Optimal Timing of College Subsidies: Enrollment, Graduation, and the Skill Premium*

Kazushige Matsuda

August 3, 2019

[Link to the latest version]

Abstract

In the United States, half of college enrollees drop out before earning a bachelor's degree. This paper examines the effect of a college subsidy scheme, in which the subsidy amount varies across years in college, on college dropout and wage inequality. I find that by increasing college subsidies from freshmen

*Matsuda: Hitotsubashi University, Graduate School of Economics, 2-1 Naka, Kunitachi, Tokyo, 186-8601, Japan, matsuda.kazushige@gmail.com. I am extremely grateful to my advisor, Richard Rogerson, for his advice, support, and encouragement. I am grateful to Mark Aguiar, Titan Alon, Kosuke Aoki, Gregor Jarosch, Ryo Jinnai, Sagiri Kitao, Nobuhiro Kiyotaki, Dirk Krueger, Alexander Ludwig, Tomohide Mineyama, Ezra Oberfield, Kota Ogasawara, Esteban Rossi-Hansberg, and Gianluca Violante for helpful comments. I also received valuable feedback from audiences at various seminars. The author of this paper, Kazushige Matsuda, had financial support from the Princeton University Fellowship. The author declares that he has no relevant material or financial interests that relate to the research described in this paper.

to seniors, the number of college graduates increases. In addition, the skill premium decreases more than the case in which the total budget for current constant subsidies is increased by 50%. The scheme is welfare improving, despite the fact that enrollment decreases.

Keywords: Education, Financial Aid, College Dropout, Skill Premium

JEL Classification: E24, I22, J24

1 Introduction

Wage inequality has been increasing in the United States. For instance, the skill premium—the wage premium of college graduates compared with high school graduates—increased from 50% in 1980 to 90% in 2000. A large literature (e.g., Goldin and Katz (2009) and Katz and Murphy (1992)) argues that the skill premium rises at least in part because the increase in the supply of college graduates does not keep pace with the increase in the demand for skilled labor.[1] Within this framework, we can reduce the skill premium by increasing college graduates in the economy. With this in mind, the literature (e.g., Krueger and Ludwig (2016)) often recommends increasing college subsidies to induce college enrollment. However, in the United States, while over 70% of high school graduates enroll in college, more than half drop out before earning a bachelor's degree[2]. Enrollment does not necessarily lead to graduation, and simply increasing subsidies might serve to increase college

[1] The literature assumes that schooling is not just a signal but also increases human capital. Kroch and Sjoblom (1994) provide empirical evidence for this.

[2] Two-year college graduates who do not transfer are counted as dropouts. Horn and Skomsvold (2011) show that the ultimate goal of over 80% of freshmen at community colleges is a bachelor's or higher degree. The "sheepskin effect" of associate degrees is not high (see Kane and Rouse (1995)), and only 5% of enrollees at 2-year colleges graduate and do not transfer (see Trachter (2015)).

dropouts rather than college graduates. It is important that we consider changing the structure of college subsidies without increasing the total budget.

In this paper, I examine a new college subsidy scheme in which the subsidy amount varies with years of college ("year-dependent subsidies"), i.e., different subsidies for freshmen, sophomores, and so on. In contrast, the literature has only considered subsidies that are constant across years in college ("year-invariant subsidies"). Subsidies that vary by year will have differential impacts on enrollment and graduation, unlike constant subsidies, as the following example demonstrates. After graduating from high school, individuals decide to enroll in college based on their high school grade point average (GPA) or high school ability. People with a high GPA want to enroll, but one's high school GPA is not necessarily the same as his or her college GPA. After enrolling, some students learn that their college GPA or college ability is low and drop out. Consider back-loaded subsidies in this setting: higher subsidies for juniors and seniors and lower subsidies for freshmen and sophomores. People who expect to drop out before earning the higher subsidies in later years stop enrolling due to the lower subsidies for early years in college. In contrast, the marginal college dropout now finds it worthwhile to continue, since subsidies for later periods increase. Therefore, the number of college graduates increases while enrollment decreases. The opposite case holds for front-loaded subsidies. Year-dependent subsidies can affect enrollment and graduation in different ways, unlike increasing or decreasing constant subsidies. The question of this paper is how year-dependent subsidies affect enrollment and graduation and how the timing of subsidies will maximize the number of college graduates and welfare.

I build a lifecycle general equilibrium model with credit constraints and endogenous enrollment and dropout decisions. Agents are heterogeneous with regard to initial asset, high school ability, and college ability. College ability affects the psychic cost of college attendance and earnings after graduation. Agents observe high school ability but cannot observe college ability before enrollment. In order to be consistent with the empirical findings of Stinebrickner and Stinebrickner (2012), agents are overly optimistic with regard to college ability before enrollment. Agents learn their college abilities after enrollment and decide to drop out or not. These educational decisions shape the aggregate skill in the economy, which in turn determines the skill premium, though imperfect substitution between skilled and unskilled labor. I calibrate the model to match the enrollment, graduation, and skill premium of the United States given the current policy. Using the model, I examine how year-dependent subsidies have differential impacts from the current constant subsidies on enrollment, graduation, the skill premium, and the expected utilitarian social welfare function. The focus of this paper is on the relative sizes of college subsidies across years in college, and I fix the total budget of college subsidies at the current level.

This paper's main findings are as follows. First, back-loaded subsidies maximize the number of college graduates and welfare. Second, by switching to back-loaded subsidies with the same total budget, the number of college graduates increases and the skill premium decreases more than the case in which the total budget for constant subsidies is increased by 50%. Third, back-loaded subsidies improve social welfare by 0.15% of lifetime consumption at the steady state, without increasing the government budget. Back-loaded subsidies reduce enrollment—which was excessive due to excessive optimism—and prevent low-ability people from deriving the psychic cost of college attendance, which improves welfare. This result suggests that increasing enrollment is not necessarily welfare-improving, and that policies aiming to increase graduation matter. Fourth, the reduced skill premium leads to a decrease in the difference in wages between college graduates and college dropouts, which also reduces the uncertainty of wages from uncertain college ability, which is beneficial for risk-averse agents.

The model in this paper is based on sequential papers by Stinebrickner and Stinebrickner (2003, 2008, 2014) and Stinebrickner and Stinebrickner (2012). They use a longitudinal survey of students at Berea College that asks approximately 12 times per year about their expectations of college GPA, enjoyability of college, and their expectations on earnings after graduation for different college GPA levels. They find that credit constraints do not play a major role in college dropout and that learning ability is a key factor.[3] They argue that college ability affects students' enjoyability and expected earnings after graduation. They also find that students are overly optimistic before enrollment. I bring their empirical findings into a quantitative heterogeneous agent model to examine the effect of year-dependent subsidies. Optimism plays a key role for welfare implications, and one of the contributions of this paper is to calibrate this effect using available data.

There is a large literature on the effect of college subsidies. For example, Bovenberg and Jacobs (2005) theoretically derive the effect of college subsidies. Abbott, Gallipoli, Meghir, and Violante (2013) emphasize the effect of subsidies on parental transfers in a quantitative overlapping-generations model. Krueger and Ludwig (2016) analyze the optimal income tax and college subsidies and show that

[3] Stinebrickner and Stinebrickner (2012) also found that the effects of students' health or parental job loss on college dropout are not statistically significant.

the mixture of a less progressive labor income tax and a larger amount of subsidies than in the current state are optimal for a utilitarian social welfare function. They abstract from college dropout and do not consider year-dependent subsidies.

This paper relates to a large literature that examines college subsidies in a model of college dropout. Caucutt and Kumar (2003) and Akyol and Athreya (2005) show the normative implication of college subsidies with exogenous college dropout risk. Hanushek, Leung, and Yilmaz (2004) analyze the effect of various college aid plans with exogenous college dropout risk. Ionescu (2011), Garriga and Keightley (2007), and Chatterjee and Ionescu (2012) study the effect of education policies on educational decisions with endogenous college dropout. This paper considers year-dependent subsidies, which have not been considered in the literature.

The rest of the paper is organized as follows. Section 2 outlines the model and defines an equilibrium. Section 3 renders the model quantitative by calibration and estimation. Section 4 presents the results, and in Section 5 I discuss the findings and offer concluding remarks.

2 Model

The model has four main building blocks. The first is year-dependent subsidies, which are new to the literature. Although subsidies are assumed to be constant across years in the current state, I examine the effect of year-dependent subsidies as a counterfactual exercise in the results section.

The second is a model of endogenous enrollment and graduation decisions based on Garriga and Keightley (2007). After high school graduation, heterogeneous individuals make an enrollment decision based on their initial asset and high school ability. College enrollees learn their college abilities after enrollment and decide to drop out of college or not. College ability determines the psychic cost of college attendance and earnings after college graduation.

The third building block is a general equilibrium framework with an aggregate production function that features imperfect substitution between skilled and unskilled labor. Educational decisions aggregate to the supply of skill and the skill premium.

The fourth building block is a full-blown overlapping-generations life cycle with intergenerational linkage. Workers choose labor supply endogenously, which determines the skill premium. Individuals give birth to children with intergenerationally correlated ability and make an endogenous wealth transfer to their children. College subsidies can crowd out endogenous wealth transfers from parents to their children. After retirement, agents receive a pension. I approximate the U.S. current system of pension with progressive benefits. It is important to discuss the effect of college subsidies on inequality given the progressivity of other policies, including pension. A full-blown lifecycle model is necessary to capture these mechanisms.

Since I focus on a stationary equilibrium in which the cross-sectional allocation within each cohort is invariant and prices are constant, I do not include time subscripts in the description of the economy.

2.1 Demography

The economy is inhabited by a continuum of overlapping-generations individuals. Age is indexed by $j \in \{1, 2, \ldots, J\}$. Each individual has one offspring. At the beginning of age 1 (biological age 18), individuals become economically independent as high school graduates.

The timeline is as follows. At the beginning of age 1, individuals make enrollment decisions. Once they do not enroll in college, they cannot enroll later. One period in the model corresponds to 2 years. Because college typically requires 4 years in reality, college graduation requires 2 periods in the model. At the beginning of age 2, a college enrollee makes a decision about whether to stay in college until graduation or drop out. Once an individual finishes their schooling, they will be one of three types: high school graduate ($e = HS$) for those who do not enroll at age 1, college dropout ($e = CD$) for those who do not continue college at the beginning of age 2, and college graduate ($e = CG$) for those who finish 2 periods of college.

After that, they face a standard life cycle problem with income risk. Individuals give birth to children at age $j_f = 7$, which is biological age 30 ($j_f = (30 - 18)/2 + 1 = 7$). At age $j_b = 16$ (at biological age 48), their children become economically independent and parents transfer wealth to their children. No transfers are allowed at other ages[4]. Individuals retire at age $j_r = 25$ (at biological age 66), and the maximum age is $J = 42$ (at biological age 100). Individuals survive with probability $\varphi_j \in [0, 1]$ between age j and $j + 1$. I assume $\varphi_j = 1$ for $j \in [0, j_r - 1]$. The survival rate between j_r and $J - 1$ is taken from the US Life Tables 2000.

[4]If transfers are allowed at other ages, such as age 2, the state variables of parents must include their children's state variables, and solving the individual's problem becomes formidable. Transfers from parents at other ages change the result if credit constraints bind for their children. As we will see later, the credit limit for age 1 is tighter than the limit for age 2, and it is unlikely that allowing transfers from parents at age 2 changes the outcome.

2.2 Preferences

When an individual becomes economically independent at age 1, he or she has preferences represented by

$$\mathbb{E}_1\left[\sum_{j=1}^{J}\tilde{\beta}_{j-1}u(c_j,\ell_j)-\sum_{j=1}^{2}\tilde{\beta}_{j-1}d_j(\boldsymbol{s}_j^c)\lambda_j(\theta_c,\phi)+\tilde{\beta}^{j_b-1}\nu V_0\right]$$

where

$$u\left(c,\ell\right)=\frac{\left(c^{\mu}\ell^{1-\mu}\right)^{1-\gamma}}{1-\gamma}$$

$$\lambda_j(\theta_c,\phi)=\lambda+\lambda^{\theta}\theta_c+\lambda_j^{\phi}\phi.$$

The first term is the expected discounted sum of instant utility and c_j is consumption and ℓ_j is leisure at age j. $\mathbb{E}_1$ is the expectation operator conditional on the information at the beginning of age 1. Individuals are endowed with 1 unit of time each period. At age $j\in[j_f,j_b-1]$, individuals live with their children and consumption is discounted by $1+\zeta$, where ζ is an adult equivalence parameter. β is the time discount rate.[5]

The second term is the expected psychic cost of college attendance and $d_1(\boldsymbol{s}_1^c)$ is an indicator function that is one if the individual enrolls at age 1 and $d_2(\boldsymbol{s}_2^c)$ is an indicator function for continuing college at age 2. As in Cunha, Heckman, and Navarro (2005), the psychic cost of education is an important factor for determining educational choice. The psychic cost of college attendance depends on two components: college ability θ_c and college taste ϕ. Individuals observe their college taste ϕ before enrollment, which is fixed over the lifetime, while the coefficient λ_j^{ϕ} can

[5] $\tilde{\beta}$ is the effective time discount rate, taking survival into account: $\tilde{\beta}_j=\beta^j\left(\prod_{k=1}^{j}\varphi_k\right)$.

vary across periods (different loading).

The two factors θ_c and ϕ for the psychic cost of college attendance are necessary to match the data in which, within a category of high school ability and family income, there is heterogeneity in terms of enrollment decisions (some enroll, but others do not). To explain this, I need college taste ϕ, which is unobservable to econometricians. I explain ability and college taste in more detail in Section 2.6.

The third term is parental altruism, where V_0 is the expected lifetime utility of their children at the beginning of age 1. I will explain the value function in detail later. Individuals enjoy their children's expected lifetime utility with a weight ν. This is a motive for transfers from parents to children.

2.3 Goods Sector

There exists a representative firm producing the final good from capital K and aggregate labor services H following a production function

$$Y = F(K, H) = K^{\alpha} H^{1-\alpha}.$$

I follow Katz and Murphy (1992) by modeling aggregate labor services H as an aggregator of two skill levels: skilled labor H^S and unskilled labor H^U.

$$H = (a^S (H^S)^{\rho} + (1 - a^S)(H^U)^{\rho})^{\frac{1}{\rho}},$$

where $a^U = 1 - a^S$ and $\frac{1}{1-\rho}$ is the elasticity of substitution. a^S is the relative productivity of skilled labor. This representative firm rents capital at price $r + \delta$, where r is the interest rate and δ the depreciation rate, and hires two skills of labor at wages w^S and w^U. Markets for output and inputs are competitive, so that the first-order conditions for profit maximization yield

$$r = \alpha \left(\frac{K}{H}\right)^{\alpha-1} - \delta$$

$$w^s = (1-\alpha)a^s \left(\frac{K}{H}\right)^{\alpha} \left(\frac{H}{H^s}\right)^{1-\rho} \quad \text{for } \; s = S, U.$$

There are two types of skill in production and three levels of education. As in the literature on the skill premium, I assume that high school graduates provide unskilled labor and that college graduates provide skilled labor. I assume that college dropouts provide unskilled labor. Torpey and Watson (2014)[6] present the proportion of jobs in the United States by required education level, such as "Bachelor's degree," "Associate's degree," "Some college, no degree," "High school diploma or equivalent," or "Less than high school." They show that only 5% of jobs require "Some college, no degree" or "Associate's degree," which implies that most college dropouts take jobs requiring less than "High school diploma or equivalent." For convenience, I define the price of effective labor by college graduates, college dropouts, and high school graduates as $w^{CG} = w^S$ and $w^{HS} = w^{CD} = w^U$.

Effective labor per hour is denoted by $\varepsilon^e_j(\theta, \eta)$, which depends on education e, age j, ability θ, and idiosyncratic productivity η. The stochastic productivity shock η is mean-reverting and follows an education-specific Markov chain $\pi^e(\eta'|\eta) > 0$, and Π^e denotes its invariant distribution function. Ability θ depends on education levels. Ability θ is high school ability θ_h for high school graduates and college

[6]They use May 2013 data from the Occupational Employment Statistics survey (employment data) and Employment Projections program (occupational education-level designations) of the U.S. Bureau of Labor Statistics.

dropouts, while it is college ability θ_c for college graduates.

2.4 College

There is a representative college. To provide a college enrollee with 1 period of education requires κ units of skilled labor, which means that college enrollees receive education from professors who are college graduates. I assume that education does not require any capital or unskilled labor.[7]

The profit of the college is

$$p_e E - w^S \kappa E,$$

where E is the measure of college enrollees and p_e denotes tuition. Colleges are competitive and there is free entry. This implies, in equilibrium with positive units of students, that $p_e = w^S \kappa$. In the United States, colleges receive subsidies from the government, which enables the sticker tuition price to be smaller than the actual education cost. I reinterpret this situation as follows: Colleges do not receive any subsidies; instead, college enrollees receive subsidies. At the same time, they must pay the full education cost. In both cases, enrollees pay p_e less the subsidy to colleges.

2.5 Financial Markets

The financial market is incomplete. There is no insurance market against idiosyncratic risks, but individuals can self-insure using risk-free assets with interest r.

[7] Archibald and Feldman (2014) argue that college tuition reflects the wages of college graduates.

Lenders incur the cost of overseeing borrowers, and the cost per unit of capital is $\iota > 0$. With the non-arbitrage condition, the interest rate for borrowing workers is $r^- = r + \iota$. In addition, the borrowing limit for workers of education level e is $\underline{A}^e$, and retired individuals have no access to loans.

The cost of overseeing borrowing college enrollees is $\iota + \iota^s$. With the non-arbitrage condition, the interest rate for borrowing enrollees is $r^s = r + \iota + \iota^s = r^- + \iota^s$. The borrowing limit for college enrollees is $\underline{A}^c_j$ at age j.

2.6 Individual Problems

The lifecycle of individuals is basically composed of education, working, and retirement stages. Although college enrollees can also work, I call individuals who are not in college and are not retired "workers." Likewise, I call this period the "working stage."

2.6.1 Education Stage

Enrollment

At the beginning of $j = 1$, individuals become independent as high school graduates and their first decision is whether to enroll in college. I define V_0 to be the value function.

$$V_0(a, \theta_h, \eta, q, \phi) = \max\{\underbrace{V_1^c(a, \theta_h, \eta, q, \phi)}_{\text{enrolling}}, \underbrace{V_1(a, HS, \theta_h, \eta)}_{\text{not enrolling}}\}$$

An individual's initial state is composed of initial assets a, high school ability θ_h, an idiosyncratic transitory productivity η from Π^{HS}, parents' (family) income level q, and education taste ϕ.

Two types of ability are distinct but correlated: high school ability θ_h and college ability θ_c. Individuals observe high school ability through high school GPA or test scores during high school, but do not observe college ability until the beginning of age 2 after enrollment. Learning about college ability after enrollment is a key factor in college dropout. There is a large literature that emphasizes the importance of learning ability for schooling outcomes (e.g., Manski (1989); Altonji (1993); Arcidiacono (2004); Cunha et al. (2005); Stange (2012); Stinebrickner and Stinebrickner (2012)). College abilities are correlated with high school abilities, and

$$\theta_c = \theta_h + \epsilon_c \text{ where } \epsilon_c \sim N(0, \sigma_c^2).$$

σ_c is an important parameter for enrollment. If uncertainty σ_c about college ability is large, returns to graduation can be very large or small. If it turns out to be large after enrollment, enrollees can stay in college to earn high returns to graduation. If it turns out to be low, enrollees can drop out of college to dismiss the low returns to graduation. This asymmetry of returns increases the option value of enrollment.

In order to be consistent with the empirical finding of Stinebrickner and Stinebrickner (2012)[8], I assume that college enrollees are overly optimistic about their college abilities. They use a longitudinal survey of students, which asks each student her expectation of GPA multiple times. First, they show that students' expectations for their college GPAs before the first semester are higher than their actual GPAs on average, which suggests excessive optimism. Second, they show that col-

[8] As more evidence on optimism, Zafar (2011) argues that optimism matters in decisions about college major.

lege enrollees revise their expectations downward after enrollment, which suggests that they learn about their college abilities after enrollment. Third, students who drop out in early years were the most optimistic, and had the largest downward revisions of their expectations. Given θ_h, college enrollees expect that

$$\theta_c = \underbrace{\mu_c(\theta_h)}_{\text{bias}} + \underbrace{\theta_h + \epsilon_c}_{\text{actual ability}} \ \text{ where } \epsilon_c \sim N(0, \sigma_c^2)$$

where $\mu_c(\theta_h)$ is the bias. If $\mu_c(\theta_h)$ is positive, enrollees are overly optimistic about their college abilities. Furthermore, the bias can depend on high school ability, and I assume that $\mu_c(\theta_h) = \mu_c^0 + \mu_c^1\theta_h$. I assume that the variance of the residual term σ_c^2 is identical to the actual one.

Initial wealth a is endogenously determined as a transfer from their parents. If idiosyncratic productivity η is high, there is a good outside option to work and they don't want to enroll. Family income level q affects college subsidies, as will be seen later.

If an individual enrolls, she enters the first half of college where the value is V_1^c. If she does not enroll, she starts working as a high school graduate with value V_1.

First half of college

The value of being in the first half of college V_1^c is

$$\begin{aligned} V_1^c(a, \theta_h, \eta, q, \phi) = \max_{c,h,a',y} & \; u(c, 1 - h - \bar{h}) - \mathbb{E}_{\theta_c|\theta_h}\lambda_1(\theta_c, \phi) \\ & + \beta\mathbb{E}_{\theta_c|\theta_h}\mathbb{E}_{\eta'} \max\{\underbrace{V_2^c(a', \theta_c, \eta', q, \phi)}_{\text{continue}}, \underbrace{V_2(\tilde{a}(a'), CD, \theta_h, \eta')}_{\text{dropout}}\} \end{aligned}$$

subject to

$$c + a' + p_e - s_1(q) = a + y - T(c, a, y)$$

$$y = w^{HS}\varepsilon_1^{HS}(\theta_h, \eta)h, \;\; a' \geq -\underline{A}_1^c \;\; c \geq 0, \;\; 0 \leq h \leq 1 - \bar{h}$$

$$\theta_c = \theta_h + \mu_c(\theta_h) + \epsilon_c, \;\; \epsilon_c \sim N(0, \sigma_c^2) \text{ (perceived process)}, \;\; \eta' \sim \Pi^{CD}.$$

Going to college requires a fraction $\bar{h}$ of time, tuition p_e, and psychic cost $\lambda_j(\theta, \phi)$ for each enrollment period. c is consumption, h is labor hours, y is labor earnings, and a' is next-period assets. The total tax $T(c, a, y)$ depends on consumption, asset holdings, and earnings. College enrollees receive subsidies $s_j(q)$ dependent on family income q. They can work as high school graduates during the first half of college.

At the end of the period, college enrollees observe their college abilities θ_c and a new idiosyncratic productivity η' drawn from Π^{CD}. College enrollees decide whether to drop out of college after this. Those who observe a low college ability drop out for two main reasons. First, a low college ability leads to high psychic cost of college attendance until graduation. Second, the labor productivity of college graduates depends on college ability: A low college ability leads to low labor earnings if they graduate. Stinebrickner and Stinebrickner (2014) empirically show that these two are the main avenues through which learning about college ability is a key factor of college dropout.

If the individual drops out, her education level becomes college dropout ($e = CD$).[9] If the individual does not drop out, she proceeds to the second half of college

[9] After dropping out or graduating, all of the student loan is refinanced into a single bond that carries interest rate r^-. $\tilde{a}(a')$ is the transformation from the asset position during college to the position after college so that the total payment is identical. When making this calculation, I assume

with value V_2^c.

Second half of college

The Bellman equation for the second half of college is

$$V_2^c(a, \theta_c, \eta, q, \phi) = \max_{c,h,a',y} u(c, 1-h-\bar{h}) - \lambda_2(\theta_c, \phi) + \beta \mathbb{E}_{\eta'} V_3(\tilde{a}(a'), CG, \theta_c, \eta')$$

subject to

$$c + a' + p_e - s_2(q) - y + T(c, a, y) = \begin{cases} (1+r)a & \text{if } a \geq 0 \\ (1+r^s)a & \text{if } a < 0 \end{cases}$$

$$y = w^{CD} \varepsilon_2^{CD}(\theta_c, \eta) h, \;\; a' \geq -\underline{A}_2^c \;\; c \geq 0, \;\; 0 \leq h \leq 1 - \bar{h}, \;\; \eta' \sim \Pi^{CG}.$$

These individuals can work as college dropouts. At the end of the period, they complete college and acquire education level $e = CG$ and draw a new idiosyncratic productivity η' from Π^{CG}.

2.6.2 Working Stage

The Bellman equation for workers is[10]

$$V_j(a, e, \theta, \eta) = \max_{c,h,a',y} u\left(\frac{c}{1 + \mathbf{1}_{\mathcal{J}_f}\zeta}, 1-h\right) + \beta \mathbb{E}_{\eta'|\eta} V_{j+1}(a', e, \theta, \eta')$$

that fixed payments would have been made for 20 years (10 periods) after dropout and

$$\tilde{a}(a') = a' \times \frac{r^s}{1-(1+r^s)^{-10}} \times \frac{1-(1+r^-)^{-10}}{r^-}.$$

[10] After retirement, idiosyncratic labor productivity shocks are no longer a state variable. Thus the Bellman equation for the last period of workers is $V_{j_r-1}(a, e, \theta, \eta) = \max_{c,h,a',y} u(c, 1-h) + \beta V_{j_r}(a', e, \theta)$.

subject to

$$c + a' - y + T(c, a, y) = \begin{cases} (1+r)a & \text{if } a \geq 0 \\ (1+r^-)a & \text{if } a < 0 \end{cases}$$

$$y = w^e \varepsilon_j^e(\theta, \eta)h, \;\; a' \geq -\underline{A}^e \;\; c \geq 0, \;\; 0 \leq h \leq 1, \;\; \eta' \sim \pi^e(\cdot|\eta)$$

where $\mathbf{1}_{\mathcal{J}_f}$ is an indicator function that is one when individuals live with their children ($j \in [j_f, j_b - 1]$). At each period, idiosyncratic productivity η transitions according to π^e.

2.6.3 Transfer

At the end of age j_b, the individual's children become independent. The Bellman equation is

$$V_{j_b}(a, e, \theta, \eta) = \max_{c(\theta'_h), h(\theta'_h), a'(\theta'_h), y(\theta'_h)} \mathbb{E}_{\theta'_h|\theta}\{u(c(\theta'_h), 1 - h(\theta'_h)) + \tilde{V}_{j_b}(a'(\theta'_h), e, \theta, \theta'_h, \eta)\}$$

subject to

$$c(\theta'_h) + a'(\theta'_h) - y(\theta'_h) + T(c(\theta'_h), a'(\theta'_h), y(\theta'_h)) = \begin{cases} (1+r)a & \text{if } a \geq 0 \\ (1+r^-)a & \text{if } a < 0 \end{cases}$$

$$y(\theta'_h) = w^e \varepsilon_j^e(\theta, \eta)h(\theta'_h), \;\; a' \geq -\underline{A}^e \;\; c(\theta'_h) \geq 0, \;\; 0 \leq h(\theta'_h) \leq 1$$

$$\theta'_h \sim N(m + m_\theta \theta, \sigma_h^2),$$

where

$$\tilde{V}_{j_b}(a, e, \theta, \theta'_h, \eta) = \max_{b \in [0,a]} \beta \mathbb{E}_{\eta'|\eta} V_{j_b+1}(a - b, e, \theta, \eta') \\ + \nu \mathbb{E}_{\eta'', \phi} V_0(b, \theta'_h, \eta'', \tilde{q}(w^e \varepsilon^e_{j_b}(\theta, \eta)), \phi)$$

subject to

$$\eta' \sim \pi^e(\cdot|\eta), \;\; \eta'' \sim \Pi^{HS}, \;\; \phi \sim N(0, 1)$$

for all θ'_h.

At the end of the period, parents choose their transfer of wealth to their children b. Before making any decisions, parents observe their children's high school ability θ'_h, which is normally distributed with mean $m + m_\theta \theta$ and standard deviation σ_h. The high school ability of children is formed partly as a result of genetics, which leads to a correlation between parents' ability θ and children's high school ability θ'_h. In addition, as Cunha and Heckman (2007), Cunha (2013), and Daruich (2018) suggest, parents with high ability earn high income, which increases early educational investment and improves their children's high school ability.

Parents observe neither their children's initial idiosyncratic productivity η'' drawn from Π^{HS} nor their college taste ϕ drawn from the standard normal distribution. Consumption, leisure, asset holdings, and parental transfers can depend on θ'_h. The lifetime utility of their children depends on family income level q, which is a function of the potential labor income of the parents.[11]

[11]Note that parental income is not the actual labor income. Parents can control the actual labor income by adjusting their working hours. In this setting, this manipulation of parental income is not allowed and parental income is a function of "potential" income, which is labor earnings if they

2.6.4 Retirement Stage

After retirement at age j_r, individuals provide no labor. The Bellman equation is

$$V_j(a, e, \theta) = \max_{c,a'} u(c, 1) + \beta \varphi_j V_{j+1}(a', e, \theta)$$

subject to

$$c + a' = (1 + r)\varphi_{j-1}^{-1} a + p(e, \theta) - T(c, \varphi_j^{-1} a, 0)$$

$$a' \geq 0 \;\; c \geq 0.$$

The sources of income are interest payments and retirement benefits $p(e, \theta)$. In the United States, retirement benefits are determined by labor earnings before retirement (see Appendix C). To capture this, retirement benefits depend on individuals' abilities and education. The asset inflated by φ_{j-1}^{-1} reflects that the assets of expiring individuals are distributed within cohorts (perfect annuity market).

2.7 Government

The government collects tax $T(c, a, y)$ from individuals and spends the revenues on college subsidies G_e, other government consumption G_c, and retirement benefits. Government consumption G_c is exogenous and proportional to aggregate output Y

spend 35% of their time working. Thus the family income mapping is

$$\tilde{q}(w^e \varepsilon_j^e(\theta, \eta)) = \begin{cases} 1 & \text{if } w^e \varepsilon_j^e(\theta, \eta) \times 0.35 \in [0, q_1] \\ 2 & \text{if } w^e \varepsilon_j^e(\theta, \eta) \times 0.35 \in [q_1, q_2] \\ 3 & \text{else} \end{cases},$$

where q_1 and q_2 correspond to \$30,000 and \$80,000.

so that $G_c = gY$. The total budget for college subsidies is

$$G_e = \sum_{j=1,2} \int_{S_j^c} s_j(q) d\mu_j^c.$$

The tax function is

$$T(c, a, y) = \tau_c c + \tau_k r a \mathbf{1}_{a \geq 0} + \tau_l y - d\frac{Y}{N},$$

where the proportional consumption tax rate is τ_c and the proportional capital income tax rate is τ_k, which is levied only on positive net worth. The government gives a lump-sum transfer $d\frac{Y}{N}$ to each individual, where N is the measure of all individuals. This reflects the progressive income tax. τ_l is the proportional labor income tax rate.

2.8 Equilibrium

The model includes J overlapping generations and is solved numerically to characterize a stationary equilibrium in which the cross-sectional allocation is invariant. In equilibrium, individuals maximize expected lifetime utility, the representative firm and college maximize profits, the government budget is balanced each period, and prices clear all markets. A stationary equilibrium is defined in Appendix A and computation in Appendix B.

3 Calibration

This section describes how I calibrate the model. There are two sets of parameters: (1) those that are estimated outside of the model or fixed based on the literature and (2) the remaining parameters to match key moments, given the first set of parameter values.

Prices are normalized such that the average annual income of high school graduates at age 48 is $51,933.

3.1 Labor Productivity Process

I assume labor productivity as

$$\ln \epsilon_j^e(\theta, \eta) = \ln \epsilon^e + \ln \psi_j^e + \epsilon_\theta^e \theta + \ln \eta.$$

ϵ^e is the intercept of log labor productivity. I normalize $\epsilon^{HS} = \epsilon^{CG} = 1$ and calibrate ϵ^{CD} to match the wage premium of college dropouts, as explained later.

I estimate ψ_j^e, the age profile of workers at education level e, from the Panel Study of Income Dynamics (PSID; see Appendix D for sample selection and estimated age profile parameters). I use the PSID because it starts from a nationally representative cross-section and the average age of the sample does not change with the calendar year, unlike the National Longitudinal Survey of Youth (NLSY). Coefficients can vary across education levels.

I use the NLSY79 to identify the effect of ability on labor productivity ϵ_θ^e because the NLSY79 reports the Armed Forces Qualification Test (AFQT) score. I remove the age profile estimated from the PSID from the wages of each individual

	HS	CD	CG
log AFQT	.61	.74	1.31
	(.32)	(.32)	(.24)

Table 1: Estimated ability slope ϵ_θ^e of labor productivity

Source: NLSY79. See the Data Appendix for details.

in the NLSY79 and estimate the effect of ability. The ability used in the wage process differs by education level. For high school graduates and college dropouts, θ is high school ability, which is approximated by $\ln$AFQT80. I use the distribution of $\ln$AFQT80 as a targeted moment for the distribution of high school ability later. η is an idiosyncratic productivity shock uncorrelated with θ_h, and I can estimate the coefficients ϵ_θ^{HS} and ϵ_θ^{CD} using $\ln$ AFQT80. For college graduates, θ is college ability. College ability is a composite of college GPA, quality of college, college major, and other factors and is hard to measure. I instrument college ability using the law of motion connecting high school ability and college ability $\theta_c = \theta_h + \epsilon^c$. I can express log labor productivity as

$$\ln \epsilon^e + \ln \psi_j^e + \epsilon_\theta^e \theta_c + \ln \eta = \ln \epsilon^e + \ln \psi_j^e + \epsilon_\theta^e \theta_h + (\ln \eta + \epsilon_\theta^e \epsilon_c).$$

Since θ_h is uncorrelated with $\ln \eta + \epsilon_\theta^e \epsilon_c$, I can estimate the coefficient ϵ_θ^{CG} using $\ln$AFQT80 [12]. Table 1 shows the estimated coefficients on ability for each education level. As in the literature, returns to education are higher for high ability.

Finally, I estimate the process of the residual $\pi^e(\eta'|\eta)$. I assume $\pi^e(\eta'|\eta)$ is a Markov chain with two states, η_H and η_L, specific to each education level e. It has

[12] Since students with high ϵ_c are self-selected as college graduates or college dropouts, I estimate using Heckman two-step estimators.

	HS	CD	CG
ρ^e	0.9390	0.9545	0.9479
σ_η^{e2}	0.0166	0.0208	0.0248

Table 2: Estimated parameters of the residual labor productivity process

Source: PSID. See the Data Appendix for details.

exactly the same persistence and conditional variance as the AR(1) process:

$$\ln \eta' = \rho^e \ln \eta + \epsilon_\eta^e, \quad \epsilon_\eta^e \sim N(0, \sigma_\eta^{e2}).$$

After filtering out the age effects, I employ a minimum distance estimator with a fixed effect and a measurement error. Since ability should be included in the fixed effect and I do not need the ability variable, I use the PSID. I use as moments the covariances of the wage residuals at different lags and age groups, separately for each education level. In Appendix D, I discuss sample selection and details of the estimation procedures. Table 2 contains the estimated parameters.

3.2 Intergenerational Ability Transmission

To estimate the conditional mean of intergenerational ability transmission, I regressed children's ability on parents' ability in the NLSY79 to obtain the estimated value of m_θ 0.46[13]. In Appendix E, I discuss sample selection and details of the estimation procedures.

If the ability term θ for parents were the same as high school ability θ_h for everyone, parameters m and σ_h would be identified automatically outside of the

[13]For college graduates, θ is college ability, but I use $\ln$ AFQT80, as an instrument, as in the estimation of the labor productivity process.

model given the observed distribution of high school ability. In this model, however, the abilities θ of some parents are high school ability and others are college ability at age j_b, when children become independent. Since people with good college ability self-select into college graduates, the mean of the parents' ability at age j_b is not the same as the mean of high school ability. For the same reason, the standard deviation of parents' ability θ depends on the self-selection of people into college graduates and is determined inside the model. I calibrate these later to match the observed distribution of high school ability.

3.3 Subsidies and Loans

I measure the cost of education from the US Department of Education's Digest of Education Statistics. As in Jones and Yang (2016), education cost is the education and general (E&G) category, which excludes dormitories and hospitals. Education cost per student is $17,489 in 2000.

Since the federal Pell Grant program, which is the largest source of subsidies, is need-based and only a small fraction of state subsidies are merit-based (less than 18%, according to Abbott et al. (2013)), I assume that subsidies are not merit-based.

I adopt Abbott et al. (2013) for the cost for college enrollees and the subsidy system of the United States (see the third column of Table 3 for the sum of federal and state subsidies). The cost of college for enrollees is set to $\$6,710$. It follows that the government subsidizes the education sector by the difference between the full cost of education above and the cost for enrollees: $\$17,489 - \$6,710 = \$10,779$. In the model, subsidies for college enrollees $\bar{s}(q)$ are the sum of this subsidy and the subsidies to students in Table 3. In the current system, college subsidies are

q	family income	subsidies to students	subsidies to colleges	total $\bar{s}(q)$
1	- \$30,000	\$2,820	\$10,779	\$13,599
2	\$30,000 - \$80,000	\$668	\$10,779	\$11,447
3	\$80,000 -	\$143	\$10,779	\$10,922

Table 3: Subsidies and family income

Source: Subsidies to students are from Abbott et al. (2013). Subsidies to colleges are from Jones and Yang (2016).

constant across periods in college and $s_1(q) = s_2(q) = \bar{s}(q)$.

The largest federal loan program in the United States is the Federal Family Education Loan Program. Of the federal loans, the Stafford loan program is the most common for undergraduates, so I focus on Stafford loans. A Stafford loan can be either subsidized or unsubsidized, according to whether interest payments are due during college, but borrowers must pay interest after college in either case, so I focus on unsubsidized loans. Students' interest rate is the prime rate plus 2.3% ($= \iota^s$, annual). I assume that students face a borrowing limit $\bar{A}_j$ dependent on age. Annual Stafford loan limits in 2000 are \$2,625 and \$3,500 for freshmen and sophomores. The loan limit for the first half is assumed to be \$6,125 ($= \$2,625 + \$3,500$). The loan limit for the second half is \$23,000, which is the aggregate Stafford loan limit. Borrowing limits for workers are based on self-reported limits on unsecured credit by education level from the 2001 Survey of Consumer Finances.

3.4 Government Policy

Government consumption and investment over GDP in the United States in 2000 is 17.8% (Bureau of Economic Analysis). Since government expenditure on tertiary education in the United States in 2000 is 0.7% of GDP (from the OECD), g is set to $17.8\% - 0.7\% = 17.1\%$. The tax on consumption and capital income are $\tau_c = 0.07$ and $\tau_k = 0.27$, respectively (see McDaniel (2007)). τ_l is determined to balance the government budget.

3.5 Production

The elasticity of substitution is a key parameter that determines the relationship between the supply of labor and the skill premium. I set ρ so that the elasticity is 1.41, as in Katz and Murphy (1992).

3.6 Preferences

I choose $\gamma = 4$ a priori, but the coefficient of relative risk aversion is $\sigma\mu + 1 - \mu \approx 2$ —which is in the range of the literature—with the value for μ calibrated later. I set the study time $\bar{h} = 0.25$ following Babcock and Marks (2011) and Abbott et al. (2013). I set the adult equivalence ζ to 0.30 following Fernández-Villaverde and Krueger (2007) and Krueger and Ludwig (2016).

3.7 The Remaining Parameters

Given the parameter values set outside the model summarized in Table 4, I choose 27 moments in Table 5[14] and minimize the average Euclidean percentage deviation of the model from the data[15] to calibrate the 17 remaining parameters in the first

[14]Skill premiums are from full-time workers in the Current Population Survey (CPS) IPUMS (Flood, King, Rodgers, Ruggles, and Warren (2018)).

[15]For the mean of high school ability, I chose 5.04—which is the mean of $\ln$ AFQT80 in the data—for the denominator of the percent deviation. I do not take the percent deviation for the

Parameters	Interpretation	Value
γ	Coef of relative risk aversion = 2	4
$\bar{h}$	Study time	0.25
ζ	Adult equivalence scale	0.3
α	Capital share of GDP	33.3%
δ	Depreciation (annual)	7.55%
ρ	Elasticity of substitution in production 1.41	0.2908
ι^s	Stafford interest premium (annual)	2.3%
$\underline{A}_1^c$	Borrowing constraint for 1st half (Stafford loan)	\$6,125
$\underline{A}_2^c$	Borrowing constraint for 2nd half (Stafford loan)	\$23,000
$\underline{A}^{HS}$	Borrowing constraint, HS (SCF)	\$17,000
$\underline{A}^{CD}$	Borrowing constraint, CD (SCF)	\$20,000
$\underline{A}^{CG}$	Borrowing constraint, CG (SCF)	\$34,000
τ_c	Consumption tax rate	7%
τ_k	Capital income tax rate	27%
g	Gov cons to GDP ratio	17.1%

Table 4: Parameters determined outside the model

column of Table 6. Optimism is a key driver of college dropouts, and I try to match the difference between the average graduation rate students expect before enrollment and the actual graduation rate. According to Stinebrickner and Stinebrickner (2012), on average, respondents believe at the time of enrollment that there is an 86% chance of graduating, while approximately 60% of students actually graduate. The percent difference is 43%($= 0.86/0.60 - 1$), and it helps identify the mean of the bias of expectation of college ability μ_c^0. Given the parameters of optimism and observed graduation rate, the aggregate enrollment rate helps identify the standard deviation of college ability σ_c, which governs the option value of enrollment. The slopes in the enrollment rate across ability and family income help identify the slope of bias μ_c^1 and the loading of college taste at the first period λ_1^{ϕ}. Although it

enrollment and graduation rates.

would be best to use data on the slope of bias across different high school ability, college ability is different from college GPA and unobservable. I cannot identify the slope using the relation between observed high school GPA and the observed bias on college GPA. Graduation rates across ability and family income help identify the parameters of psychic cost $(\lambda, \lambda^{\theta}, \lambda_2^{\phi})$. Observed enrollment and graduation rates are taken from the NLSY97. The observed wage premiums of college graduates and college dropouts from the Current Population Survey help identify the parameters of productivity of labor (a^S, ϵ^{CD}). The moments from the 8th to 15th rows in Table 5 are associated with the education cost κ; utility parameters (μ, β, v); lump-sum transfer d; overseeing cost ι; and intergenerational ability parameters (m, σ_h) for each.[16] I normalize the mean of high school ability to zero without loss of generality.

The third column of Table 6 presents the calibrated values. The calibrated value of μ_c^0 is positive and enrollees are optimistic about their college ability on average. Since the standard deviation of college ability is 0.40[17], the bias for the mean ability is 48% of the standard deviation of college ability. In addition, μ_c^1 is negative, and enrollees with lower high school ability are more optimistic than enrollees with higher high school ability. These characteristics are consistent with the pattern of the bias of college GPA observed in Stinebrickner and Stinebrickner (2012).

λ^0 is positive and agents derive psychic cost from college. The average monetary value of the psychic cost from attending two periods of college is \$208,880[18].

[16] The transfer from parents is taken from Daruich (2018) using the PSID. The ratio of pre-tax to post-tax income is from Heathcote, Perri, and Violante (2010). The share of borrowers is from Abbott et al. (2013) using the Survey of Consumer Finances 2001.

[17] The square root of the sum of the variance of high school ability and σ_c^2. $0.40 = \sqrt{0.217^2 + 0.340^2}$.

[18] The monetary value of the psychic cost is calculated as the net present value of consumption

Moment	Model	Data
Expected/Actual graduation rate -1	0.431	0.433
Enrollment rate of ability quartile	(figure)	(figure)
Graduation rate of ability quartile	(figure)	(figure)
Enrollment rate of family income quartile	(figure)	(figure)
Graduation rate of family income quartile	(figure)	(figure)
Skill premium for CG	90.8%	90.2%
Skill premium for CD	19.6%	19.9%
Education cost/mean income at 48	0.320	0.33
Hours of work	33.8%	33.3%
K/Y	1.298	1.325
Transfer/mean income at 48	67.0%	66%
Log pre-tax/post-tax income	61.2%	61%
Borrowing non-retirees older than age $j = 2$	6.59%	6.8%
Mean of AFQT	-0.0135	0
Standard deviation of AFQT	0.217	0.213

Table 5: Moments matched and model fit

It is comparable in magnitude with the psychic cost reported in Cunha et al. (2005). A negative λ^1 implies that the psychic cost is smaller for agents with high ability.

The standard deviation of college ability is 0.40 and 90% greater than that of high school ability. There is much uncertainty regarding college ability given high school ability. This leads to a high option value of enrollment, which is consistent with Stange (2012) and Trachter (2015).

Model fit is presented in Table 5 and Figures 1 and 2. In general, the model fits

loss c_λ satisfying

$$\sum_{j=1}^{J} \tilde{\beta}^{j-1} u(\bar{C} - c_\lambda, \bar{L}) = \sum_{j=1}^{J} \tilde{\beta}^{j-1} u(\bar{C}, \bar{L}) + \lambda(0,0) + \beta\lambda(0,0),$$

where $\bar{C}$ and $\bar{L}$ are the aggregate consumption and leisure per individual and $\lambda(0,0)$ is the psychic cost of college attendance for the people with the average college ability ($\theta_c = 0$) and average taste ($\phi = 0$).

Parameter	Description	Value
μ_c^0	college ability bias intercept	0.190
μ_c^1	college ability bias slope	-0.409
λ	psychic cost intercept	23.2
λ^θ	psychic cost slope	-241
λ_1^ϕ	first period college taste	64.1
λ_2^ϕ	second half college taste	41.3
a^S	productivity of skilled labor	0.457
ϵ^{CD}	productivity of CD	1.02
σ_c	s.d. of college ability	0.340
κ	education cost	0.226
μ	consumption share of preference	0.418
β	time discount rate	0.938
v	altruism	0.0948
d	lump-sum transfer ratio	0.125
ι	borrowing wedge ($r^- = r + \iota$)	18.0%
m	intergenerational ability transmission intercept	-0.0471
σ_h	intergenerational ability transmission s.d.	0.171

Table 6: The remaining parameters

well, considering the over-identification of 17 parameters against 27 moments. In the data, ability is correlated with enrollment and graduation more than family income, and the model captures this pattern. Although graduation rates across family income are somewhat flatter than the data, they capture the key pattern. Enrollment and graduation rates are higher for the second quartile than for the third quartile in the model, because there are only three bins for family income q and there is a jump in college subsidies when people cross the threshold of family income.

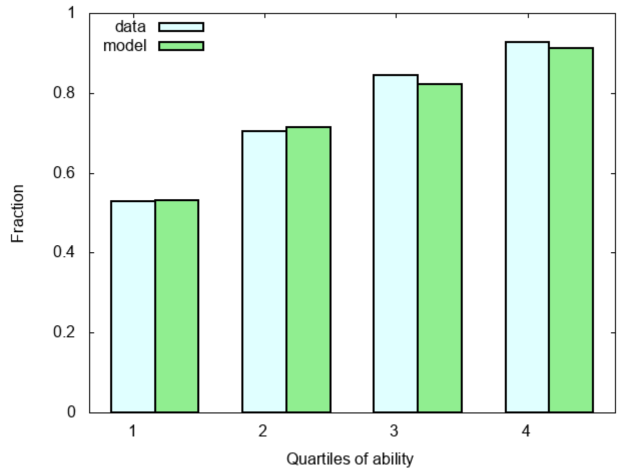
3.8 Validation Exercises

In this subsection, I compare the model simulation with the data for the moments I do not target when calibrating.

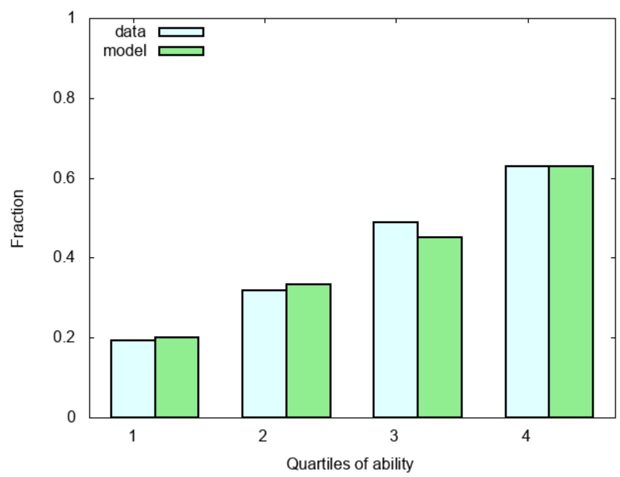
The Partial Equilibrium Effect of Year-invariant Subsidies

The elasticity of enrollment with regard to tuition or subsidies has been examined in the micro empirical literature. I simulate the partial equilibrium response of enrollment to a $1,000 increase in subsidies for all years in college and family income evenly, and compare with estimates from the empirical literature. All prices and the distribution at age 1 are fixed at the current level, and these additional subsidies are given to only one generation.

The aggregate enrollment rate of the affected generation increases by 1.05 percentage points in the simulation. The micro emprirical literature has estimates of the effect of subsidies on enrollment by Dynarski (2002), Kane (1994), and Cameron and Heckman (2001). While this literature argues that the enrollment rate of groups that benefit from an additional subsidy of $1,000 increases by between 3 and 6 percentage points, Hansen (1983) and Kane (1994) argue that there is less evidence of



(a) Enrollment rate

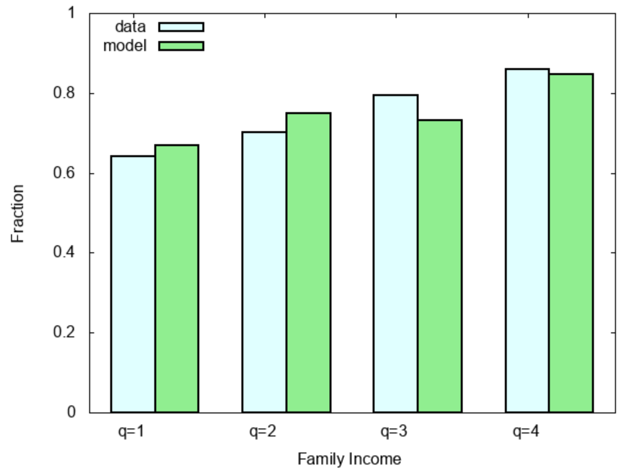


(b) Graduation rate

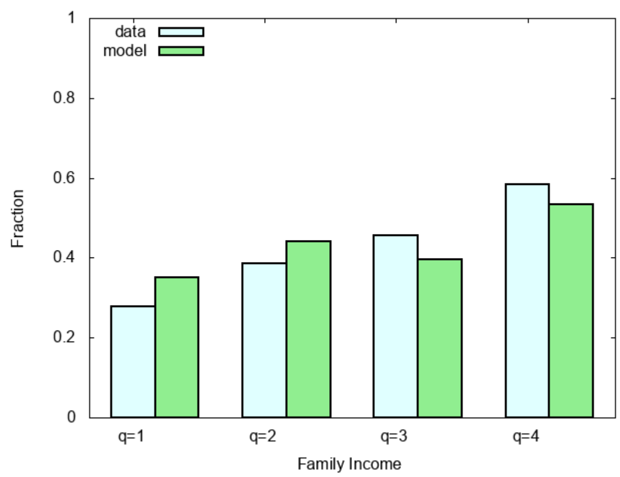
Figure 1: Model fit: Enrollment and graduation rates for each ability quartile

Data Source: NLSY97. I use the sample of only 25-year-old high school graduates. Ability is the log of AFQT score using the definition from the NLSY79. Scores are adjusted by age, as in Altonji et al. (2012) and Castex and Kogan Dechter (2014).

a rise in college enrollment of targets of the Pell Grant program (see Kane (2006) for a survey of the literature). Therefore the simulation is broadly in the range of the literature. In addition, the increase in enrollment is smaller in the model than



(a) Enrollment rate



(b) Graduation rate

Figure 2: Model fit: Enrollment and graduation rates for each family income quartile

Data Source: NLSY97. I use the sample of only 25-year-old high school graduates. Family income is defined as the average of parental income at 16 and 17 if both are available. I use one if both are not available.

in the data, which implies that this calibration is a more conservative choice. If the response of enrollment were high, the effect of changing college subsidies would also be high and I would overestimate the effect of switching to year-dependent subsidies.

In the simulation, the share of college graduates increases by 0.45 percentage points and that of college dropouts increases by 0.60 percentage points. This is consistent with Dynarski (2008), Castleman and Long (2016), Scott-Clayton (2011), Scott-Clayton and Zafar (2019), Denning, Marx, and Turner (2017), and Bettinger, Gurantz, Kawano, Sacerdote, and Stevens (2019), who all find a positive effect of subsidies on graduation.

The Sluggish Increase in College Graduates

The increase in college graduates has been sluggish in the United States since 1980, despite the increase in the skill premium during the same period. In this subsection, I examine how well the model can explain this sluggish increase by targeting the skill premiums of 1980 and 2000 in the United States. The benchmark calibration is targeted to the United States in 2000, and I assume that only the productivity of skilled labor a^S and productivity of college dropouts ϵ^{CD} change in the model between 1980 and 2000. In particular, I set the values of a^S and ϵ^{CD} to match the college graduate wage premium of 46.2% and the college dropout wage premium of 12.1%, as observed in 1980 in the United States, with the other parameter values fixed. I compute the steady state with the new values to replicate 1980 in the United States. The first two rows of Table 7 show the wage premiums for college graduates and dropouts in the model and the data. By definition, the changes in the model and the data are the same. I compare the change in the share of college graduates and dropouts with the data.[19]

The third and fourth rows of Table 7 show the change in the share of college

[19] I use the Current Population Survey IPUMS (Flood et al. (2018)) for the wage premiums in 1980 and the change in the shares of college graduates and dropouts between 1980 and 2000.

	1980	2000	change (model)	change (data)
college graduate premium	46.2%	90.9%	44.7pp	43.2pp
college dropout premium	12.1%	19.6%	7.5pp	7.4pp
share of college graduates	28.0%	32.9%	4.9pp	4.98pp
share of college dropouts	42.8%	41.3%	-1.5pp	2.41pp

Table 7: Change in the share of college graduates and dropouts

graduates and dropouts between 1980 and 2000. As the college graduate premium increases by 44.7 percentage points from 1980 to 2000, the third column shows that the share of college graduates increases by 4.9 percentage points. The model can explain the sluggish increase in the share of college graduates.

College dropouts increase in the data while decreasing in the model. In the model, the increase in the college graduate wage premium induces more people to graduate from college than the increase in enrollment. Bound, Lovenheim, and Turner (2010) document the drop in the graduation rate in the United States and argue that the supply side of higher education played a key role. For example, they argue that the resources of colleges—either expenditures per student or student-faculty ratios—decreased, which also reduced the graduation rate. Moreover, there is a shift to community college as an initial institution, which has a lower completion rate than other types of institutions. This paper abstracts from changes on the supply side of higher education. The exercise of this paper is changing the college subsidy system given the resources of colleges per student.

4 Results

This section is composed of three exercises. In the first exercise, I increase overall spending without changing the structure of subsidies, financed by the proportional labor income tax, and examine how this affects enrollment, graduation, and the skill premium. In the second exercise, I keep total spending fixed but choose subsidies by year to maximize the number of college graduates in the stationary equilibrium and compare the effect with the first exercise. In the third exercise, I keep total spending fixed and choose subsidies by year to maximize the utilitarian social welfare function in the stationary equilibrium as a normative analysis.

4.1 The Effect of Year-invariant Subsidies

As a benchmark case, I examine the general equilibrium effect of a permanent change in the total budget of the current college subsidy scheme, which I call year-invariant subsidies. Table 8 shows how the enrollment rate, defined as the number of college enrollees, the number of college graduates, and the skill premium change, as the government subsidies budget increases. $\bar{G}_e$ denotes the current level of total government budget for college subsidies. I consider an increase in the budget for college subsidies G_e from $0.75\bar{G}_e$ to $2\bar{G}_e$. The proportional labor income tax rate τ_l is adjusted to the change in the budget. Subsidies across college years and family income proportionally change with G_e fixed.

Both the enrollment rate and the share of college graduates increase as the total budget increases. Since skilled and unskilled labor are incomplete substitutes and the supply of skilled labor increases, the skill premium decreases. As the total bud-

G_e	0.75 $\bar{G}_e$	$\bar{G}_e$	1.5$\bar{G}_e$	2$\bar{G}_e$
enrollment rate	72.7%	74.2%	77.2%	77.8%
share of college graduates	32.1%	32.9%	34.2%	35.0%
skill premium	95.0%	90.9%	82.8%	78.3%
τ_ℓ	36.0%	36.2%	36.6%	37.0%

Table 8: Education and the skill premium for the current college subsidy scheme

get of college subsidies increases, the labor income tax rate to balance the budget increases.

4.2 The Effect of Year-dependent Subsidies

In this subsection, I derive the year-dependent subsidies that maximize the number of college graduates and show how they affect enrollment, graduation, and the skill premium. I fix total spending at the current level $\bar{G}_e$, and only allow the relative sizes of subsidies to differ across college years. The maximization problem is formulated as

$$\max_{g_1>0, g_2>0, \tau_\ell} \int_{S_2^c} d\mu_2^{CG}$$

subject to

$$\int_{S_1^c} g_1 \bar{s}(q) d\mu_1^c + \int_{S_2^c} g_2 \bar{s}(q) d\mu_2^c = \bar{G}_e$$

and the government budget constraint. The new subsidies are $s_1(q) = g_1\bar{s}(q)$ and $s_2(q) = g_2\bar{s}(q)$, where $\bar{s}(q)$ is the current college subsidy system. In this problem, the government chooses the general levels of college subsidies g_j for each period j compared with the current system. If I increase subsidies for the first half of college

$s_j(q)$	year-invariant $\bar{G}_e$	year-dependent $\bar{G}_e$
$s_1(1)$	\$13,599	\$4
$s_1(2)$	\$11,447	\$4
$s_1(3)$	\$10,922	\$3
$s_2(1)$	\$13,599	\$42,436
$s_2(2)$	\$11,447	\$35,720
$s_2(3)$	\$10,922	\$34,082

Table 9: Year-dependent subsidies maximizing the number of college graduates

g_1, the general level of subsidies for the second half g_2 must decrease.[20]

Note that I do not allow changes in relative subsidies across different family income and ability within a year in college from the current system. For example, the ratio of subsidies for $q = 1$ and $q = 2$ in the first half of college is fixed at the current state. This paper focuses on how the year-dependency of college subsidies affects educational choices and abstracts from analysis of the optimal need-based and merit-based subsidies[21].

Table 9 shows the annual amount of subsidies in the two cases. The first column is identical to the case of year-invariant subsidies with $G_e = \bar{G}_e$. The second column is the year-dependent subsidies that maximize the number of college graduates. The first three rows are college subsidies at the first half of college across family income level $q = 1$ to 3 from the top to the bottom. The next three rows are college subsidies in the second half. Optimal year-dependent subsidies are back-loaded: Subsidies are more generous for the second half than for the first half. Optimal year-dependent subsidies for the first half are negligible[22].

[20]Since the composition of education level and labor productivity of workers changes under the new subsidy system, aggregate labor income changes and τ_ℓ must be adjusted to balance the government budget, even though the budget for college subsidies is fixed.

[21]See Findeisen and Sachs (2016) for the optimal need-based and merit-based subsidies.

[22]Note that I search only the positive values of g_1 and g_2 for the maximization problem, and the

year-invariant/dependent	invariant $\bar{G}_e$	dependent $\bar{G}_e$	invariant $1.5\bar{G}_e$
enrollment rate	74.2%	68.7%	77.2%
share of college graduates	32.9%	34.5%	34.2%
skill premium	90.9%	82.6%	82.8%
τ_ℓ	36.2%	36.3%	36.6%

Table 10: Education and the skill premium for the optimal subsidy scheme

Table 10 displays the enrollment rate, share of college graduates, and skill premium for each case. Year-dependent subsidies reduce the enrollment rate by 5.5 percentage points and increase the share of college graduates by 1.6 percentage points.[23] In contrast, the third column presents the year-invariant case of $G_e = 1.5\bar{G}_e$ again. Year-dependent subsidies increase the share of college graduates and reduce the skill premium more than the case in which the total budget for year-invariant subsidies is increased by 50%. Changing the structure of college subsidies is as effective as increasing the budget by 50%. The required increase in the tax rate for the year-dependent subsidies is 0.1%, while the tax rate must increase by 0.4% in the case of increasing the budget by 50%.

The mechanism of the effect of year-dependent subsidies is as follows. In the current system, over 70% of those who graduate from high school enroll in college. Increasing enrollment will basically encourage more people to enroll who are likely to drop out. This means that the enrollment margin is not so important from the perspective of getting people to graduate. The marginal person who drops out is better able to benefit from college than the marginal person who does not enroll. It is

negligibly positive amounts of the year-dependent subsidies for the first period are computational artifacts.

[23] If the prices and wages are fixed at the current level, the enrollment rate and share of college graduates are 68.8% and 34.9% each. The general equilibrium effect offsets the increase in the share of college graduates by 20%. Considering the general equilibrium effect, therefore, is important.

easier to create incentives for the marginal dropout to finish than to create incentives for the marginal non-enrollee to enroll and finish. Decreasing subsidies for the first period serves mainly to discourage people who are unlikely to graduate from enrolling. Higher subsidies for the second period encourages marginal dropouts to finish.

There is another mechanism of the back-loaded subsidies. In the current system, the government has paid subsidies to all of the people who enroll but later drop out. With back-loaded subsidies, the government does not need to pay high subsidies to people who drop out before the second period and can give more subsidies to the people who graduate, which increases the number of college graduates. As Table 9 shows, the sum of the subsidies to college graduates for the two periods for the middle family income is \$35,724 (= \$4+\$35,720), which is higher than the case of the current system \$22,894 (= \$11,447+\$11,447). Back-loaded subsidies are more cost-effective from the perspective of increasing the number of college graduates.

4.3 Welfare Analysis of Year-dependent Subsidies

In this subsection, I examine how year-dependent subsidies can improve welfare. I fix the total budget for college subsidies at the current level and examine how the utilitarian social welfare function improves by only varying the relative sizes of subsidies across college years.[24] The optimization problem is

$$\max_{g_1>0,g_2>0,\tau_\ell} \sum_j N_j \left(\int V_j(\boldsymbol{s}_j) d\bar{\mu}_j(\boldsymbol{s}_j) + \int V_j^c(\boldsymbol{s}_j^c) d\bar{\mu}_j(\boldsymbol{s}_j^c) \right)$$

[24]The utilitarian social welfare function considered here is the sum of lifetime utility of all of the existing agents. The optimal policy is similar even if I maximize the sum of lifetime utility of agents only at age 1.

	Current state	Optimal
$s_1(1)$	\$13,599	\$10,721
$s_1(2)$	\$11,447	\$9,025
$s_1(3)$	\$10,922	\$8,611
$s_2(1)$	\$13,599	\$19,858
$s_2(2)$	\$11,447	\$16,716
$s_2(3)$	\$10,922	\$15,949

Table 11: Year-dependent subsidies maximizing social welfare.

subject to

$$\int_{S_1^c} g_1 \bar{s}(q) d\mu_1^c + \int_{S_2^c} g_2 \bar{s}(q) d\mu_2^c = G_e$$

and the government budget constraint. N_j is the relative population of age j calculated from survival rates φ_j. I assume that the government that implements the optimal policy recalculates the expected lifetime utility without excessive optimism regarding college ability, which is different from the lifetime utility agents expect before enrollment.

In Table 11, optimal college subsidies are back-loaded and the amount for the second half is twice the subsidy for the first half. The first and second rows of Table 12 show that enrollment decreases by 0.4 percentage points and the share of college graduates increases by 0.7 percentage points by switching to the optimal policy. The skill premium decreases by 3.6 percentage points.

To examine the welfare gain under the optimal policy, I use lifetime consumption equivalence as a summary measure of welfare. Let $\tilde{V}_j(\boldsymbol{c}, \boldsymbol{h}; \boldsymbol{s}_j)$ be expected lifetime utility at age j with the path of consumption $\boldsymbol{c}$ and leisure $\boldsymbol{h}$ with the state $\boldsymbol{s}_j$, in which agents in the initial period have no optimism when calculating the

	Current state	Optimal
share of college enrollees	74.2%	73.8%
share of college graduates	32.9%	33.6%
skill premium	90.9%	87.3%
τ_ℓ	36.2%	36.2%
welfare gain		+0.15%

Table 12: The effect of optimal year-dependent subsidies.

lifetime value. Then lifetime consumption equivalence is defined as ω_{tot}, such that

$$\sum_j N_j \int_{S_0} \tilde{V}_j(\boldsymbol{c}^B, \boldsymbol{h}^B; \boldsymbol{s}_j) d\mu_j^B = \sum_j N_j \int_{S_0} \tilde{V}_j((1+\omega_{tot})\boldsymbol{c}^A, \boldsymbol{h}^A; \boldsymbol{s}_j) d\mu_j^A$$

where $\boldsymbol{c}^A$, $\boldsymbol{h}^A$, and μ_j^A are the consumption path, the leisure path, and the measure of the current state and $\boldsymbol{c}^B$, $\boldsymbol{h}^B$, and μ_j^B are the ones after the change in policies.

The fifth row of Table 12 shows that the lifetime consumption equivalence of the optimal year-dependent subsidies is 0.15%. Although the size is modest, we do not need to increase the total budget for the year-dependent subsidies. As the fourth row of Table 12 shows, the tax rate does not need to increase to get the welfare gain, unlike increasing the size of year-invariant subsidies.

To examine why welfare improves, I define the lifetime consumption equivalence only of newborns in the economy and decompose it into three parts, as in Benabou (2002): (i) *a level effect* that measures the gain in aggregate consumption, leisure, and the psychic cost of college attendance; (ii) *an uncertainty effect* that measures the effect of the volatility of consumption and leisure paths across states and over time on the utility of risk-averse agents; and (iii) *an inequality effect* that

	Total	Level	Uncertainty	Inequality
Optimal	$+0.07\%$	$+0.15\%$	$+0.04\%$	-0.09%

Table 13: Welfare decomposition.

measures the distribution at the beginning of age 1.[25] I follow the decomposing process of Abbott et al. (2013).

Table 13 shows that the total welfare gain for newborns is 0.07%. There is a positive level effect of 0.15%. While output, capital, and consumption do not change (see the first three rows of Table 14), psychic cost is a key factor. In the current system, individuals are overly optimistic and there is an excessively large amount of college enrollees. Optimal back-loaded subsidies screen people who enroll and reduce enrollment. By reducing subsidies for the first half, the marginal enrollee with low ability stops enrolling, which reduces their psychic cost. Reducing psychic cost for people with low ability has a significant effect on welfare. This is consistent with Cunha et al. (2005), who argue that psychic cost is a sizable component of lifetime utility. The welfare implication of the back-loaded subsidies depends largely on optimism. In Appendix F, I calibrate the case without optimism and examine how the assumption about optimism matters for the optimal policy.

The uncertainty effect is 0.04%, as there is less uncertainty under the optimal policy. Due to a smaller skill premium, there is less difference in wages between college graduates and dropouts. The policy can reduce the uncertainty of lifetime income from dropout decisions by the risk of college ability.

The inequality effect is -0.09%, as there is more inequality across heterogeneous agents at age 1 under the optimal policy. This is counterintuitive, because

[25]The sum of these three effects is not necessarily the total welfare effect.

	Current state	Optimal
Y	0.318	0.318
K	0.413	0.413
C	0.211	0.211
w^S	0.355	0.352
w^U	0.405	0.408
std c	0.129	0.129
std a	0.478	0.475
std h	0.0834	0.0833
std wage	0.544	0.540

Table 14: Aggregates under optimal year-dependent subsidies

	$q = 1$	$q = 2$	$q = 3$
$\theta = 1$	$+0.6\%$	$+0.1\%$	$+0.5\%$
$\theta = 2$	$+0.2\%$	-0.4%	$+0.5\%$
$\theta = 3$	-0.8%	-0.3%	$+0.5\%$
$\theta = 4$	-0.9%	-0.0%	$+0.4\%$

Table 15: Lifetime consumption equivalence variation for newborns

the difference in wages of skilled and unskilled labor decreases and the standard deviations of consumption, assets, hours, and wages per hour also do not increase, as shown in Table 14. Although inequality as of period 1 increases, cross-sectional inequality in the economy does not increase under the optimal policy.

To see why inequality at age 1 increases, I calculate the welfare gain for each ability and family income level in Table 15[26].

Given family income, the welfare gain is large for those with low ability who are unlikely to graduate from college. Since the price of unskilled labor increases, as in Table 14, the welfare of agents with low ability increases more than other

[26]The distribution of ability differs in the current state and the optimal case, because the share of college graduates changes the mean ability of the future generation. Each ability quartile in the table is the quartile of the current state.

	% of subsidy loss
Subsidies	-100%
Labor income	+24%
(Price of an hour of working)	+13%
(Leisure)	(-0.061)
Transfer from parents	+0.03%
Reducing savings	+65%
Less tuition	+4%
Consumption	-7%

Table 16: Change in each item of income

agents.

One exception is middle family income $q = 2$, where middle ability people lose the most. Under optimal back-loaded subsidies, high school graduates gain welfare because the price of unskilled labor increases, college dropouts lose welfare because subsidies for the first half decrease, and the effects are ambiguous for college graduates because total subsidies over the two periods in college increase, but the price of skilled labor decreases. Since middle ability people are likely to be college dropouts, they lose the most welfare.

Given the same ability level, the welfare gain is greater for high family income, which is consistent with the negative inequality effect[27]. College enrollees from poor families get less transfer from parents, and the borrowing constraint for the first period ($\$6,125$) is tighter than for the second period ($\$23,000$). It follows that reducing subsidies for the first half can reduce consumption by agents from poor families during the first period of college.

In order to see how people from poor families adjust to the loss of college sub-

[27] While the welfare loss of agents from poor families ($q = 1$) is large, the fraction of poor families is only 6% and the contribution to the social welfare is small.

sidies for the first period, Table 16 shows the average change in each part of income and consumption for an individual with $\theta_h = 0$ (mean ability), $q = 1$, $\eta = \eta_H$, and $\phi = 0$ at the first half of college. The loss of subsidies for the first half of college does not lead to the same amount of loss of consumption. Labor income increases and covers 24% of the loss of college subsidies. First, they provide unskilled labor during the first half of college and the price of unskilled labor increases. Second, agents work for longer hours to mitigate the loss of college subsidies. As the fourth row shows, they cut their leisure by 0.061 out of the unit hour endowment. These results are consistent with the findings of Keane and Wolpin (1997) and Garriga and Keightley (2007).

Next, since college subsidies are shifted to the second half of college, they reduce savings for the second half, which covers 65% of the loss of college subsidies. Tuition decreases due to the lower price of skilled labor, which covers 4% of the loss of college subsidies. In total, agents can mitigate the loss of college subsidies by 93% $(= 100\% - 7\%)$ for consumption. Since they are from poor families, there is a negligible increase in transfers from parents.

4.4 Correcting Bias

A large part of the welfare gain of back-loaded subsidies originates in excessive optimism about college ability. If the government can provide information to students to correct the bias on college ability before the enrollment decision, it can improve welfare and we might not need to rely on back-loaded subsidies. In this subsection, I show what the welfare gain is by correcting bias and compare it with back-loaded subsidies without correcting bias.

	Current state	Correcting bias	Optimal
share of college enrollees	74.2%	45.5%	45.8%
share of college graduates	32.9%	26.2%	26.0%
skill premium	90.9%	124%	125%
τ_ℓ	36.2%	36.4%	36.4%
welfare gain		-9.28%	-9.25%

Table 17: Education and the skill premium when correcting bias

The second column of Table 17 shows a welfare loss from correcting bias with the current subsidies. The enrollment rate drops significantly. Without optimism, enrollment is excessively small because there is a borrowing constraint and no insurance available for the risk of college ability. In the current system, while optimism leads to excessive enrollment, optimism also cancels out the effect of the tight credit limit or the absence of insurance for the risk of college ability. In total, the loss from a large skill premium and excessively small enrollment is higher than the gain from avoiding the excessive enrollment due to optimism.[28]

To consider combining year-dependent subsidies with correcting bias, I solve the optimal policy problem in Section 4.3 without bias—that is, $\mu_c(\theta_h) = 0$ for all θ_h. The solution is shown in the second column of Table 18. The optimal subsidy is *front-loaded*: greater for the first period than the second period. As shown in the third column of Table 17, by subsidizing college in the first period, the policy can increase enrollment. However, the welfare gain is still negative compared to the current case. Using optimal back-loaded subsidies is more beneficial than correcting bias. While optimism is likely to induce excessive college enrollment, it

[28] Correcting bias reduces the initial expected value agents have in mind, even with the allocation fixed. However, this is not the origin of the welfare loss of correcting bias. The welfare of the optimal policy is calculated by the government, which does not have optimism even before correcting bias.

	Current state	Optimal
$s_1(1)$	\$20,344	\$21,750
$s_1(2)$	\$17,124	\$18,308
$s_1(3)$	\$16,339	\$17,469
$s_2(1)$	\$20,344	\$17,808
$s_2(2)$	\$17,124	\$14,990
$s_2(3)$	\$16,339	\$14,302

Table 18: Optimal subsidies with correcting bias

can also avoid excessively small enrollment.

5 Conclusion

The skill premium has been expanding in the United States, and policymakers often consider educational subsidies as a tool to increase college enrollment and decrease inequality. However, enrollment does not necessarily lead to graduation, and it is important to understand how policy can affect graduation. This paper quantitatively assesses the effects of year-dependent subsidies on enrollment, graduation, and the skill premium compared to year-invariant subsidies. Switching to back-loaded subsidies, with the total budget fixed, can increase the fraction of college graduates and reduce the skill premium more than the case in which year-invariant subsidies are increased by 50%. Back-loaded subsidies improve welfare without increasing the total budget for college subsidies and increasing taxes. As a result of back-loaded subsidies, enrollment decreases. This result shows that a policy that increases enrollment is not necessarily welfare-improving, and that we should distinguish the effect of policy on enrollment from graduation.

The welfare implication of this paper is also important in terms of evaluating existing policies that have year-dependent aspects. Some universities offer free tuition for the final semester if the person graduates in four years, and these are back-loaded subsidies. This result suggests that such a policy is welfare-improving.

Although this paper focuses on college, the mechanics are applicable to other education levels such as postgraduate education. Increasing subsidies for postgraduate education might lead to an increase in workers with higher education, which affects the distribution of skill and wages. Varying the amount of subsidies by years within earlier education might have similar effects. Furthermore, age-dependent subsidies to human capital investment after finishing schooling could be beneficial under a similar mechanism. Subsidies that are dependent on education levels and age have the potential to be an important policy tool.

References

Abbott, Brant, Giovanni Gallipoli, Costas Meghir, and Giovanni L Violante. 2013. Education policy and intergenerational transfers in equilibrium. Tech. rep., National Bureau of Economic Research.

Akyol, Ahmet and Kartik Athreya. 2005. Risky higher education and subsidies. *Journal of Economic Dynamics and Control* 29, no. 6:979–1023.

Altonji, Joseph G. 1993. The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics* 11, no. 1, Part 1:48–83.

Altonji, Joseph G, Prashant Bharadwaj, and Fabian Lange. 2012. Changes in the characteristics of american youth: Implications for adult outcomes. *Journal of Labor Economics* 30, no. 4:783–828.

Archibald, Robert B and David H Feldman. 2014. *Why does college cost so much?* Oxford University Press.

Arcidiacono, Peter. 2004. Ability sorting and the returns to college major. *Journal of Econometrics* 121, no. 1-2:343–375.

Babcock, Philip and Mindy Marks. 2011. The falling time cost of college: Evidence from half a century of time use data. *Review of Economics and Statistics* 93, no. 2:468–478.

Benabou, Roland. 2002. Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica* 70, no. 2:481–517.

Bettinger, Eric, Oded Gurantz, Laura Kawano, Bruce Sacerdote, and Michael Stevens. 2019. The long-run impacts of financial aid: Evidence from california's cal grant. *American Economic Journal: Economic Policy* 11, no. 1:64–94.

Bound, John, Michael F Lovenheim, and Sarah Turner. 2010. Why have college completion rates declined? an analysis of changing student preparation and collegiate resources. *American Economic Journal: Applied Economics* 2, no. 3:129–57.

Bovenberg, A Lans and Bas Jacobs. 2005. Redistribution and education subsidies are siamese twins. *Journal of Public Economics* 89, no. 11-12.

Cameron, Stephen V and James J Heckman. 2001. The dynamics of educational attainment for black, hispanic, and white males. *Journal of Political Economy* 109, no. 3:455–499.

Castex, Gonzalo and Evgenia Kogan Dechter. 2014. The changing roles of education and ability in wage determination. *Journal of Labor Economics* 32, no. 4:685–710.

Castleman, Benjamin L and Bridget Terry Long. 2016. Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation. *Journal of Labor Economics* 34, no. 4:1023–1073.

Caucutt, Elizabeth M and Krishna B Kumar. 2003. Higher education subsidies and heterogeneity: A dynamic analysis. *Journal of Economic Dynamics and Control* 27, no. 8:1459–1502.

Chatterjee, Satyajit and Felicia Ionescu. 2012. Insuring student loans against the financial risk of failing to complete college. *Quantitative Economics* 3, no. 3:393–420.

Cunha, Flavio. 2013. Investments in children when markets are incomplete. *Review of Economic Studies* .

Cunha, Flavio and James Heckman. 2007. The technology of skill formation. *American Economic Review* 97, no. 2:31–47.

Cunha, Flavio, James Heckman, and Salvador Navarro. 2005. Separating uncertainty from heterogeneity in life cycle earnings. *oxford Economic papers* 57, no. 2:191–261.

Daruich, Diego. 2018. The macroeconomic consequences of early childhood development policies. *FRB St. Louis Working Paper* , no. 2018-29.

Denning, Jeffrey T, Benjamin M Marx, and Lesley J Turner. 2017. Propelled: The effects of grants on graduation, earnings, and welfare. Tech. rep., National Bureau of Economic Research.

Dynarski, Susan. 2002. The behavioral and distributional implications of aid for college. *American Economic Review* 92, no. 2:279–285.

———. 2008. Building the stock of college-educated labor. *Journal of Human Resources* 43, no. 3:576–610.

Fernández-Villaverde, Jesús and Dirk Krueger. 2007. Consumption over the life cycle: Facts from consumer expenditure survey data. *The Review of Economics and Statistics* 89, no. 3:552–565.

Findeisen, Sebastian and Dominik Sachs. 2016. Optimal need-based financial aid. Tech. rep., Mimeo, University of Mannheim/European University Institute.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren. 2018. Integrated public use microdata series, current population survey: Version 6.0. [database]. *Minneapolis: MN: IPUMS* .

Garriga, Carlos and Mark P Keightley. 2007. A general equilibrium theory of college with education subsidies, in-school labor supply, and borrowing constraints .

Goldin, Claudia Dale and Lawrence F Katz. 2009. *The race between education and technology*. Harvard University Press.

Hansen, W Lee. 1983. Impact of student financial aid on access. *Proceedings of the Academy of Political Science* 35, no. 2:84–96.

Hanushek, Eric A, Charles Ka Yui Leung, and Kuzey Yilmaz. 2004. Borrowing constraints, college aid, and intergenerational mobility. Tech. rep., National Bureau of Economic Research.

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L Violante. 2010. Unequal we stand: An empirical analysis of economic inequality in the united states, 1967–2006. *Review of Economic dynamics* 13, no. 1:15–51.

Horn, Laura and Paul Skomsvold. 2011. Community college student outcomes: 1994-2009. *Washington, DC: National Center for Education Statistics* .

Ionescu, Felicia. 2011. Risky human capital and alternative bankruptcy regimes for student loans. *Journal of Human Capital* 5, no. 2:153–206.

Jones, John Bailey and Fang Yang. 2016. Skill-biased technical change and the cost of higher education. *Journal of Labor Economics* 34, no. 3:621–662.

Kane, Thomas J. 1994. College entry by blacks since 1970: The role of college costs, family background, and the returns to education. *Journal of political Economy* 102, no. 5:878–911.

———. 2006. Public intervention in post-secondary education. *Handbook of the Economics of Education* 2:1369–1401.

Kane, Thomas J and Cecilia Elena Rouse. 1995. Labor-market returns to two-and four-year college. *The American Economic Review* 85, no. 3:600.

Katz, Lawrence F and Kevin M Murphy. 1992. Changes in relative wages, 1963–1987: supply and demand factors. *The Quarterly Journal of Economics* 107, no. 1:35–78.

Keane, Michael P and Kenneth I Wolpin. 1997. The career decisions of young men. *Journal of Political Economy* 105, no. 3:473–522.

Kroch, Eugene A and Kriss Sjoblom. 1994. Schooling as human capital or a signal: some evidence. *Journal of Human Resources* :156–180.

Krueger, Dirk and Alexander Ludwig. 2016. On the optimal provision of social insurance: Progressive taxation versus education subsidies in general equilibrium. *Journal of Monetary Economics* 77:72–98.

Manski, Charles F. 1989. Schooling as experimentation: a reappraisal of the post-secondary dropout phenomenon. *Economics of Education review* 8, no. 4:305–312.

McDaniel, Cara. 2007. Average tax rates on consumption, investment, labor and capital in the oecd 1950-2003. *Manuscript, Arizona State University* 19602004.

Scott-Clayton, Judith. 2011. On money and motivation a quasi-experimental analysis of financial incentives for college achievement. *Journal of Human resources* 46, no. 3:614–646.

Scott-Clayton, Judith and Basit Zafar. 2019. Financial aid, debt management, and socioeconomic outcomes: Post-college effects of merit-based aid. *Journal of Public Economics* .

Stange, Kevin M. 2012. An empirical investigation of the option value of college enrollment. *American Economic Journal: Applied Economics* 4, no. 1:49–84.

Stinebrickner, Ralph and Todd Stinebrickner. 2008. The effect of credit constraints on the college drop-out decision: A direct approach using a new panel study. *American Economic Review* 98, no. 5:2163–84.

———. 2014. Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model. *Journal of Labor Economics* 32, no. 3:601–644.

Stinebrickner, Ralph and Todd R Stinebrickner. 2003. Understanding educational outcomes of students from low-income families evidence from a liberal arts college with a full tuition subsidy program. *Journal of Human Resources* 38, no. 3:591–617.

Stinebrickner, Todd and Ralph Stinebrickner. 2012. Learning about academic ability and the college dropout decision. *Journal of Labor Economics* 30, no. 4:707–748.

Torpey, Elka and Audrey Watson. 2014. Education level and jobs: Opportunities by state. *Career Outlook, US Bureau of Labor Statistics* .

Trachter, Nicholas. 2015. Stepping stone and option value in a model of postsecondary education. *Quantitative Economics* 6, no. 1:223–256.

Zafar, Basit. 2011. How do college students form expectations? *Journal of Labor Economics* 29, no. 2:301–348.

A Stationary Equilibrium

Let $\boldsymbol{s}_j^c \in S_j^c$ be the age-specific state vector for college enrollees and $\boldsymbol{s}_j \in S_j$ for workers and retirees. I also define the age-specific state vector for workers and retirees conditional on education e as $\boldsymbol{s}_j^e \in S_j^e$.

Definition 1 *A stationary equilibrium is a list of value functions of workers and college enrollees* $\{V_j(\boldsymbol{s}_j), V_j^c(\boldsymbol{s}_j^c)\}$*; decision rules of enrollment and graduation* $\{d_j(\boldsymbol{s}_j^c)\}$*; decision rules of consumption, asset holdings, labor hours, output, and parental transfers of workers* $\{c_j(\boldsymbol{s}_j), a'_j(\boldsymbol{s}_j), h_j(\boldsymbol{s}_j), y_j(\boldsymbol{s}_j), b(\boldsymbol{s}_j)\}$*; decision rules of college enrollees* $\{c_j^c(\boldsymbol{s}_j^c), a_j'^c(\boldsymbol{s}_j^c), h_j^c(\boldsymbol{s}_j^c), y_j^c(\boldsymbol{s}_j^c)\}$*; aggregate enrollees, capital, and labor inputs* $\{E, K, H^S, H^U\}$*; prices* $\{r, w^S, w^U, p_e\}$*; policy* τ_ℓ*; and measures* $\boldsymbol{\mu} = \{\mu_j^c(\boldsymbol{s}_j^c), \mu_j(\boldsymbol{s}_j), \mu_j^e(\boldsymbol{s}_j^e)\}$ *such that*

1. *Taking prices and policy as given, value functions* $\{V_j^c(\boldsymbol{s}_j^c), V_j(\boldsymbol{s}_j)\}$ *solve the individual Bellman equations and* $\{c_j(\boldsymbol{s}_j), a'_j(\boldsymbol{s}_j), h_j(\boldsymbol{s}_j), y_j(\boldsymbol{s}_j)\}$, $b(\boldsymbol{s}_j)$, $\{d_j(\boldsymbol{s}_j^c), c_j^c(\boldsymbol{s}_j^c), a_j'^c(\boldsymbol{s}_j^c), h_j^c(\boldsymbol{s}_j^c), y_j^c(\boldsymbol{s}_j^c)\}$ *are associated decision rules.*

2. *Taking prices and policy as given,* K*,* H^{HS}*,* H^{CG} *solve the optimization problem of the firm and* E *solves the optimization problem of the college.*

3. *The government budget is balanced.*

$$G_c + G_e + \sum_{j=j_r}^{J} \int_{S_j} p(e,\theta) d\mu_j = \sum_{j=1,2} \int_{S_j^c} T(c_j^c(\boldsymbol{s}_j^c), a_j^c(\boldsymbol{s}_j^c), y_j^c(\boldsymbol{s}_j^c)) d\mu_j^c$$
$$+ \sum_{j} \int_{S_j} T(c_j(\boldsymbol{s}_j), a_j(\boldsymbol{s}_j^s), y_j(\boldsymbol{s}_j^s)) d\mu_j^s$$

where

$$G_c = gF(K,H)$$

$$G_e = \sum_{j=1,2} \int_{S_j^c} s_j(q) d\mu_j^c.$$

4. *Labor, asset, and education markets clear.*

$$H^S + \kappa E = H^{CG}$$

$$H^U = H^{HS} + H^{CD}$$

where

$$H^{CG} = \sum_{j=3}^{j_r-1} \int_{S_j^{CG}} \epsilon_j^{CG}(\theta,\eta) h_j(\boldsymbol{s}_j) d\mu_j^{CG}$$

$$H^{CD} = \sum_{j=2}^{j_r-1} \int_{S_j^{CD}} \epsilon_j^{CD}(\theta,\eta) h_j(\boldsymbol{s}_j) d\mu_j^{CD} + \int_{S_2^c} \epsilon_2^{CD}(\theta,\eta) h_2^c(\boldsymbol{s}_2^c) d\mu_2^c$$

$$H^{HS} = \sum_{j=1}^{j_r-1} \int_{S_j^{HS}} \epsilon_j^{HS}(\theta,\eta) h_j(\boldsymbol{s}_j) d\mu_j^{HS} + \int_{S_1^c} \epsilon_1^{HS}(\theta,\eta) h_1^c(\boldsymbol{s}_1^c) d\mu_1^c$$

and

$$K = \sum_{j=1}^{J} \int_{S_j} a_j'(\boldsymbol{s}_j) d\mu_j + \sum_{j=1,2} \int_{S_j^c} a_j'^c(\boldsymbol{s}_j^c) d\mu_j^c$$

$$E = \sum_{j=1,2} \int_{S_j^c} d\mu_j^c.$$

5. *Measures $\boldsymbol{\mu}$ are reproduced for each period: $\boldsymbol{\mu}(S) = Q(S, \boldsymbol{\mu})$ where $Q(S, \cdot)$ is a transition function generated by decision rules and exogenous laws of motion, and S is the generic subset of the Borel-sigma algebra defined over the state space.*

B Computation of Stationary Equilibrium

This section describes the method of computing an equilibrium.

1. Starting from an initial vector of aggregate variables $\boldsymbol{w} = \left(\frac{K}{H}, \frac{H^S}{H}, H, \tau_\ell\right)$, compute the prices r, w_S, w_U, p_e and pension $p(e, \theta)$ required for individual decision problems.

2. Given these variables, solve individuals' decision problems. This step consists of sub-steps.

   (a) Solve backward the Bellman equations for age $j = J, \ldots, j_b + 1$. The number of grids for assets is 30 and for high school ability and college ability is 5[29]. The number of grids for college taste is 30. I apply the endogenous grid method.

   (b) Given an initial guess of the value function of newborns V^0, solve backward the individual problems from $j = j_b, \cdots, 1$ for value functions and policy functions. This leads to a new V_0.

[29]The grids of assets depend on age. The range of the grids for high school ability is $[-.55, .55]$ and that for college ability is $[-0.55 - 1.75\sigma_c, 0.55 + 1.75\sigma_c]$. The range of grids for college ability is broader because of the higher variance. That of college taste is $[-2, 2]$.

(c) I implement a Howard-type improvement algorithm: That is, with the decision rules fixed, update V_0 until the guess and the value functions converge.

(d) Given the converged V_0, solve the decision rules of individuals until convergence.

3. I interpolate linearly assets and ability to 80 and 21.

4. Starting from an initial measure μ_0 and given decision rules, solve forward from μ_0 to μ_J and update μ_0 until convergence.

5. Given the measures, derive the new aggregate variables K, H, H^S, and τ_ℓ from the government budget constraint and go back to step 2.

C Pension

The average lifetime income is

$$\hat{y}(e,\theta) = \frac{\sum_{j=2}^{j_r-1} w^e \varepsilon_j^e(\theta,1)\bar{h}}{j_r - 2}$$

where $\bar{h} = 0.333$.

The pension formula is given by

$$p(e,\theta) = \begin{cases} s_1\hat{y}(e,\theta) & \text{for } \hat{y}(e,\theta) \in [0, b_1) \\ s_1 b_1 + s_2(\hat{y}(e,\theta) - b_1) & \text{for } \hat{y}(e,\theta) \in [b_1, b_2) \\ s_1 b_1 + s_2(b_2 - b_1) + s_3(\hat{y}(e,\theta) - b_2) & \text{for } \hat{y}(e,\theta) \in [b_2, b_3) \\ s_1 b_1 + s_2(b_2 - b_1) + s_3(b_3 - b_2) & \text{for } \hat{y}(e,\theta) \in [b_3, \infty) \end{cases}$$

where $s_1 = 0.9$, $s_2 = 0.32$, $s_3 = 0.15$, $b_1 = 0.22\bar{y}$, $b_2 = 1.33\bar{y}$, $b_3 = 1.99\bar{y}$, and $\bar{y} = \$28,793$ annually.

D Labor Productivity Process

To identify the effect of age on wages, I use the Panel Study of Income Dynamics (PSID). I use data for the waves from 1968 to 2014 (from 1997, the PSID has become biannual). I use the SRC sample of heads whose age is between 25 and 63, which leads to 11,512 samples. I restrict observations to those with positive hours of labor for the individual. I keep only people who do not report extreme changes in hourly wages (changes in log earnings larger than 4 or less than -2) or extreme hourly wages (less than \$1 or larger than \$400). I keep only people with 8 or more year observations, which leads to 3,518 samples. Quadratic age polynomials are separately estimated by education group with year dummies. High school graduates are those with 12 years as the highest grade completed. College dropouts are those with the highest grade completed between 13 and 15. College graduates are those with the highest grade completed greater than or equal to 16. Estimation results are

	High school graduates	College dropouts	College graduates
Age	.0530181	.0684129	.0955783
	(.0030501)	(.0040353)	(.0036997)
Age^2	-.0005314	-.0006872	-.0009521
	(.0000356)	(.0000474)	(.0000429)

Table 19: Age profile estimates of each education level

Source: PSID. The methodology is explained in the main text.

in Table 19. I take the average of the productivity of the corresponding two years for the productivity of j in the model and normalize the process so that productivity at the first period after the education stage is unity.

For the law of motion of residuals, I use the same PSID sample and use the residuals of the regression for the age profile. For estimation, I normalize job experience to 0 as age minus 18 for high school graduates, age minus 20 for college dropouts, and age minus 22 for college graduates and apply a minimum distance estimator for different lags and different experience years of the residuals for age 25 to 40. I assume there is a measurement error from an identical and independent distribution. I also assume there is a fixed effect and estimate the persistence ρ^e, the variance of the residual σ^e_η, the variance of the fixed effect, and the variance of the measurement error for each education level.

To identify the effect of ability on wages, I use the NLSY79 after filtering out the age effect using the PSID. Ability is approximated by the log of the AFQT80 raw score. To estimate the coefficient on ability in effective labor, I use the NLSY79 of 11,864 people. For the ability regression, I restrict to samples aged between 25 and 63, which leads to 11,627 people. After the age effect is filtered out, I regress hourly wages on ability for each education level (HS, CD, and CG). As in the selection of the PSID, I keep only people who do not report extreme changes in hourly wages (changes in log earnings larger than 4 or less than -2) or extreme hourly wages (less than \$1 or larger than \$400). I keep only people with 8 or more year observations, which leads to 3,851 people. I exclude enrolled students and hours worked per week less than 20. I also control for dummies for each year.

To handle the selectivity bias problem, I use Heckman two-step estimators. For high school graduates, I assume a linear selection equation of $\ln$AFQT80 and year dummies and the whole sample consists of people whose education level is higher than high school graduates. Among people who graduate high school, those with less ability are self-selected as high school graduates. For college dropouts and college graduates, I assume a linear selection equation of $\ln$AFQT80 and year dummies, and the whole sample contains both college dropouts and graduates. Of those who enroll in college, people with high ability are self-selected as college graduates and people with low ability as college dropouts.

E Intergenerational Ability Transmission

To estimate the transmission of ability from parents to children, I rely on data from the NLSY79 to approximate parents' ability and the "NLSY79 Child & Young Adult" for children. The NLSY79 Child & Young Adult survey started in 1986 and has occurred biennially since then. This survey provides information on test scores of the children of the women in the NLSY79 dataset. Test scores reported include the PIAT math, the PIAT reading recognition, and the PIAT reading comprehension.

There are 11,521 children born to 4,934 female respondents of the NLSY79.

To focus on cognitive ability, I use the PIAT math to approximate the high school ability of children. In particular, I use the standardized PIAT math score, which adjusts for different ages at which the test is taken and is comparable across age. If there are multiple PIAT math scores for a child, I use only the latest score. I exclude children whose PIAT math scores are missing. This leaves me with 9,232 children born to 4,055 mothers.

I use AFQT scores to measure mothers' ability. In particular, I only use respondents whose AFQT scores and education levels are both present. I focus on people with high school degrees. This leaves me with 6,193 children born to 2,828 mothers.

F Calibration without Optimism

In this paper, I assume that students are overly optimistic about their college abilities before enrollment, which is a key factor in the large college dropout rate in the United States. In this section, I examine a different approach to explain the large college dropout rate in the United States: a large option value of college enrollment with no optimism. I assume $\mu_c(\theta_h) = 0$ for all θ_h, and assume that the standard deviation of college ability depends on high school ability and

$$\sigma_c(\theta_h) = \sigma_c \exp(\sigma_c^\theta \theta_h).$$

Table 20 displays the calibrated values under the specification without optimism. As we can see, the intercept of the standard deviation of college ability σ_c is larger than the case with optimism. I need a high standard deviation and a high

Parameter	Description	Value
λ	psychic cost intercept	-16.6
λ^{θ}	psychic cost slope	287
λ_1^{ϕ}	first period college taste	-68.8
λ_2^{ϕ}	second half college taste	-40.0
a^S	productivity of skilled labor	0.435
ϵ^{CD}	productivity of CD	0.985
σ_c	s.d. of college ability intercept	0.721
σ_c^{θ}	s.d. of college ability slope	0.158
κ	education cost	0.422
μ	consumption share of preference	0.422
β	time discount rate	0.931
v	altruism	0.0630
d	lump-sum transfer ratio	0.131
ι	borrowing wedge ($r^- = r + \iota$)	18.7%
m	intergenerational ability transmission intercept	-0.0384
σ_h	intergenerational ability transmission s.d.	0.0764

Table 20: Parameters calibrated without optimism

option value to match the high college dropout rate. The slope of the standard deviation is positive and the uncertainty about college ability is higher for higher high school ability.

Table 21 and Figures 3 and 4 display moments. The graduation rate of low high school ability is excessively high in this formulation without optimism. In order to match the high dropout rate of low-ability people, the model requires a high standard deviation of college ability. Then too many people draw high college ability enough to stay and graduate. To match the low graduation rate of high-ability people, the model requires high psychic cost for low-ability people. Then enrollment decreases and college dropout also decreases. While the effect of increasing the standard deviation of college ability increases college enrollment, it also increases

Moment	Model	Data
Enrollment rate of ability quartile	(figure)	(figure)
Graduation rate of ability quartile	(figure)	(figure)
Enrollment rate of family income quartile	(figure)	(figure)
Graduation rate of family income quartile	(figure)	(figure)
Skill premium for CG	90.7%	90.2%
Skill premium for CD	20.1%	19.9%
Education cost/mean income at 48	0.308	0.33
Hours of work	33.3%	33.3%
K/Y	1.241	1.325
Transfer/mean income at 48	67.2%	66%
Log pre-tax/post-tax income	60.5%	61%
Borrowers	6.07%	6.3%
Mean of AFQT	0.0880	0
Standard deviation of AFQT	0.204	0.213

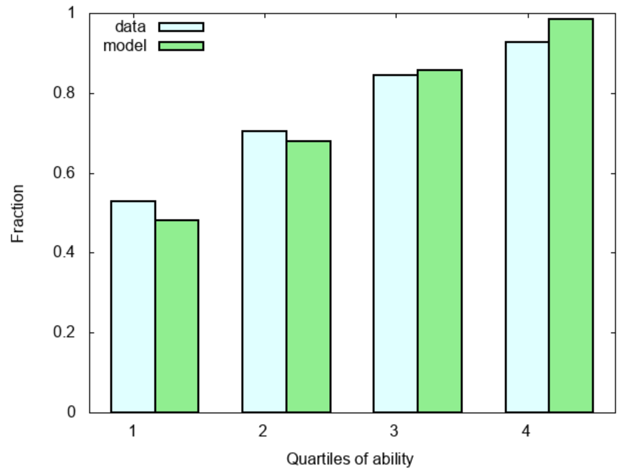
Table 21: Moments matched without optimism

psychic cost to match the high college dropout and decreases college enrollment at the same time, offsetting the first effect. Instead, the best match requires a *low* standard deviation of college ability and low college psychic cost to match the high enrollment. To summarize, a high option value without optimism does not explain the high dropout rate in the data compared with the model with optimism.

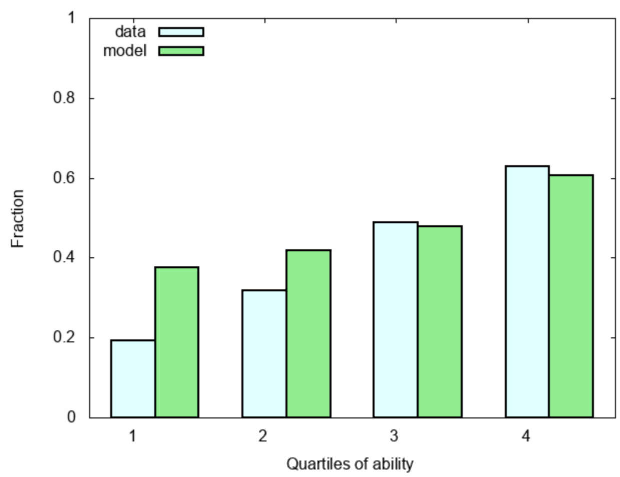
The optimal policy with this formulation is shown in Table 22. The optimal

	Current state	Optimal
$s_1(1)$	\$13,600	\$14,153
$s_1(2)$	\$11,448	\$11,913
$s_1(3)$	\$10,923	\$11,367
$s_2(1)$	\$13,600	\$12,478
$s_2(2)$	\$11,448	\$10,503
$s_2(3)$	\$10,923	\$10,021

Table 22: Optimal subsidies without optimism



(a) Enrollment rate

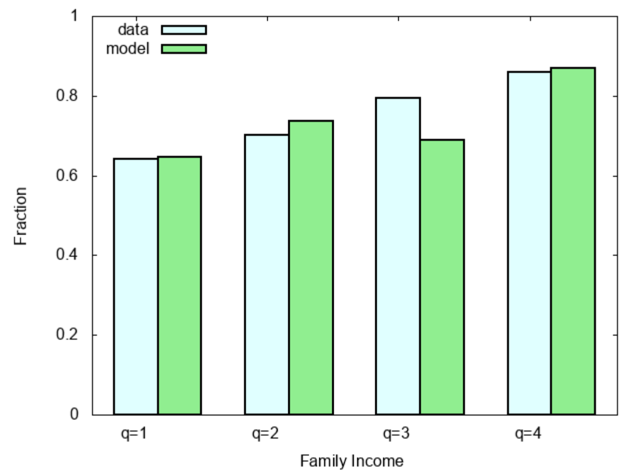


(b) Graduation rate

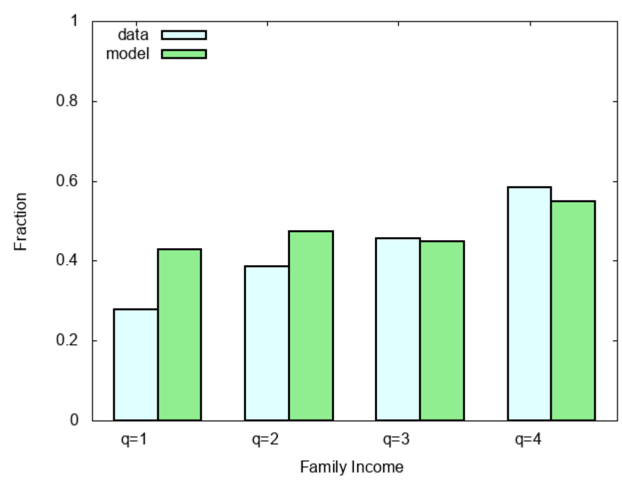
Figure 3: Model fit: Enrollment and graduation rate for each ability quartile

Data Source: NLSY97. I use the sample of only 25-year-old high school graduates. Ability is the log of AFQT score using the definition from the NLSY79. Scores are adjusted by age, as in Altonji et al. (2012) and Castex and Kogan Dechter (2014).

policy is now *front-loaded*. Without optimism, enrollment is excessively low in the current state, as in the case with bias correction. This implies that the assumption regarding optimism matters, and one of the contributions of this paper is to calibrate



(a) Enrollment rate



(b) Graduation rate

Figure 4: Model fit: Enrollment and graduation rates for each family income quartile

Data Source: NLSY97. I use the sample of only 25-year-old high school graduates. Family income is defined as the average of parental income at 16 and 17 if both are available. I use one if both are not available.

this effect using available data.