 1900–1920, we observe a decline in the college premium of about 5 percentage points. The premium hits a minimum for cohorts born around 1920–1925 and increases thereafter. This timing is consistent with much of the literature, which finds the contemporaneous college premium at its lowest around 1950, when individuals from the 1920–1925 birth cohorts would have been 25–30 years old (Autor et al., 1998; Goldin and Katz, 2010; Restuccia and Vandenbroucke, 2013).

4.1.4. Federal student loan limits

The first generally available Federal student loans were implemented via the Higher Education Act of 1965, and became known later as the Stafford loan program.[18] Initially the government provided loans up to $1,500 per year during each year of undergraduate study. In the decades following the loan limits were raised several times. To set the Federal borrowing

[17] A few parts of this set-up are worth noting. First, our formulation takes a "cohort view" of the age-time-cohort problem, essentially assuming that there are only age and cohort effects, not direct time effects. Second, because the Census data begins in 1940, we do not observe earnings before this time. However, the experience quartics are estimated across all cohorts, so we can make out of sample predictions for individuals whose earnings are unobserved at some ages. Lastly, note also that this formulation allows wages differ across men and women, but not the college earnings premium. This is consistent with empirical evidence (Hubbard, 2011).

[18] While the National Defense Education Act of 1958 also included student loans, these were not generally available to all college students, so we do not include them here.

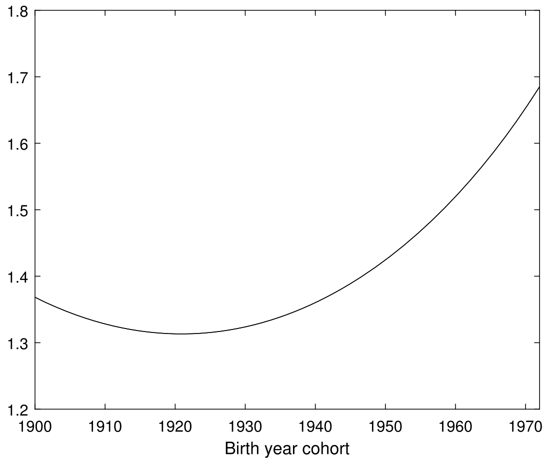

Fig. 4. Lifetime earnings ratio: college vs high school.

Table 2
Federal student loan limits.

Calendar years	Federal limit	Affected birth cohorts
1918–1966	$0	1900–1948
1967–1973	$6,000	1946–1955
1974–1977	$7,500	1953–1959
1978–1983	$10,000	1957–1965
1984–1993	$29,250	1963–1972
1993–1994	$35,125	1972

Table notes: The last column is the birth cohorts under study (1900–1972) that are subjected to the given federal limit at any time during their college career.

limit, $\tilde{k}_t$, we compute the combined annual subsidized and unsubsidized Stafford loan limits available for four years of college. Table 2 lists the years and the nominal loan limits we take as exogenous to the model. Note that our last birth cohort in 1972 attends college 1990–1994, so we do not include loan limit changes after that time.

4.2. Parameters estimated from modern data and fixed across cohorts

4.2.1. Probability of passing college

We next need to set the annual probability of passing college by ability, $\pi(\alpha)$. Our proxy for ability in the NLSY79 data is an individual's AFQT percentile score. Note that $\pi(\alpha)$ is a reduced form way to capture college non-completion for any reason, including failure and voluntary drop-out, so we first calculate the *cumulative* probability of college completion among individuals in each AFQT percentile bin. From this we calculate the annual probability equal to the cumulative probability raised to the one-fourth power. Thus, we assume there are four independent annual opportunities for failure, since the length of college is $C = 4$. Finally, we estimate the function $\pi(\alpha)$ by linear regression of the form $\pi(\alpha) = \beta_0 + \beta_1\alpha + \epsilon$, where α is the AFQT percentile. The estimated coefficients are $\beta_0 = 0.6198$ and $\beta_1 = 0.0034$, with $R^2 = 0.55$. Fig. 5 plots the likelihood of graduating college by AFQT percentiles in the NLSY79 data, and the model-implied likelihood of college graduation, that is, $\pi(\alpha)^4$.[19,20]

4.2.2. Joint distribution of assets and ability

The next step is to construct a realistic joint distribution of initial assets and ability. While the estimation strategy is detailed completely in Appendix B, we briefly outline it here.

We require marginal distributions of ability and parental transfers, along with any potential correlation between the two. However, to our knowledge, no single data source contains information on both a measure of innate ability and parental

[19] One concern with our procedure might be time-varying dropout risk for a given ability level, which we would not capture here. Existing historical evidence suggests college dropout rates have remained stable between the 1930s and 2000s, and we discuss this in more detail in Section 6 after laying out the quantitative results.

[20] Note that while $\pi(\alpha)$ is constant over college years, the failure rate declines with years of college because surviving students are higher ability on average. This is a function of the fact that $d\pi/d\alpha > 0$. Anticipating the results, we find dropout rates consistent with available evidence. For the 1962 cohort, our benchmark model predicts dropout rates of 20.0, 15.1, 11.5, and 8.9 percent over the four years of college. Hendricks and Leukhina (2017) report 16.6, 14.5, 8.3, and 4.9 for the same cohort using High School and Beyond data.

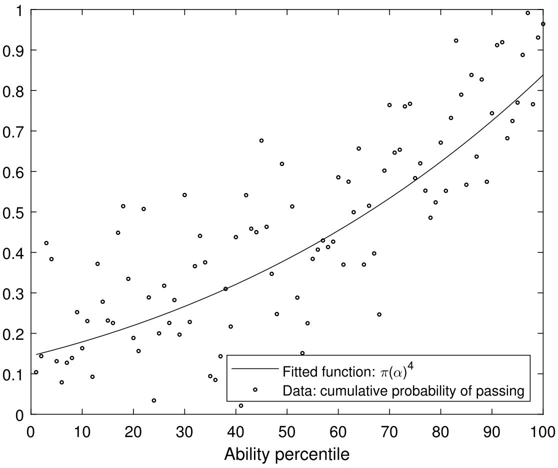

Fig. 5. Probability of graduating from college by AFQT percentile.

transfers during the time period we consider. Our strategy is therefore to merge information from two separate datasets: the NLSY79 and the National Center for Education Statistics' High School & Beyond (HSB) dataset. NLSY79 includes information about students' grades and ability (i.e., AFQT score), but no information on parental transfers. HSB, on the other hand, includes information about grades and transfers for high school seniors (17–18 years old, or age 0 in the model). We utilize the fact that both datasets include high school grades. We first estimate the empirical link between grades and ability using information in the NLSY. We then impose that same relationship in HSB data, thus allowing us to construct a measure of predicted ability in the HSB. Finally, we derive the joint distribution of ability and parental transfers within the HSB data.

Ability distribution We first estimate the grade-ability relationship in the NLSY. We assume GPA is lognormally distributed, then use a linear regression to link log GPA to ability (proxied in the NLSY by AFQT score). By our distributional assumption on GPA, ability is therefore normally distributed as $\alpha^{NLSY} \sim N(\beta_0 + \beta_1 \mu_{g,NLSY},\ \beta_1^2 \sigma_{g,NLSY}^2 + \sigma_\varepsilon^2)$, where β_0, β_1, and σ_ε are the estimates and error standard deviation from the linear regression, and $\mu_{g,NLSY}$ and $\sigma_{g,NLSY}$ are the associated parameters of the GPA distribution.

Next, we project the estimated relationship between grades and ability onto grade information in HSB. Again assuming that HSB GPA is lognormal, this implies

$$\alpha_{HSB} \sim N(\beta_0 + \beta_1 \mu_{g,HSB},\ \beta_1^2 \sigma_{g,HSB}^2 + \sigma_\varepsilon^2). \tag{4.2}$$

Note that the parameters β_0, β_1, and σ_ε are estimated from the NLSY, but $\mu_{g,HSB}$ and $\sigma_{g,HSB}$ must be estimated from the HSB grade data.

The estimation of $\mu_{g,HSB}$ is the last step of constructing the marginal ability distribution. HSB includes grade bins, not grades themselves, which requires us to estimate the distribution using masses in each discrete bin, not the underlying continuous distribution. Figure 12 in Appendix B shows that we match the distribution well with parameters $\mu_{g,HSB} = 4.404$ and $\sigma_{g,HSB} = 0.096$. Using these estimates in equation (4.2) gives us the implied marginal distribution of ability in the HSB dataset.

Transfer distribution The empirical distribution of parental transfers is best matched with a gamma distribution. The shape and scale parameters are chosen to minimize the sum of squared errors between that empirical c.d.f. of average transfers (normalized by the mean) and that of the estimated gamma distribution. The best fit parameters are a shape parameter of 0.24 and a scale parameter of 4.44. Figure 13 in Appendix B plots the estimated and empirical distributions together, and shows that we match the data well. Lastly, we assume that the average transfer increases at the same rate as average income in the United States.

Joint distribution of ability and transfers The last step is to compute the joint distribution of the two marginals we created above. We use a Frank copula to accomplish this, which takes the form

$$C(u, v) = \frac{-1}{\rho} \log\left(1 + \frac{(\exp(-\rho u) - 1)(\exp(-\rho v) - 1)}{\exp(-\rho) - 1}\right).$$

The parameter ρ governs the dependence of draws. Our joint distribution of α and k_0 can therefore be written as

$$H(\alpha, k_0) = C\,[F(\alpha), G(k_0)]$$

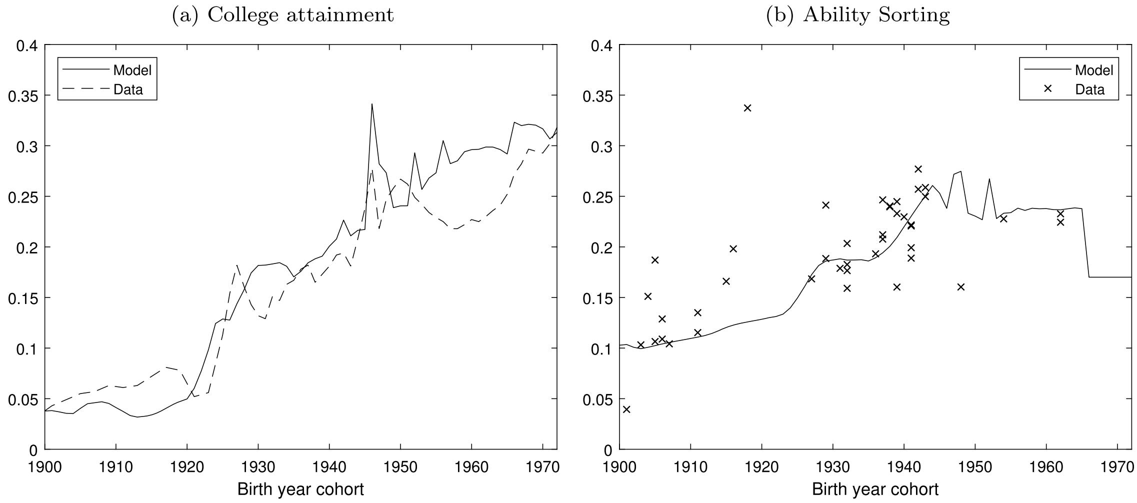

Fig. 6. Benchmark model results.

where F and G are the cumulative distribution functions of the normal and gamma respectively. We are left to calibrate ρ, which roughly implies a positive correlation between the two series when $\rho > 0$. Note, however, that while our procedure gives the marginal distribution of ability in HSB, it does not provide individual-level estimates of ability. We therefore use the implied relationship to draw individual realizations of ability for students in the HSB data. We then compute the Kendall rank coefficient and repeat this simulation 1000 times. The average Kendall rank coefficient is 0.04, implying that high school graduates with higher ability on average have higher initial asset holdings. This implies a copula parameter of $\rho = 0.35$.

4.3. Parameters set within the model

We are now left to set only the borrowing constraint γ and the variance of ability signals $\sigma_{\varepsilon,t}$. Conceptually, ability signals should be more precise (i.e., the variance should be smaller) when students have more and better information about their own ability and likelihood of passing college. For cohorts born post-World War II, standardized admissions tests like the Scholastic Aptitude Test (SAT) have provided an important signal of student ability and college-readiness. Beale (1970) reports that standardized testing grew throughout the first half of the 20th century, and by 1960 most colleges considered SAT scores "absolutely essential" to the admissions process. In fact, data from Hendricks et al. (2018) show that 1962 (i.e., the 1944 birth cohort) was the first year in which the number of standardized tests taken (SAT and ACT combined) exceeded the number of new college freshmen.

Given this evidence, we assume that $\sigma_{\varepsilon,t}$ has been constant for all cohorts born after World War II and that it declined linearly for all cohorts born prior to World War II. By imposing this structure, we need only choose the initial value of $\sigma_{\varepsilon,t}$ for the first birth cohort in 1900, and the constant value for all cohorts born in 1945 and later. We do so two steps. First, we jointly choose γ and the initial value of $\sigma_{\varepsilon,t}$ so that the model matches the college completion rate of 3.8 percent and average ability difference of 10 percentage points for the first birth cohort in 1900, as shown in Fig. 1. The values which match the data are $\gamma = 0.0536$ and $\sigma_\varepsilon = 4.64$. Holding γ fixed across cohorts, we then choose the constant value of $\sigma_{\varepsilon,t}$ for 1945–1972 cohorts to match the average ability difference of the 1962 birth cohort, which results in $\sigma_\varepsilon = 0.65$. This choice is made for consistency with the fact that we estimated $\pi(\alpha)$ and $F(\alpha, k_0)$ using data for cohorts born circa 1960–1965; however, we could also match other post-1945 cohorts without materially affecting the results.

5. Quantitative results

With the calibrated model in hand, we now simulate the model to study attainment and ability sorting. First, Section 5.1 lays out our benchmark model fit. We find that the model matches both attainment and sorting well for pre-1950 cohorts. We then turn to understanding how we generate our benchmark results in Section 5.2, highlighting critical time-varying factors (college costs, earnings premia, ability signal precision) and their interaction with borrowing constraints.

5.1. Benchmark model fit

We begin with our benchmark model fit for the time series of attainment and ability sorting. Overall, the model replicates well the long-run trends over much of the 20th century. Fig. 6 plots the overall time series from both model and data, and we discuss attainment and sorting in turn.

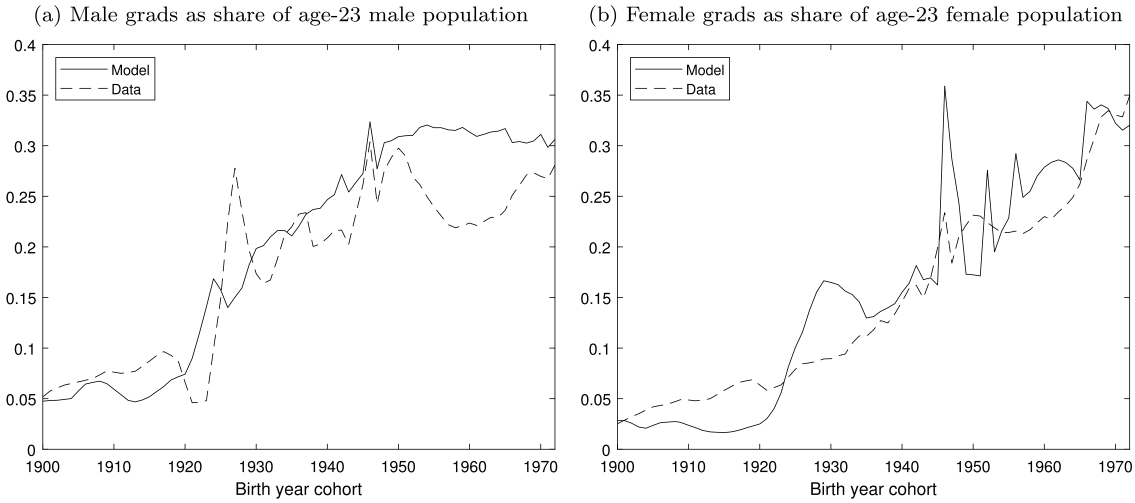

Fig. 7. College attainment by sex.

College attainment Fig. 6a plots the time series for college attainment in both the model and the data. The model accurately captures overall growth for cohorts 1900–1972. On average, the annual growth in attainment is 3.00 percent in the data, compared to 2.98 percent in the model. The model does miss the data, however, during some shorter intervals. Most notably, despite matching average growth for the 1950–1970 cohorts, the model misses the dramatic decline of attainment for cohorts in the 1950s and the rebound among 1960s cohorts.[21] For pre-1950 cohorts, the model generally replicates the data well, but fails to predict the slight increase in attainment for 1910–1920 cohorts. However, it tracks pre-1910 cohorts well, matches the sharp increase among 1920s cohorts, then again tracks the data closely through the late 1940s cohorts.[22] We discuss the 1910–1920 and 1950–1965 cohorts further in Appendix C, where we conduct two experiments that are capable of substantially closing the gaps between model and data.

Ability sorting Fig. 6b turns to ability sorting, where we plot the difference in average ability between college and non-college individuals.[23] The ability gap for birth cohorts in 1900 was about 10 percentage points, and this grew to almost 25 percentage points for cohorts born in the late 1940s and early 1950s. Our model matches this trend as well and is able to capture the entirety of the growth over this period.[24] In Appendix C, we compare the time series of average ability for college and non-college individuals separately and show that they each match their respective data counterparts well.

College attainment by sex In addition to matching aggregate college attainment, the model matches the long-run trends by sex, despite this not being targeted in the calibration. Fig. 7 plots college attainment as in Fig. 6a but for males and females separately. For males, the model tracks the data closely for 1900–1950 cohorts, net of the brief spike in attainment following World War II that shows up for the late 1920s birth cohorts. This is unsurprising since all individuals in our model face the same average cost of college; no one receives free tuition as was provided by the G.I. Bill. We extend the model to consider G.I. Bill subsidies and their effect on average costs in Appendix D.1. The model also generates a slowdown in college attainment for males after the 1950 birth cohorts, but not a decline of the magnitude seen in the data.

For females, the model predicts that college attainment continues rising for post-1950 birth cohorts, similar to what we see in the data. The introduction of Federal loans plays an important role in increasing female attendance for later birth cohorts by loosening borrowing constraints for females, whose lower expected lifetime earnings subjects them to relatively tighter natural borrowing limits. One place where the benchmark model deviates notably from the data is the over-prediction of female attainment growth during the 1920–1930 birth cohorts. This is caused by rapidly falling average college costs in the model, some of which is due to subsidies from the World War II G.I. Bill. We revisit this issue in Appendix D.1, as well, and show that much of this deviation can be corrected after accounting for the fact that G.I. Bill education subsidies were almost exclusively utilized by men.

Lastly, the model accurately predicts a decline in the male share of college graduates. In the data, men account for two-thirds of college graduates in the 1900 birth cohort, and this falls rather steadily until women captured the majority

[21] This is a common issue when trying to understand attainment for these cohorts (Card and Lemieux, 2001a), and is a function the strong growth in the college earnings premium faced by these cohorts. See Bound et al. (2010) and Castro and Coen-Pirani (2016) for detailed investigations of this phenomenon.

[22] In Appendix D we provide more details on the variation in graduation rates and enrollment rates over time.

[23] Specifically, the graph is the average IQ of a student who completes at least one year of college minus the average IQ of a student who completes high school but does not complete a year of college. This measure is used in Hendricks and Schoellman (2014) and Hendricks et al. (2018) as well.

[24] One could potentially argue that we under-predict the ability gap for cohorts in the mid-1910s, but unlike attainment, it is difficult to construct an exact time series of ability. We are therefore hesitant to put more year-over-year emphasis on this series and instead focus on the overall trend.

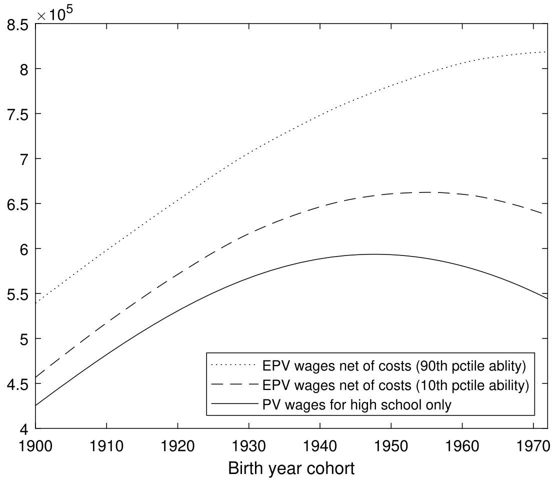

Fig. 8. Expected present value of income net of education costs.

beginning with cohorts born in the 1960s. For comparison, the 1900–1920 cohorts in our model average 66 percent male share of college graduates. Over the 1920–1950 cohorts this falls to 57 percent, and for 1950–1972 cohorts it falls further to 53 percent. We also find that females make up a majority of college graduates for all cohorts after 1966 in the model.

Having now established that the model matches the trends in both attainment and ability sorting, we next turn to understanding the model features that allow us to match these trends.

5.2. Factors generating attainment and sorting

We begin by asking which individuals would prefer to attempt college purely from the perspective of higher lifetime income. Put differently, we ask how, in the absence of any borrowing constraint, college cost and wage changes interact to change attendance. Fig. 8 plots the expected lifetime earnings net of college costs for two individuals: one in the tenth ability percentile and one in the ninetieth ability percentile, and compares it to the present value of lifetime income if either decides not to attempt college.[25]

As expected, Fig. 8 shows that higher ability individuals have higher lifetime earnings, as they are more likely to complete college and earn higher wages. More importantly, however, it also shows that low ability individuals would find it beneficial to attempt college. That is, for all ability levels and cohorts, the expected present value of lifetime income (net of college costs) is above the expected present value of not attempting college. The expected net lifetime earnings of a tenth ability percentile individual is 7 percent higher if he attempts college in the 1900 birth cohort, rising slightly to 8 percent in the 1920 cohort and 11 percent by the 1950 cohort.[26]

This result is critical for interpreting how the time series variation we introduce impacts the decision to enroll in college. In the absence of any borrowing constraint, the variation we observe would play *no* role in understanding enrollment decisions. Thus, time-varying college costs, wages, and ability signals change attainment through their ability to impact the financing of a college education. We now ask how these time-varying factors help us generate model-predicted changes in ability and sorting, highlighting their interaction with borrowing constraints.

5.2.1. Time series variation and its interaction with borrowing constraints

We provide three counterfactual economies in which each model time series is held fixed. Specifically, the first constructs a counterfactual series for college costs that is a constant fraction of average income. The second fixes the college graduate and some-college earnings premia over time. The last assumes that the precision of ability signals remains constant over time, instead of declining over time. We use these exercises to show the importance of each component for understanding our results.

Table 3 provides the results for each counterfactual economy, and includes the attainment and ability gap for various years. The full time series are available in Fig. 9. The left panel shows the changing time series for each counterfactual, while our two main outcomes of interest are in the middle (college attainment) and right (ability sorting) panels. We discuss these two outcomes in turn.

[25] Fig. 8 shows the calculation for men, but the figure for women is identical with only a level difference.

[26] The same numbers for the ninetieth percentile are 26, 23, and 32 percent for the 1900, 1920, and 1950 cohorts. The decline between the 1900 and 1920 cohorts is a function of rising college costs which high ability individuals are relatively more likely to pay.

Table 3
Levels of attainment and the ability gap, various years.

Birth years	Benchmark	Counterfactual economies		
		Constant college costs	Constant earnings premium	Low signal variance
Panel A: college attainment				
1900	0.038	0.035	0.064	0.042
1920	0.050	0.091	0.197	0.075
1940	0.201	0.156	0.271	0.203
1950	0.241	0.217	0.284	0.241
1972	0.313	0.313	0.283	0.313
Panel B: ability gap				
1900	0.103	0.097	0.170	0.336
1920	0.128	0.130	0.170	0.308
1940	0.220	0.194	0.191	0.278
1950	0.230	0.216	0.243	0.230
1972	0.170	0.170	0.243	0.170

Accounting for the time series of attainment Panel A of Table 3 provides the college attainment results. The main result is that college costs and the college earnings premia play a critical role in matching attainment. Interestingly, however, the high signal variance experienced by earlier birth cohorts plays almost no role.

In light of our previous result of binding borrowing constraints, that college costs play an important role should be unsurprising. The counterfactually lower price for 1910–1920 cohorts increases attainment for the 1920 cohort by 4 percentage points relative to the baseline. The counterfactual model then under-predicts attainment by 4.5 percentage points in 1940, as the cost of college declines dramatically for post-1930 cohorts in the baseline model. Thus, the steep decline in costs for post-1920 cohorts is critical for matching the attainment data for 1920–1950 cohorts.

Fixing the some-college and college earnings premia causes a drastic over-prediction of attainment for pre-1950 cohorts. By 1920, the counterfactual model with fixed earnings premia over-predicts attainment by over a factor of four (0.197 compared to 0.050) and continues until the 1950 cohort. Fig. 9d shows that this over-prediction occurs for the same cohorts in which our counterfactual economy implies higher earnings premia relative to the benchmark model. This link is slightly more subtle, as it is not the case that individuals are responding directly to the higher income available to college graduates (recall that the previous result in Fig. 8).[27] Instead, the critical implication of the changing earnings premia is its impact on the level of borrowing allowed. Since the borrowing limit is a fraction of lifetime income, higher earnings premia for pre-1950 cohorts implies looser borrowing constraints, and thus higher attainment. Thus, the relatively low earnings premia for pre-1950 cohorts plays an important role in keeping attainment low.[28]

Lastly, unlike college costs and earnings premia, the precision of ability signals has almost no impact on attainment. For 1900 cohorts, the counterfactual economy predicts attainment of 0.042, close to the benchmark result of 0.038. Fig. 9h shows some effect between the 1910–1920 cohorts, but otherwise it has little impact in the full time series. Again, we can point to our previous result that all individuals would prefer college attendance in the absence of borrowing constraints. Low precision in ability signals changes which individuals believe themselves to be high ability. Those individuals internalize the impact of ability on the available borrowing limit, and thus attend college. Thus, while ability signal precision changes *who* goes to college, it does not have a substantial affect on the *number* of people who attend. Next, however, we show that this has important implications for ability sorting.

Overall, our results show the critical interaction between costs and earnings premia for generating the growth in college attainment over time. The low earnings premia is necessary to counteract the large decline in college costs for cohorts born in the 1920s and 1930s. As the college earnings premium begins to rise for post-1930 cohorts, increasing college costs provide a countervailing effect that smooths the growth in attainment over time. Thus, both effects are critical to generate the full time series of attainment, but which force provides upwards pressure on attainment varies depending on the time period.

Accounting for the time series of ability Panel B of Table 3 turns to the ability gap, while the full time series are available in the right panels of Fig. 9.

Despite the fact that cost variation impacts college attainment, it has little impact on the ability gap. For 1900–1920 cohorts, the counterfactual model follows the benchmark model closely. It predicts an ability gap of 0.097 for the 1900 cohort (0.103 in the benchmark model) and 0.130 for the 1920 cohort (0.128 in the benchmark). Though it slightly under-predicts ability relative to the benchmark for the 1930 and 1940 cohorts, the magnitude is quite small (Fig. 9c).

[27] In all counterfactuals, expected present value of lifetime earnings net of college costs for college attendees is higher than those who do not attempt college, as required for this result to continue to hold.

[28] Put somewhat more concretely, financial institutions observe the increasing income afforded to a student who chooses to enter college, and thus understand that these students will have higher financial resources to repay any required loans. Thus, they respond by allowing college attendees to borrow more.

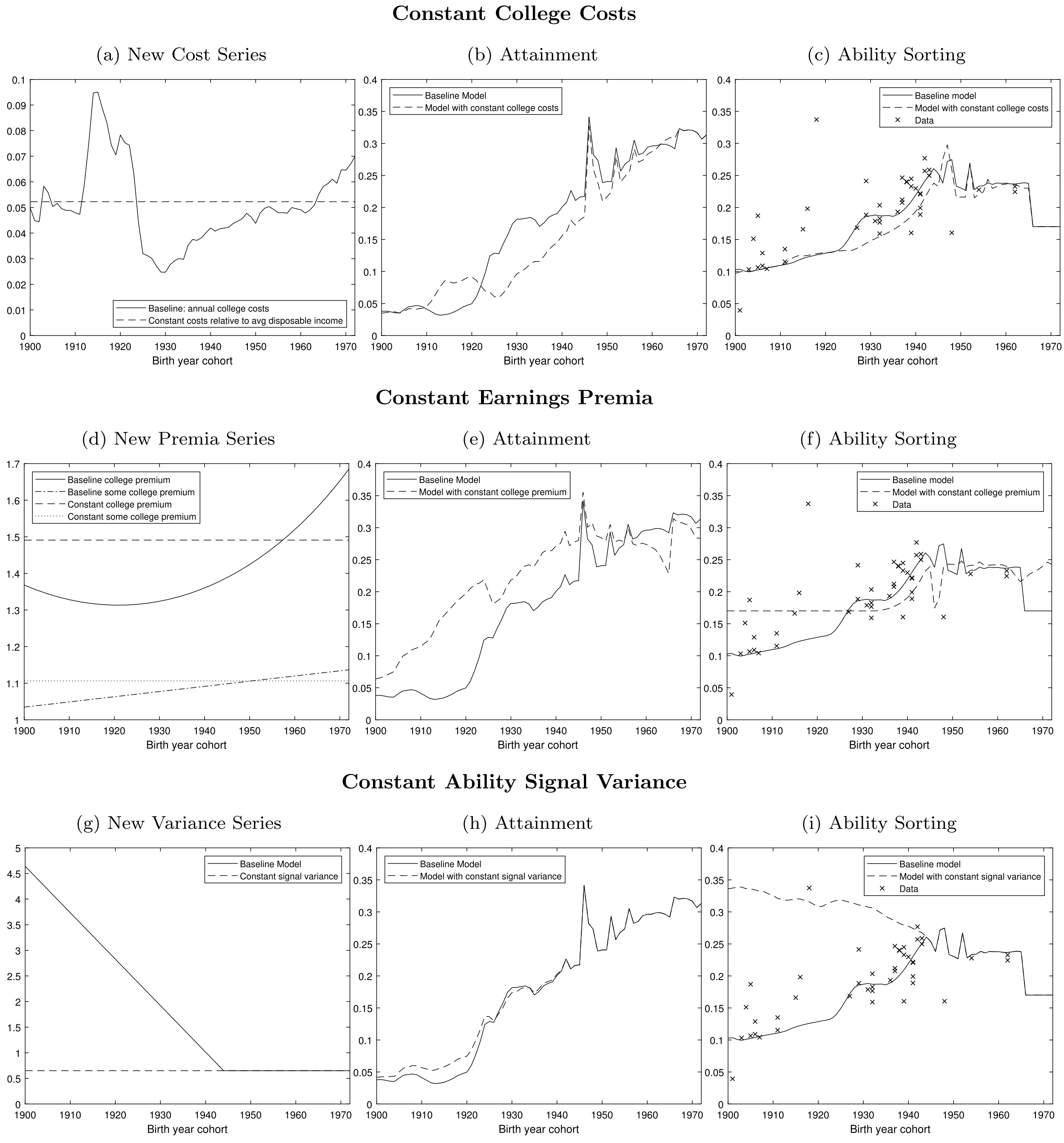

Fig. 9. Counterfactual economies.

The result is driven by the fact that changes in college costs affect marginal ability individuals. The rationale is the same used in our discussion of attainment – high ability individuals can endogenously borrow more, so are less likely to be constrained by costs. Therefore, the strong decline in costs faced by cohorts in the 1930s induces marginal students to enter college, thus lowering the average ability of college and non-college individuals. Fig. 10 shows this result graphically. When we hold costs artificially high for cohorts in the 1930s, both average college and non-college ability rise as fewer marginal students can enter college. Costs therefore have a limited impact on ability as changes shift both curves simultaneously. Put differently, this is entirely a function of the *type* of individual who switches to college in response to college cost variation.

Unlike college costs, earnings premia play a quantitatively important role in generating increased sorting for pre-1930 cohorts (Fig. 9f). Fixing the earnings premia causes a 65 percent over-prediction of the ability gap for the 1900 cohort relative to the benchmark model (0.170 to 0.103). This continues through the 1920 cohorts, and the model predicts an

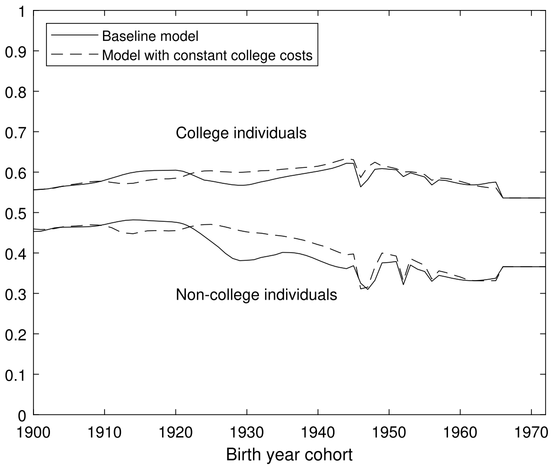

Fig. 10. Avg. ability percentile for college and non-college individuals.

ability gap of 0.170 for the 1920 cohort (0.128 in the benchmark). The rationale again relies on the impact of borrowing constraints. Because low premia generate tighter borrowing constraints, a number of high ability students are unable to attend college. The higher earnings premia in the counterfactual economy thus allows a larger share of the population to attend college. In fact, the premia are high enough in this counterfactual economy that everyone attempts college for cohorts up through the mid 1930s. As attendance increases, the set of non-college individuals becomes increasingly comprised of first-year college dropouts. These dropouts are negatively selected on ability, thus inducing a larger ability gap at higher returns to college. Indeed, comparing the time series of the earnings premia (Fig. 9d) and the ability gap (Fig. 9f), we observe this result in the benchmark model as well.

Finally, the increasing precision of ability signals is not only important to generate the magnitude of the increase in sorting over time, but the slope of the relationship as well. In fact, setting the signal variance for all cohorts to its post-WWII level, the model predicts a decline in the ability gap from 0.336 to 0.230 between the 1900 and 1950 cohorts, compared to an increase from 0.103 to 0.230 in the benchmark model. Signal precision therefore makes little difference in the total number of people who attend college, but plays a critical role in *who* goes to college. The result again relies on the fact that borrowing constraints play a critical role. With high signal variance, some low ability individuals believe themselves to be high ability. They internalize the impact this has on their capacity to pay for college through ability's impact on the borrowing constraint, and thus attend college. As ability signals become more precise, this mismatch between ability and college attendance plays a critical role in generating the proper time series of the ability gap, as Fig. 9i shows.

Taken together, our results imply a critical interaction across all three time varying features of the economy in generating time series variation in attainment and the ability gap. Variation over time in college costs and earnings premia are critical for generating the proper time series of college attainment, while ability signals play little role. However, college costs play little role in understanding ability sorting over time. The earnings premia play a quantitatively important role for pre-1930 cohorts, but changes in the ability signal precision play a critical role in generating not only the quantitative magnitude, but in fact the sign of the change over time. Missing any one of these features would render the model unable to jointly match the time series of attainment and ability.

6. Discussion and robustness

Before concluding, we discuss a number of possible alternatives and policy details from our model. In doing so we draw heavily on historical evidence regarding college dropout risk, student aid, and loan availability. Lastly, we briefly discuss a number of alternative exercises in the model that are detailed more completely in Appendix D.

6.1. Does risk of non-completion change over time?

Our baseline calibration involves a constant risk of dropout given ability. This is a function of our inability to observe a complete time series of the relationship between ability and college exit rates. However, we can provide some aggregate evidence on college graduation rates over time. In modern data, we observe that the 1996 starting class of college freshmen had a five year graduation rate of 50.2 percent (Table 326.10 in the 2016 *Digest of Education Statistics*). This is the earliest aggregate data available in the DES, and it follows just after our 1972 birth cohort would have completed college. It is also a reasonable benchmark to keep in mind since our time series for college completion is measured at age 23, i.e., five years after the typical high school graduate enters college at age 18.

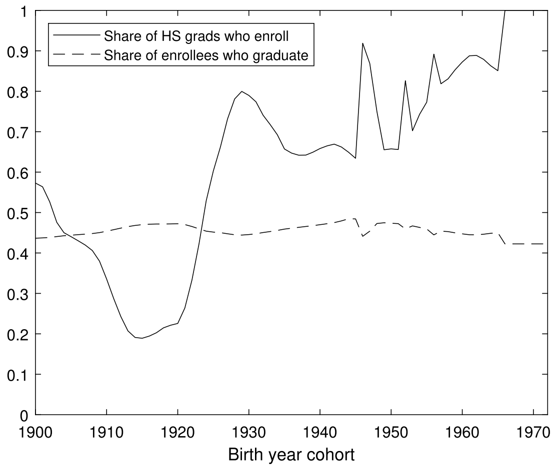

Fig. 11. Baseline model enrollment and graduation rates.

Earlier micro evidence suggests similar graduation rates. Iffert (1958) followed more than 13,000 college freshmen in 1950 (i.e., the 1932 birth cohort) and calculated a graduation rate of 51 percent. Even earlier, McNeely (1937) followed more than 15,000 students across 25 institutions who were first-time college freshmen in 1931–1932 and found a 58 percent graduation rate after five years. Further evidence comes from the 1942–1944 *Biennial Survey of Education*. The average graduation rate across 1931–1940 high school graduates (1913–1922 birth cohorts) who subsequently entered college was 48.6 percent. Fig. 11 plots the college enrollment and graduation rates in our baseline model. The graduation rate is just under one-half and relatively stable across cohorts, despite substantial swings the enrollment rate. Thus, despite a slight underestimation of the graduation rate, our assumption seems consistent with available empirical evidence.[29]

A final related question is whether college non-completion is primarily due to academic or other reasons, such as individual financial shocks. Modern evidence shows that relatively few students leave college purely financial reasons and often much more for academics, even among the lowest income families (Stinebrickner and Stinebrickner, 2008). For a historical perspective, McNeely (1937) surveyed students on their reasons for exiting college. For the 1931–1932 freshmen class (birth cohort 1918), he finds that only 12.4 percent of those exiting reported financial difficulties as the main reason for leaving university. Furthermore, he follows credit accumulation rates of leavers versus stayers, and concludes: “the predominant proportion of the students leaving the universities were deficient in their academic work to the extent of failing to earn the same credit-hours for which they registered in the semester or quarter when they left the institutions.” Note that this is during the Great Depression, where we would presumably expect excessive financial burdens of college costs. Given this evidence, we find it reasonable to abstract from financial shocks and maintain the assumption in our structural model that college risk is only a function of student ability.[30]

6.2. Historical evidence of cost and borrowing opportunities for college

Our model shows that the substantial variation in college costs plays an important role in understanding the empirical trends for early twentieth century birth cohorts. We now detail some of the policy changes that generate this variation.

Direct college costs Several factors contributed to keeping college costs down during the pre-World War II era. First, there was a large increase in the supply of higher education institutions. In the two decades from 1920 to 1940, the number of higher education institutions in the U.S. increased from 1,041 to 1,713. Institutions faced increasing competitive pressure during this era of expansion, and therefore found it difficult to raise tuition and fees, particularly during the depression. Harris and Edwin (1962), for example, relates the experience of Sweet Briar College which, “...in 1931–1932 had to increase its tuition from $800 to $1,000 to cover its costs. But applications dropped from 647 to 292.”[31]

Student aid Because our measure of college costs is net of student aid, our price series does not necessarily reflect the sticker price of college. Any increase in aid generosity would cause our series to grow more slowly than the sticker price of college, for example. We briefly discuss here the evolution of student aid over time. Aggregate data on student aid

[29] Recall that we estimate the function $\pi(\alpha)$ using data from the NLSY79, which covers birth cohorts from the early 1960s. Blandin and Herrington (2018) compare the NLSY79 graduation rates to the NLSY97 cohorts born 1980–1984 and show that they are also very similar.

[30] However, we also note that student ability and financial resources are positively correlated in the data, and therefore it will be true that poorer students face higher risk on average.

[31] As noted by Harris and Edwin (1962), many public institutions were also restricted from raising costs as a matter of public policy.

expenditures were first reported in the *Biennial Surveys of Education* for the 1951–1952 academic year, at which time they amounted to 9 percent of the $448 million paid by students for tuition and fees. By 1960 student aid had grown to represent 15 percent of student payments, and by 1970, over 22 percent. Thereafter, it leveled off and fluctuated around 20 percent until 1990, when our last cohort started college. In general, student aid in the post-World War II era seems to have played a nontrivial role in dampening what would have otherwise been even steeper increases in college costs.

The earliest direct evidence we have found on the characteristics of students receiving scholarships comes from Iffert (1958), who studied the 1950 entering class of college freshman. Despite the fact that 25 percent of students in that study received scholarship aid at some point during their college career, only 0.6 percent received enough to cover 80–100 percent of their cost. For about two-thirds of those receiving aid, the amount covered less than 20 percent of their expenses. Moreover, there was no strong relationship between aid awards and students' test scores or high school class rank. Iffert concluded that "scholarships are not reserved exclusively for scholars."

Pre-1950 data is more difficult to find systematically, however some information is available. Harris and Edwin (1962) calculated student aid as a share of tuition payments to be 5 percent in 1940, consistent with the trend for later years described above. A U.S. Office of Education Bulletin from 1931 (Ratcliffe, 1931) reports that the number of available scholarships for the 1927–1928 academic year were only sufficient to cover 4 percent of total undergraduate enrollment.[32] Collectively these data seem to confirm a prevailing viewpoint of the time, as noted in the 1924–1926 *Biennial Survey of Education*, that those desiring education, "...must do so under the modern one-price policy—the same cost to all for the same service."

Overall, scholarships and aid seem to be of limited importance in informing our cost series for cohorts born before 1950. Based on the above historical evidence, we adopt the "one-price" assumption in our baseline quantitative model. However, different states certainly had different college tuition costs, so our assumption of no heterogeneity is potentially too strong. In the Appendix, we conduct a robustness check in which we introduce variation in college prices disciplined by observed cross-state price variation. We discuss this more below.

Student loans Having discussed college costs and aid, we now turn to the issue of student loans as a means of financing those costs. To our knowledge the first aggregate data collected by the U.S. government on student loan funds is for 1940, as reported in various issues of the *Biennial Surveys of Education*. In that year the value of total student loan funds was $28.9 million, which amounted to 14.5 percent of revenue from student fees. By 1958, student loan funds had grown to $71.2 million, but only represented 7.6 percent of revenue from student fees.

Prior to 1940 we are not aware of any official data on student loans, but Harris and Edwin (1962) reports one estimate that around $10 million in student loan funds were available during the early 1930s.[33] We emphasize two main takeaways from this historical evidence. First, while relatively small compared to the modern era, loans were available to help finance college at least as far back as the 1920s. Second, student loans from endowed funds and private sources grew until they were essentially supplanted by federal loan programs. We try to capture these two main features in our quantitative model using a version of the natural borrowing constraint that grows along with real income growth and is superseded by federal student loans for later cohorts.[34]

6.3. Robustness and further experiments

We perform a number of different experiments and robustness checks on our quantitative results. We list them briefly here, and the full details are in Appendix D.

First, we consider the differential impact of World War II GI Bill subsidies for veteran men versus women and non-veteran men. In particular, we provide free tuition in the model to a subset of men designated as veterans, and we utilize an alternative measure of costs that more accurately estimates the actual costs faced by women and non-veteran men who did not benefit directly from the GI Bill. Second, we consider heterogeneity in college costs. Our main motivation in this exercise is the large observed variation in public college costs across states, so we seek to determine whether results are robust to the assumption of a single college price per cohort. Third, we allow ability to directly affect earnings after any education is

---

[32] Even this limited availability of aid presented here is most likely too generous an interpretation. Ratcliffe (1931) further notes that, "This requirement is one that is often made by the institutions in granting students aid from their own funds, generally after the students have been enrolled at the institution for a sufficient length of time to demonstrate their worthiness."

[33] We do not to suggest that loans actually accounted for such a large share of the funds used to pay for college, but this much was at least *available* to borrow. Harris and Edwin (1962) reports that loans accounted for about one percent of all undergraduate student charges in 1956, and we have no evidence to suggest they were significantly larger in earlier years.

[34] It is interesting to note, however, that the design of these loan programs differed quite a bit from their modern counterparts. Loan terms were generally shorter than modern federal student loans. Harris and Edwin (1962) compiled information on student loan funds for a small set of public and private universities in 1955–1956 and found loan repayment periods typically ranging three to six years. Some individual universities also arranged for students to obtain financing through local banks. Earlier evidence indicates even shorter loan terms. For example, Indiana University bulletins throughout the 1920s and 1930s describe the available student loan funds. The typical terms were for one to two years at three percent interest. Loan repayments and accrued interest were generally used to grow the funds and provide for future lending. Though outside the scope of this paper, an interesting path for future research is the relationship between the design of early twentieth century loan programs and their impact on ability sorting and attainment. Lochner and Monge-Naranjo (2011) highlight the importance of this topic in modern loan programs.

completed. Here we want to determine whether returns to innate ability, in addition to the average returns to education, affect overall attainment and ability sorting. Lastly, we consider an experiment that allows for the ability distribution of high school graduates to vary across cohorts. The objective here is to model selection in high school dropouts, which could be important given the large growth in high school graduation discussed in Section 2. In all cases, the results change slightly, but the overall message remains the same. These results and discussion surrounding them are available in Appendix D.

7. Conclusion

We develop a life-cycle decision model to investigate long-run trends in college completion and ability sorting for the 1900–1972 birth cohorts. Key features of the model include unobservable ability, heterogeneous parental transfers, and borrowing constraints. To discipline our model, we utilize historical data series including statistics on high school graduation rates, college costs, and education earnings premia. The model matches the overall trends in college attainment and ability sorting between college and non-college individuals well.

Counterfactual results demonstrate that the model's ability to match the time series of attainment and ability depends critically on the time varying factors we build into the model (college costs, earnings premia, and ability signal precision). College costs and earnings premia play an important role in matching the time series of college attainment, while ability signal precision plays almost none. On the other hand, college costs play little role in understanding ability sorting over time, while earnings premia play an important role. Without increased precision in ability signals, however, the model predicts a decline in the ability gap over time. Crucially, all of these results rely on the interaction with binding borrowing constraints.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.red.2018.07.003.

References

Athreya, Kartik, Eberly, Janice, 2016. Risk, the College Premium, and Aggregate Human Capital Investment.

Autor, David H., Katz, Lawrence F., Krueger, Alan B., 1998. Computing inequality: have computers changed the labor market? The Quarterly Journal of Economics 113 (4), 1169–1213.

Beale, Andrew V., 1970. The evolution of college admission requirements. Journal of the National Association of College Admissions Counselors 15 (3), 14–15.

Belley, P., Lochner, L., 2007. The changing role of family income and ability in determining educational achievement. Journal of Human Capital 1 (1), 37–89.

Blandin, Adam, Herrington, Christopher, 2018. Family Structure, Human Capital Investment, and Aggregate College Attainment. Virginia Commonwealth University.

Bound, John, Turner, Sarah, 2002. Going to war and going to college: did World War II and the GI Bill increase educational attainment for returning veterans? Journal of Labor Economics 20 (4), 784–815.

Bound, John, Lovenheim, Michael F., Turner, Sarah, 2010. Why have college completion rates declined? An analysis of changing student preparation and collegiate resources. American Economic Journal: Applied Economics 2 (3), 129.

Card, David, Lemieux, Thomas, 2001a. Dropout and enrollment trends in the postwar period: what went wrong in the 1970s? In: Risky Behavior Among Youths: An Economic Analysis. University of Chicago Press, pp. 439–482.

Card, David, Lemieux, Thomas, 2001b. Going to college to avoid the draft: the unintended legacy of the Vietnam War. The American Economic Review 91 (2), 97–102.

Castro, Rui, Coen-Pirani, Daniele, 2016. Explaining the evolution of educational attainment in the United States. American Economic Journal: Macroeconomics 8 (3), 77–112.

Cunha, Flavio, Heckman, James, Navarro, Salvador, 2005. Separating uncertainty from heterogeneity in life cycle earnings. Oxford Economic Papers 57 (2), 191–261.

Garriga, Carlos, Keightley, Mark P., 2007. A General Equilibrium Theory of College with Education Subsidies, In-School Labor Supply, and Borrowing Constraints. Working Paper. Federal Reserve Bank of St. Louis.

Goldin, Claudia, Katz, Lawrence F., 2010. The Race Between Education and Technology. Harvard University Press.

Harris, Edwin Seymour, 1962. Higher Education: Resources and Finance. McGraw-Hill.

Hendricks, Lutz, Leukhina, Oksana, 2017. How risky is college investment? Review of Economic Dynamics 26, 140–163.

Hendricks, Lutz, Leukhina, Oksana, 2018. The return to college: selection bias and dropout. International Economic Review. https://doi.org/10.1111/iere.12297, forthcoming.

Hendricks, Lutz, Schoellman, Todd, 2014. Student abilities during the expansion of US education. Journal of Monetary Economics 63, 19–36.

Hendricks, Lutz, Herrington, Chris, Schoellman, Todd, 2018. A Long Run View on College Access and Attendance Patterns. Working Paper.

Hubbard, William H., 2011. The phantom gender difference in the college wage premium. The Journal of Human Resources 46 (3), 568–586.

Iffert, Robert E., 1958. Retention and Withdrawal of College Students. US Office of Education Bulletin.

Ionescu, Felicia, 2009. The federal student loan program: quantitative implications for college enrollment and default rates. Review of Economic Dynamics 12 (1), 205–231.

Kaboski, Joseph P., 2005. Supply Factors and the Mid-Century Fall in the Skill Premium. Working Paper. University of Notre Dame.

Keane, Michael, Wolpin, Kenneth, 2001. The effect of parental transfers and borrowing constraints on educational attainment. International Economic Review 42 (4), 1051–1103.

Keller, Elisa, 2014. The slowdown in American educational attainment. Journal of Economic Dynamics and Control 46, 252–270.

Lochner, Lance J., Monge-Naranjo, Alexander, 2011. The nature of credit constraints and human capital. The American Economic Review 101 (6), 2487–2529.

McNeely, John H., 1937. College student mortality. US Office of Education Bulletin 11.

Ratcliffe, Ella B., 1931. Scholarships and fellowships. US Office of Education Bulletin 15.

Restuccia, Diego, Vandenbroucke, Guillaume, 2013. The evolution of education: a macroeconomic analysis. International Economic Review 54 (3), 915–936.

Ruggles, Steven, Trent Alexander, J., Genadek, Katie, Geoken, Ronald, Schroeder, Matthew B., Sobek, Matthew, 2010. Integrated Public Use Microdata Series: Version 5.0 [Machine-Readable Database]. University of Minnesota, Minneapolis.
Snyder, Thomas D. (Ed.), 1993. 120 Years of American Education: A Statistical Portrait. United States Department of Education.
Stinebrickner, Ralph, Stinebrickner, Todd, 2008. The effect of credit constraints on the college drop-out decision: a direct approach using a new panel study. The American Economic Review 98 (5), 2163–2184.