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

Kazushige Matsuda†

March 7, 2019

[Link to the latest version]

Abstract

In the United States, a half of college enrollees drop out before earning a bachelor's degree. This paper examines the effect of a new college subsidy scheme whose amount varies across years on enrollment, graduation, and the skill premium compared to the current system in which the subsidy is constant across years. I find that switching to back-loaded subsidies with the same total budget increases the number of college graduates and decreases the skill premium more than the case with increasing the total budget of the current subsidies by 50%, and are welfare improving despite the fact that enrollment decreases.

*I am extremely grateful to my advisor Richard Rogerson for crucial 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, Gianluca Violante for helpful comments. I also receive valuable comments feedback from the audiences at various seminars.

†km11@princeton.edu.

1 Introduction

Wage inequality has been increasing in the United States. The skill premium—the wage premium of college graduates to high school graduates—has increased from 50% in 1980 to 90% now. A large literature (ex. Goldin and Katz (2007) and Katz and Murphy (1992)) argues that the skill premium rises because the increase in the supply of college graduates does not catch up with the increase in the demand for skilled labor. In this framework, we can reduce the skill premium by increasing college graduates in the economy. With this in mind, the existing literature often suggests policies seeking to increase college enrollment: it often equates enrollment with graduation. However, in the United States, while over 70% of high school graduates enroll in college, more than half of them drop out before earning a bachelor's degree (See Table 1)[1]. Enrollment does not necessarily lead to graduation and we should treat enrollment and graduation as two different margins. It is important for us to understand how policy can affect graduation separately from enrollment.

In this paper, I propose a new college subsidy scheme in which the amount of subsidies vary with years of college ("year-dependent subsidies"), i.e., subsidies that differ for freshmen, sophomores, and so on. The existing literature has only considered subsidies that are constant across years in college. Subsidies that vary by year will have differential impacts on enrollment and graduation unlike constant subsidies, as the following example suggests. After they graduate from high school, individuals decide to enroll in college based on their high school GPA or high school ability. In general, people with high ability want to enroll. But one's high school GPA is not necessarily the same as his/her college GPA. After enrolling, some students learn that their college GPA

[1]Two-year college graduates who do not transfer are counted as dropout. According to a report from the National Center for Education Statistics for 1994-2009, more than 80 percent of community college freshmen say that their ultimate goal is a bachelors or higher degree (Horn and Skomsvold (2012)). The sheepskin effect of associate degrees is not high (See Kane and Rouse (1995)) and only 5% of enrollees at two-year colleges graduate and do not transfer (See Trachter (2015)).

or college ability is low and drop out. Consider back-loaded subsidies in this setting: increasing subsidies for the latter years of college and decreasing subsidies for the early years. People who expect to drop out before earning the increased subsidies for latter years stop enrolling due to the decreased subsidies for early periods. In contrast, the marginal college dropout now finds it worthwhile to continue as the subsidies for the latter periods increase. Therefore, the number of college graduates increases while enrollment decreases, and vice versa for front-loaded subsidies. Year-dependent subsidies can affect enrollment and graduation in different ways unlike constant subsidies. The question of this paper is how year-dependent subsidies affect enrollment and graduation and what timing of college subsidies will maximize the number of college graduates and welfare.

I build a life-cycle general equilibrium model with credit constraints, endogenous enrollment and dropout decisions. Agents are heterogeneous with regard to initial asset, high school ability, and college ability. College ability affects utility in college and earnings after graduation. Agents are over-optimistic with regard to college ability before enrollment in order to be consistent with the existing empirical findings. Agents learn their college ability after enrollment and decide to drop out or not. These educational decisions shape the aggregate skill in the economy and the skill premium through imperfect substitution between skilled and unskilled labor. I calibrate the model to match the enrollment, graduation, and the skill premium in 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 lifetime utility of newborns with equal weights. The focus of this paper is on relative sizes across years and I fix the total budget of college subsidies at the current level when varying subsidies across college years.

The main finding of this paper is the following. First, it is back-loaded subsidies

HGPA Quantile	% graduation
Q1	19%
Q2	31%
Q3	48%
Q4	63%
total	42%

Table 1: College graduation Rates for High school GPA Quartiles

Source: NLSY97. I use the sample of only 25 year old people. Family income is defined as the average of parental income at 16 and 17 if both are available. I use the one if only one of the two is available.

that maximize the number of college graduates and welfare. Second, by switching to the back-loaded subsidies with the same total budget, the number of college graduates increases and the skill premium decreases more than the case with increasing the total budget of constant subsidies by 50%. With the back-loaded subsidies, college enrollees have incentive to stay in college and the share of college graduates increases, which reduces the skill premium. On the other hand, enrollment decreases due to decreased subsidies for early years. Third, back-loaded subsidies improve social welfare by 0.12% of lifetime consumption at the steady state, without increasing tax. The gains come from two effects. First, there is excess enrollment because agents are over-optimistic. Back-loaded subsidies reduce enrollment and prevent low-ability people from enrolling and deriving disutility. Second, 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 to risk-averse agents.

The rest of this paper is organized as follows. Section 2 outlines the model and defines an equilibrium. Section 3 makes the model quantitative by calibration and estimation. Section 4 presents results and, in section 5, I provide discussion and concluding remarks.

1.1 Related Literature

In the macroeconomic literature, Bovenberg and Jacobs (2005) theoretically derive the effect of subsidies. Abbott, Gallipoli, Meghir and Violante (2018) emphasize the effect of subsidies on parental transfers in a quantitative overlapping generations model. Krueger and Ludwig (2016) analyze the optimal income tax and subsidies simultaneously and show that the less progressive labor income tax and a large amount of subsidies than the current state are optimal for a social utilitarian welfare function.

One of the early papers of a model with college dropout is Manski (1989), who show that dropout has an option value and college enrollees can experiment on the real value of college going. Arcidiacono, Aucejo, Maurel and Ransom (2015), Athreya and Eberly (2016), Lee, Shin and Lee (2015), and Castex (2017) analyze how introducing college dropout into models change the allocation of human capital and returns to education. Stange (2012) and Trachter (2015) quantitatively show the importance of the option value of college dropout quantitatively. Hendricks and Leukhina (2017) argue that college dropout is predictable before enrollment due to the strong correlation between high school GPA and the dropout rate.

Caucutt and Kumar (2003) and Akyol and Athreya (2005) show the normative implication of subsidies with exogenous college dropout risk. Hanushek, Leung and Yilmaz (2014) analyze the effect of various college aid with exogenous college dropout risk. Although the majority of the literature on college subsidies regard dropout as exogenous, there are some exceptions. Ionescu (2011) shows the effect of default policies of student loan on educational decisions. Garriga and Keightley (2007) show the effect of an increase in subsidies on the dropout decision and labor supply in a general equilibrium framework. Chatterjee and Ionescu (2012) argue it is welfare improving to insure student loan against exogenous financial risk of dropping out with endogenous dropout decisions. Colas, Findeisen and Sachs (2018) show that optimal college subsidies are more need-based than the current system. A difference between these literature and this paper is that they have not considered subsidies that can vary across college years.

The sequential papers of Stinebrickner and Stinebrickner (2008), Stinebrickner and Stinebrickner (2012), Stinebrickner and Stinebrickner (2014) show that learning academic ability during college is a main driver of college dropout. The model of this paper is based on their empirical findings on the reasons of college dropout.

2 Model

The model has four main building blocks. The first is year-dependent subsidies: subsidies that vary with years in college, which is a brand-new ingredient in the literature. While subsidies are constant across years in the calibration, I examine the effect of year-dependent subsidies in the following model.

The second is a model of endogenous enrollment and graduation decisions based on Garriga and Keightley (2007) and Ionescu (2011). At the first period after high school graduation, individuals make an enrollment decision based on their initial asset and high school ability. College enrollees learn their college abilities and decide to drop out of college or not. College ability is a key factor in that it determines utility in college and the returns to college graduation.

The third building block is an overlapping generations life cycle with incomplete markets with inter-generational linkage of ability and wealth, based on Abbott et al. (2018) and Daruich (2017). Individuals in the model face uncertainty with regard to college ability and labor productivity over life cycle with no insurance available. Individuals give birth to children with inter-generationally correlated ability and make an endogenous wealth transfer to their children. College subsidies can crowd-out endogenous wealth transfers from parents.

The fourth building block is a general equilibrium framework with an aggregate production function featuring imperfect substitution between skilled and unskilled labor based on Goldin and Katz (2007) and Katz and Murphy (1992). The educational decisions aggregate to the supply of skill, which determines the skill premium.

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 any time subscript in the description of the economy.

2.1 Demography

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

Figure 1 is the timeline. At the beginning of age 1, individuals make enrollment decisions. Once they do not enroll in college, they cannot enroll later. Time is discrete and one period in the model corresponds to two years. Consistent with college typically requiring four years in reality, college graduation requires two periods in the model. At the beginning of age 2, a college enrollee makes a decision about whether to continue in college or not. Once an individual finishes their schooling, they will be one of three types: high school graduates ($e = HS$) for those who do not enroll at age 1, college dropouts ($e = CD$) for those who do not continue college at the beginning of age 2, and college graduates ($e = CG$) for those who finish two 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 leave and become economically independent and parents give wealth transfers to their children.

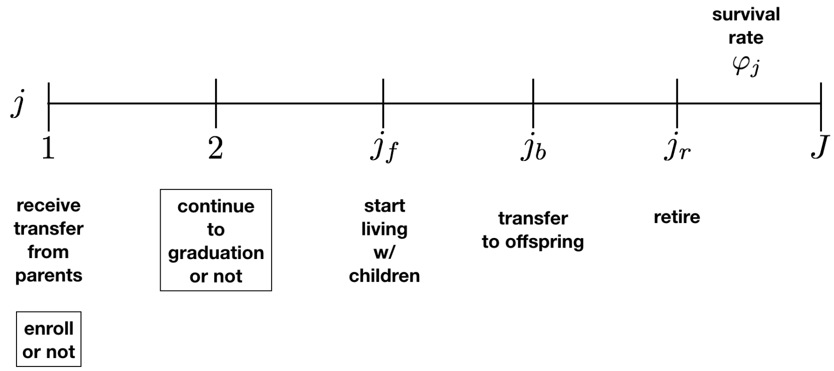


Figure 1: Timeline.

There are no transfers allowed at other ages[2]. Individuals retire at age $j_r = 25$ (at biological age 66) and the maximum age is $J = 42$ (at biological age 100). It is at age j_b when the child leaves the household with a wealth transfer from parents.

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.

2.2 Preferences

When an individual becomes economically independent at age 1, he or she has preferences represented by the sum of three components:

1. The expected discounted sum of instant utility:

$$\mathbb{E}_1 \sum_{j=1}^{J} \tilde{\beta}^{j-1} u(c_j, \ell_j) \tag{1}$$

[2] If transfers are allowed at other ages such as age 2, the state variables of parents have to include their children's state variables and solving the individuals' problem becomes formidable. Transfers from parents changes the result mainly when credit constraints bind for their children. As you will see later, the credit limit for age 1 is tighter than the limit for age 2 and it is unlikely that the transfers from parents at age 1 changes the outcome.

where

$$u\left(c,\ell\right)=\frac{\left(c^{\mu}\ell^{1-\mu}\right)^{1-\gamma}}{1-\gamma} \tag{2}$$

and c_j denotes 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. The individuals are endowed with one 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.[3]

2. The expected college utility:

$$\mathbb{E}_1 d_0(\boldsymbol{s}_0)\lambda_1(\theta_c,\phi)+\beta\mathbb{E}_1 d_1(\boldsymbol{s}_1^c)\lambda_2(\theta_c,\phi) \tag{3}$$

where

$$\lambda_j(\theta_c,\phi)=\lambda+\lambda^{\theta}\theta_c+\lambda_j^{\phi}\phi \tag{4}$$

and $d_0(\boldsymbol{s}_0)$ is an indicator function which is one if the individual enrolls and $d_1(\boldsymbol{s}_1^c)$ is an indicator function for graduation. Individuals derive this utility only while in college. As in Heckman, Lochner and Todd (2006), the psychic cost of education is an important factor determining college education. I define $\lambda_j(\theta_c,\phi)$ not as disutility but as utility without loss of generality. College utility depends on two components: college ability θ_c and college taste ϕ. ϕ is fixed over lifetime while the coefficient λ_j^{ϕ} can vary across periods (different loading).

I need the two different factors θ_c and ϕ for college utility to match the data. In the data, within the same category of high school ability and family income, there is heterogeneity in terms of enrollment decisions (it is true neither that everyone enrolls nor that no one enrolls). To explain it, I need college taste

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

ϕ unobservable to econometricians. Individuals observe college taste ϕ before enrollment. I explain ability and college taste in more detail in the individual problems section.

3. Parental altruism.

$$\beta^{j_b-1}\nu\mathbb{E}_1 V_0 \tag{5}$$

where V_0 is the expected lifetime value of their children at the beginning of age 1. I will explain the detail of the value function later. Individuals enjoy their children's lifetime utility with a weight ν. This is a motive of 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:

$$Y = F(K, H) = K^{\alpha}H^{1-\alpha} \tag{6}$$

where aggregate labor services H is a function of two skill levels: skilled labor S and unskilled labor U.

$$H = (a^S(H^S)^{\rho} + (1-a^S)(H^U)^{\rho})^{\frac{1}{\rho}} \tag{7}$$

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

$$r = \alpha \left(\frac{K}{H}\right)^{\alpha-1} - \delta \tag{8}$$

$$w^S = (1-\alpha)a^S \left(\frac{K}{H}\right)^{\alpha} \left(\frac{H}{H^S}\right)^{1-\rho} \tag{9}$$

$$w^U = (1-\alpha)(1-a^S) \left(\frac{K}{H}\right)^{\alpha} \left(\frac{H}{H^U}\right)^{1-\rho} \tag{10}$$

There are two types of skill in production while there exist three levels of education. In the literature on the skill premium as in Katz and Murphy (1992), high school graduates are assumed to provide unskilled labor and college graduates provide skilled labor. I assume college dropouts provide unskilled labor. Torpey and Watson (2013)[4] present the proportion of jobs in the United States by required education level. They show that many jobs require either "Bachelor's degree", "High school diploma or equivalent", or "Less than high school". Since the model considers only people who have graduated from high school, the important distinction is "Bachelor's degree" versus "High school diploma or equivalent". I interpret the jobs requiring the former as skilled labor and the latter as unskilled labor. In addition, they show that only 5% of jobs require "Some college, no degree" and "Associate's degree", which implies that most college dropouts take jobs requiring education level "High school diploma or equivalent". Thus I assume college dropouts provide unskilled labor. 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 $\varepsilon_j^e(\theta, \eta)$, which depends on education e, age

[4]They use the May 2013 data of Occupational Employment Statistics survey (employment data) and Employment Projections program (occupational education-level designations) by the U.S. Bureau of Labor Statistics. I assume the jobs for "Bachelor's degree", "Master's degree", and "Doctoral or professional degree" in their categories require college graduation. I assume the jobs for "Some college, no degree", "Associate's degree", and "Postsecondary nondegree award" require some college.

j, ability θ, and idiosyncratic productivity η. The stochastic productivity shock η is mean-reverting and follows an education-specific Markov chain $\pi_\eta^e(\eta'|\eta) > 0$ and Π_η^e denotes its invariant distribution function. Labor productivity of high school graduates and college dropouts depends on high school ability. Labor productivity of college graduates depends on college ability.

2.4 College

There is a representative college. To provide a college enrollee with one period of education requires κ units of skilled labor, which means college enrollees receive education from professors who themselves are college graduates. I assume education does not require any capital or unskilled labor.[5]

The profit of college is

$$p_e E - w^S \kappa E \tag{11}$$

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 governments, which enables the sticker tuition smaller than the actual education cost. I reinterpret this situation as follows: colleges do not receive any subsidy while college enrollees receive subsidies instead. At the same time, they have to pay the full education cost. In both cases, enrollees pay p_e less the subsidy for education.

[5] While this is a strong assumption, this formulation captures an important aspect of tuition. When considering policy changes, it is important to keep track of what happens to tuition. Archibald and Feldman (2011) argue that college tuition reflects wages of college graduates. Policies that affect the wages of skilled labor can also affect tuition which potentially has an effect on enrollment and graduation. While this specification is too simple, it captures the effect of the skill premium on tuition.

2.5 Financial Markets

The financial market is incomplete. There is no insurance market against idiosyncratic risks and individuals can self-insure using risk-free assets.

Lenders incur the cost of overseeing borrowers to lend capital to workers and the cost per unit of capital is $\iota > 0$. With the non-arbitrage condition, the interest rate to 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 college enrollees is $\iota + \iota^s$. With the non-arbitrage condition, the interest rate to enrollees is $r^s = r + \iota + \iota^s = r^- + \iota^s$. The borrowing limit for college enrollees is $\underline{A}_j^c$ 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 the individuals who are not in college "workers". Likewise, I call the periods when the individuals are not in college "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 or not. I define V_0 to be the value function.

$$V_0(a, \theta_h, \eta, q, \phi) = \max[\underbrace{V_1^c(a, \theta_h, \eta, q, \phi)}_{\text{enrolling}}, \underbrace{V_1(a, HS, \theta_h, \eta)}_{\text{not enrolling}}] \tag{12}$$

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.

There are two types of ability that are distinct but related to each other: high school ability θ_h and college ability θ_c. Individuals observe high school ability but do not observe college ability before the enrollment decision. Individuals observe their high school ability through high school grade point average (GPA) or test scores during high school. College ability is only observed after the first period of college. Stinebrickner and Stinebrickner (2012) present evidence that enrollees do not have perfect foresight of their college abilities before enrollment. However, college abilities are correlated with high school abilities and

$$\theta_c = \theta_h + \epsilon_c \text{ where } \epsilon_c \sim N(0, \sigma_c^2) \tag{13}$$

I assume that college enrollees are over-optimistic about their college abilities, in order to be consistent with an empirical finding of Stinebrickner and Stinebrickner (2012) that optimism is a key factor of enrollment. They have a longitudinal survey of students, which asks each student his or her expectation of GPA multiple times. First, they show that their expectations of college GPA before the first semester is higher than their actual GPAs on average, which suggests over-optimism. Second, they show that college 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 are the most optimistic and had the largest downward revisions of their expectations. Given θ_h, enrollees expect that

$$\theta_c = \underbrace{\mu_c(\theta_h)}_{\text{bias}} + \underbrace{\theta_h + \epsilon_c}_{\text{actual ability}} \quad \text{where } \epsilon_c \sim N(0, \sigma_c^2) \tag{14}$$

where $\mu_c(\theta_h)$ is the bias. If $\mu_c(\theta_h)$ is positive, enrollees are over-optimistic about their college abilities. Furthermore, the bias can depend on high school ability and I assume $\mu_c(\theta_h) = \mu_{c0} + \mu_{c1}\theta_h$. I assume that the variance of the residual term is identical to the actual one.

Initial wealth a is endogenously determined as a transfer from their parents as will be shown later. 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 seen later.

If an individual enrolls, he or she enters the first half of college where the value is V_1^c. If they do not enroll, they start working as high school graduates and its value is V_1.

First half of college

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

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

subject to

$$c + a' + p_e - s_1(q) = a + y - T(c,a,y) \tag{15}$$

$$y = w^{HS}\varepsilon_1^{HS}(\theta_h,\eta)h, \;\; a' \geq -\underline{A}_1^c \;\; c \geq 0, \;\; 0 \leq h \leq 1-\bar{h} \tag{16}$$

$$\theta_c = \theta_h + \mu_c(\theta_h) + \epsilon_c, \;\; \epsilon_c \sim N(0,\sigma_c^2) \text{ (perceived process)} \tag{17}$$

Going to college requires a fraction $\bar{h}$ of time, tuition p_e and additive utility $\lambda_j(\theta,\phi)$ for each enrolling period. c is consumption, 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 first half of college, college enrollees observe their college ability θ_c and a new idiosyncratic productivity η' drawn from Π^{CD}. College enrollees choose whether or not to drop out of college after this. If the individual drops out, his or her education level becomes college dropout ($e = CD$) and their value is V_2. After dropping out, all 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 that fixed payments would have been made for 20 years (10 periods) after dropout[6]. If the individual does not drop out, they proceed to the second half of college with value V_2^c.

Second half of college

The Bellman equation for the second half of college is

$$V_2^c(a, \theta_c, \eta, q, \phi) = \max_{c,h,a',y} u(c, 1 - h - \bar{h}) + \lambda_2(\theta_c, \phi) + \beta \mathbb{E}_{\eta'} V_3(\tilde{a}(a'), CG, \theta_c, \eta) \tag{18}$$

subject to

$$c + a' + p_e - s_2(q) - y + T(c, a, y) = \begin{cases} (1 + r)a & \text{if } a \geq 0 \\ (1 + r^s)a & \text{if } a < 0 \end{cases} \tag{19}$$

$$y = w^{CD} \varepsilon_2^{CD}(\theta_c, \eta)h, \;\; a' \geq -\underline{A}_2^c \;\; c \geq 0, \;\; 0 \leq h \leq 1 - \bar{h} \tag{20}$$

They can work as college dropout. At the end of period, they complete college and acquire education level $e = CG$ and draw a new idiosyncratic productivity η' from Π^{CG}. Student loan is refinanced into a single bond and the transformation is $\tilde{a}(a')$. The value

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

of workers at age j is V_j.

2.6.2 Working Stage

The Bellman equation for workers is[7]

$$V_j(a, e, \theta, \eta) = \max_{c,h,a',y} u\left(\frac{c}{1+\mathbb{1}_{\mathcal{J}_f}\zeta}, 1-h\right) + \beta \mathbb{E}_{\eta'|\eta} V_{j+1}(a', e, \theta, \eta') \tag{21}$$

subject to

$$c + a' - y + T(c, a, y) = \begin{cases} (1+r)a & \text{if } a \geq 0 \\ (1+r^-)a & \text{if } a < 0 \end{cases} \tag{22}$$

$$y = w^e \varepsilon^e_j(\theta, \eta) h, \;\; a' \geq -\underline{A}^e \;\; c \geq 0, \;\; 0 \leq h \leq 1 \tag{23}$$

where $\mathbb{1}_{\mathcal{J}_f}$ is an indicator function which is one when the individuals live with their children ($j \in [j_f, j_b - 1]$). Ability is $\theta = \theta_h$ for high school graduates and college dropouts. $\theta = \theta_c$ for college graduates. At each period, idiosyncratic productivity η transitions according to π^e_η.

2.6.3 Transfer

At the age j_b, the individuals' children become independent and they determine the amount of transfer. Its Bellman equation is

$$V_{j_b}(a, e, \theta, \eta) = \max_{b \in [0,a]} \mathbb{E}_{\theta'_h|e,\theta} \{\tilde{V}_{j_b}(a - b, e, \theta, \theta'_h, \eta) + \nu \mathbb{E}_{\eta'',\phi} V_0(b, \theta'_h, \eta'', \tilde{q}(w^e \varepsilon^e_j(\theta, \eta)), \phi)\} \tag{24}$$

where

$$\tilde{V}_{j_b}(a, e, \theta, \theta'_h, \eta) = \max_{c,h,a',y} u(c, 1-h) + \beta \mathbb{E}_{\eta'|\eta} V_{j_b+1}(a', e, \theta, \eta') \tag{25}$$

[7] After retirement, labor productivity is 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

$$c + a' - y + T(c, a, y) = \begin{cases} (1+r)a & \text{if } a \geq 0 \\ (1+r^-)a & \text{if } a < 0 \end{cases} \tag{26}$$

$$y = w^e \varepsilon_j^e(\theta, \eta)h, \;\; a' \geq -\underline{A}^e \;\; c \geq 0, \;\; 0 \leq h \leq 1 \tag{27}$$

At the beginning 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. The density function for the child's ability is $\pi_\theta(\theta'_h|\theta)$. Parents can observe neither their children's initial idiosyncratic productivity η'' drawn from Π^{HS} nor college taste ϕ drawn from the normal distribution $N(0,1)$. Consumption, leisure, asset holdings, and parental transfers can depend on θ'_h. The value of their children depends on family income level q which is a function of the potential labor income of the parents.[8]

2.6.4 Retirement Stage

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

$$V_j(a, e, \theta) = \max_{c, a'} u(c, 1) + \beta \varphi_{j+1} V_{j+1}(a', e, \theta) \tag{29}$$

subject to

$$c + a' = (1+r)\varphi_j^{-1} a + p(e, \theta) - T(c, \varphi_j^{-1} a, 0) \tag{30}$$

$$a' \geq 0 \;\; c \geq 0 \tag{31}$$

[8]Note that the parental income is not the actual labor income. The 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 spend 35% working. Thus the family income mapping is

$$\tilde{q}(w^e \varepsilon_j^e(\theta, \eta)) = \begin{cases} 1 & \text{if } w^e \varepsilon_j^e(\theta, \eta) \times 0.35 \in [0, q_1] \\ 2 & \text{if } w^e \varepsilon_j^e(\theta, \eta) \times 0.35 \in [q_1, q_2] \\ 3 & \text{else} \end{cases} \tag{28}$$

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

The sources of income are interests and retirement benefits $p(e,\theta)$. In the United States, retirement benefits are determined by the labor earnings before retirement (see Appendix B). To capture this, the retirement benefits depend on their ability and education. The asset inflated by φ_j^{-1} reflects that 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 subsidies G_e, other government consumption G_c and retirement benefits. Government consumption G_c is exogenous and proportional to the aggregate output $G_c = gY$. The total budget of college subsidies is

$$G_e = \sum_{j=1,2} \int_{S_j^c} s_j(q) d\mu_j^c \tag{32}$$

The tax function is

$$T(c,a,y) = \tau_c c + \tau_k r a \mathbb{1}_{a\geq 0} + \tau_l y - d\frac{Y}{N} \tag{33}$$

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 the individuals. This reflects the progressive income tax. τ_l is the proportional part of labor income tax.

2.8 Equilibrium

The model includes J overlapping generations and is solved numerically to characterize a stationary equilibrium. Stationarity implies that the cross-sectional allocation within each cohort j is invariant. In equilibrium, individuals maximize expected lifetime utility, firms maximize profits, the government budget is balanced each period, and prices clear all the markets. 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 and $\boldsymbol{s}_0 \in S_0$ for individuals at the beginning of age 1. I also define the age-specific state vector for workers and retirees conditional on education e as $\boldsymbol{s}_j^e \in S_j^e$. Computation is described in Appendix A.

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* $d_0(\boldsymbol{s}_0)$ *and graduation* $d_1(\boldsymbol{s}_1^c)$*, decision rules of consumption, asset holdings, labor, output, 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*, 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 policies as given, value functions* $\{V_j^c(\boldsymbol{s}_j^c), V_j(\boldsymbol{s}_j)\}$ *solve the individual Bellman equations and* $d_0(\boldsymbol{s}_0)$, $d_1(\boldsymbol{s}_1^c)$, $\{c_j(\boldsymbol{s}_j), a_j'(\boldsymbol{s}_j), h_j(\boldsymbol{s}_j), y_j(\boldsymbol{s}_j), b(\boldsymbol{s}_j)\}$, $\{c_j^c(\boldsymbol{s}_j^c), a_j'^c(\boldsymbol{s}_j^c), h_j^c(\boldsymbol{s}_j^c), y_j^c(\boldsymbol{s}_j^c)\}$ *are associated decision rules.*

2. *Taking prices and policies as given,* K, H^{HS}, H^{CG} *solve the optimization problem of the good sector and* E *solves the optimization problem of the education sector.*

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

$$G_c = gF(K,H) \tag{34}$$

$$G_e = \sum_{j=1,2} \int_{S_j^c} s_j(q) d\mu_j^c \tag{35}$$

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

$$H^S + \kappa E = H^{CG} \tag{36}$$

$$H^U = H^{HS} + H^{CD} \tag{37}$$

where

$$H^{CG} = \sum_{j=3}^{j_r - 1} \int_{S_j^{CG}} \epsilon_j^{CG}(\theta, \eta) h_j(\boldsymbol{s}_j) d\mu_j^{CG} \tag{38}$$

$$H^{CD} = \sum_{j=2}^{j_r - 1} \int_{S_j^{CD}} \epsilon_j^{CD}(\theta, \eta) h_j(\boldsymbol{s}_j) d\mu_j^{CD} + \int_{S_2^c} \epsilon_2^{CD}(\theta, \eta) h_2^c(\boldsymbol{s}_2^c) d\mu_2^c \tag{39}$$

$$H^{HS} = \sum_{j=1}^{j_r - 1} \int_{S_j^{HS}} \epsilon_j^{HS}(\theta, \eta) h_j(\boldsymbol{s}_j) d\mu_j^{HS} + \int_{S_1^c} \epsilon_1^{HS}(\theta, \eta) h_1^c(\boldsymbol{s}_1^c) d\mu_1^c \tag{40}$$

and

$$K = \sum_{j=1}^{j_r - 1} \int_{S_j} a_j'(\boldsymbol{s}_j) d\mu_j + \sum_{j=1,2} \int_{S_j^c} a_j'^c(\boldsymbol{s}_j^c) d\mu_j^c \tag{41}$$

$$E = \sum_{j=1,2} \int_{S_j^c} d\mu_j^c \tag{42}$$

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.*

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.

3.1 Labor Productivity Process

I assume labor productivity

$$\ln \epsilon_j^e(\theta, \eta) = \ln \epsilon^e + \ln \psi_j^e + \epsilon_\theta^e \theta + \ln \eta \tag{43}$$

where ψ_j^e is the age profile of workers at age j at education level e estimated from PSID (See Appendix C). The coefficients can vary across education levels.

The ability used in the wage process differs across education levels. For high school graduates and college dropouts, θ is high school ability which is approximated by ln AFQT80. 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 a college GPA, quality of college, college majors, and other factors. Since it is hard to measure, I instrument college ability using high school ability. From the law of motion connecting high school and college ability $\theta_c = \theta_h + \epsilon^c$, you can express the log labor productivity as

$$\ln \epsilon^e + \ln \psi_j^e + \epsilon_\theta^e \theta_c + \ln \eta = \ln \epsilon^e + \ln \psi_j^e + \epsilon_\theta^e \theta_h + (\ln \eta + \epsilon_\theta^e \epsilon_c) \tag{44}$$

Since θ_h is uncorrelated with $\ln \eta + \epsilon_\theta^e \epsilon_c$, I can estimate the coefficient ϵ_θ^{CG} using ln AFQT80 in the same way[9]. Table 2 shows the estimated coefficients on ability for each education level. As in the literature, returns to education are higher for high ability.

I assume $\pi_\eta^e(\eta'|\eta)$ is a Markov chain with two states η_H and η_L specific to each

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

	HS	CD	CG
log AFQT	.6021263	.7358609	1.306158
	(.3102392)	(.3144641)	(.2388868)

Table 2: Estimated ability slope ϵ_θ^e

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

Table 3: Estimated parameters of wage process

education level. It has exactly the same persistence and conditional variance as the AR(1) process:

$$\ln \eta' = \rho^e \ln \eta + \epsilon_\eta^e, \quad \epsilon_\eta^e \sim N(0, \sigma_\eta^{e2}) \tag{45}$$

After filtering out age effects, I employ a Minimum Distance Estimator with a fixed effect and a measurement error. I use as moments the covariances of the wage residuals at different lags and age groups, separately for each education level. In Appendix C, I discuss sample selections and the detail of the estimation procedures. Table 3 is the estimated parameters.

3.2 Intergenerational Ability Transmission

Newborns draw their high school abilities θ_h' from a normal distribution whose mean depends on the ability of their parents.

$$\theta_h' = m + m_\theta \theta + \epsilon_\theta, \;\; \epsilon_\theta \sim N(0, \sigma_h^2) \tag{46}$$

High school ability is formed partly as a result of genetics, which leads to a correlation between parents' and childrens' ability. In addition, as Cunha and Heckman (2007), Cunha (2013), and Daruich (2017) suggest, high ability parents earn higher income, which increases early educational investment and their children's high school ability.

In order to estimate the conditional mean of inter-generational ability transmission, I regressed children's ability on parents' ability in NLSY79 to obtain the parameter 0.46[10]. A standard deviation increase in parent's ability leads to an increase in children's high school ability by .46 of a standard deviation.

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 (2015), the education cost is education and general (E&G) category which excludes dormitories and hospitals. The education cost per student is \$17,187 in 2000.

Since the Federal Pell subsidy 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. (2018)), I assume subsidies are not merit-based in the status-quo.

I adopt Abbott et al. (2018) for the cost of college for enrollees and the subsidy system of the Unites States (see Table 4 for 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 cost of education above and the cost for enrollees, $\$17,187 - \$6,710 = \$10,477$. In the model, the subsidies for enrollees are the sum of this subsidy and the subsidies as in Table 4 which is denoted by $\bar{s}(q)$. In the current system, college subsidies are constant across periods in college and $s_1(q) = s_2(q) = \bar{s}(q)$.

The largest federal loan program in the US is the Federal Family Education Loan

[10]For college dropouts and college graduates, $\theta = \theta_c$ but I use ln AFQT80 as an instrument as in the estimation of labor productivity process.

q	family income	subsidies to students	subsidies to colleges	total $\bar{s}(q)$
1	- \$30,000	\$2,820	\$10,477	\$13,297
2	\$30,000 - \$80,000	\$668	\$10,477	\$11,145
3	\$80,000 -	\$143	\$10,477	\$10,620

Table 4: subsidies and family income

Program. Among federal loans, the Stafford loan program was the most common for the undergraduates so I focus on Stafford loans. A Stafford loan can be either subsidized or unsubsidized. The difference between these two is interest payments during college but borrowers have to pay interest after college for either type. I focus on unsubsidized loan. Students' interest rate is the prime rate plus 2.3% ($= \iota^s$, annual). I assume students face a borrowing limit dependent on age. The annual Stafford loan limits 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. The borrowing limits for workers are based on self-reported limits on unsecured credit by education level from 2001 Survey of Consumer Finances.

3.4 Government Policy

The government consumption and investment over GDP in the United States in 2000 is 17.8% from Bureau of Economic Analysis. Since the government expenditure on teritiary education in the United States in 2000 is 0.7% of GDP (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)).

Parameters	Interpretation	Value
γ	Coef of relative risk aversion	4
$\bar{h}$	Study time	0.25
ζ	Adult equivalence scale	0.3
α	Capital share	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)	\$20,754
$\underline{A}^{CD}$	Borrowing constraint, CD (SCF)	\$24,833
$\underline{A}^{CG}$	Borrowing constraint, CG (SCF)	\$37,832
τ_c	Consumption tax rate	7%
τ_k	Capital income tax rate	27%
g	Gov cons to GDP ratio	17.1%

Table 5: Parameters determined outside the model.

3.5 The Remaining Parameters

Given the parameter values set outside the model in Table 5, there are 16 remaining parameters: bias of expectation of college ability (μ_c^0, μ_c^1), college utility $(\lambda^0, \lambda^1, \lambda_1^\epsilon, \lambda_2^\epsilon)$, the variance of college ability σ_c, productivity of labor (a^S, ϵ^{CD}), education cost κ, utility parameters (μ, β, v), lump-sum transfer d, overseeing cost ι, and inter-generational ability parameters $(\bar{m}, \sigma_h)$.

I choose 27 moments in Table 7 and minimize the average Euclidean percentage deviation of the model from the data[11]. The enrollment and graduation rates across ability and family income and the skill premiums are the main theme of the paper. Optimism is a key driver of college dropouts and I try to match the difference between the graduation rates students expect and the actual one. According to Stinebrickner and

[11] For the mean of high school ability, I chose 5.03, which is the mean of ln AFQT80 before normalization, for the denominator of the percent deviation. I do not take the percent deviation for the enrollment and graduation rates.

Stinebrickner (2012), on average, students of the college they survey believe that there is an 86% chance of graduating while approximately 60% of students graduate. The percent difference is $43\%(= 0.86/0.60 - 1)$. Since the educational decisions are strongly dependent on ability, matching the mean and standard deviation of high school ability are also important.

The third column of Table 6 presents the calibrated values. The calibrated value of μ_c is positive. Enrollees are optimistic about their college ability on average. Since the standard deviation of college ability is 0.38[12], the bias for the mean ability is 50% of the standard deviation of college ability. In addition, enrollees with lower high school ability are more optimistic than enrollees with higher high school ability. These characteristics are consistent with the bias of college GPA observed in Stinebrickner and Stinebrickner (2012). λ^0 is negative and agents derive disutility from college. A positive λ^1 implies that the disutility is smaller for agents with high ability than agents with low ability.

The model fit is presented in Table 7 and Figures 2 and 3. In general, the model fits well considering over-identification of 16 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 the graduation rates across family income are somewhat flatter than the data, they capture the key pattern. The enrollment and graduation rates are higher for the second quartile than for the third quartile. This might be because there are only three bins for family income q and there is a jump of subsidies when people cross over the threshold of family income.

3.6 Validation Exercises

Partial Equilibrium Effect of Year-Invariant subsidies

The elasticity of enrollment with regard to tuition or subsidies has been extensively

[12] The square root of the sum of the variance of high school ability and σ_c^2. $0.38 = \sqrt{0.212^2 + 0.318^2}$

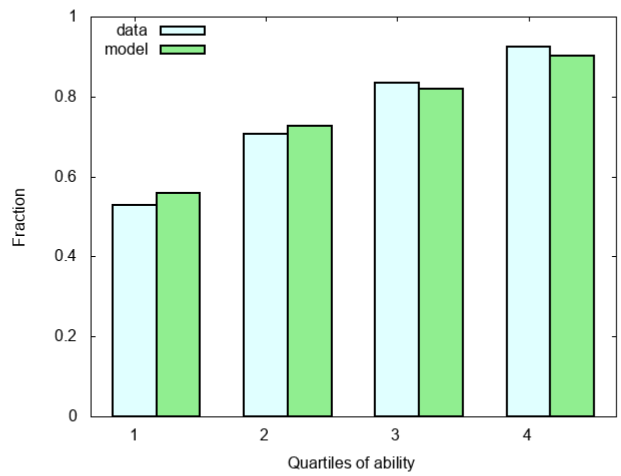
Parameter	Description	Value
μ_c^0	college ability bias intercept	0.191
μ_c^1	college ability bias slope	-0.423
λ	college utility intercept	-18.3
λ^θ	college utility slope	212
λ_1^ϕ	first period college taste	57.4
λ_2^ϕ	second half college taste	41.4
a^S	productivity of skilled labor	0.473
ϵ^{CD}	productivity of CD	1.04
σ_c	s.d. of college ability	0.318
κ	education cost	0.209
μ	consumption share of preference	0.415
β	time discount rate	0.939
v	altruism	0.0958
d	lump-sum transfer ratio	0.125
ι	borrowing wedge ($r^- = r + \iota$)	18.3%
m	intergenerational ability transmission intercept	-0.0417
σ_h	intergenerational ability transmission s.d.	0.169

Table 6: Parameters calibrated.

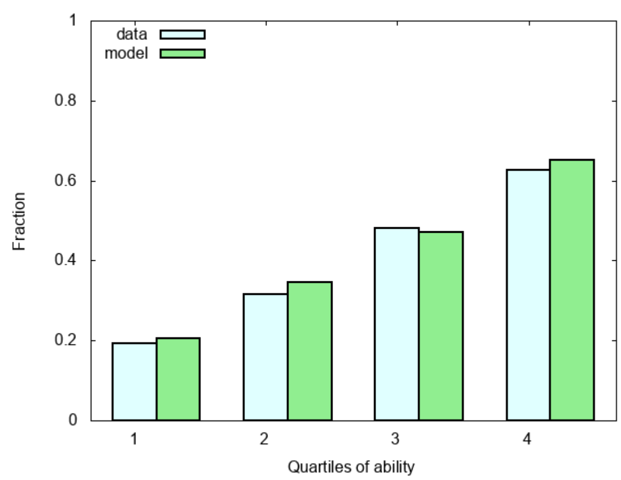
Moment	Model	Target
Enrollment rate of ability quartile	(figure)	(figure)
Graduation rate of ability quartile	(figure)	(figure)
Enrollment rate of family income quartile	(figure)	(figure)
Graduation rate of family income quartile	(figure)	(figure)
Skill premium for CG[a]	89.7%	89%
Skill premium for CD	19.8%	20%
Expected/Actual graduation rate -1	0.436	0.433
Education cost/mean income at 48	0.311	0.33
Hours of work	33.6%	33.3%
K/Y	1.310	1.325
Transfer/mean income at 48	66.7%	66%
log pre-tax/post-tax income	60.5%	61%
Borrowers	6.81%	6.3%
Mean of AFQT	-0.0077	0
Standard deviation of AFQT	0.212	0.21

Table 7: Moments matched.

[a] The skill premiums are from full-time workers in Current Population Survey (CPS) IPUMS (Flood, King, Rodgers, Ruggles and Warren (2018))



(a) Enrollment rate

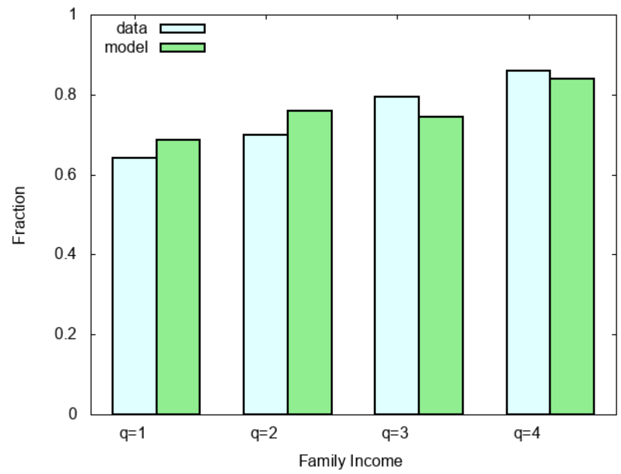


(b) Graduation rate

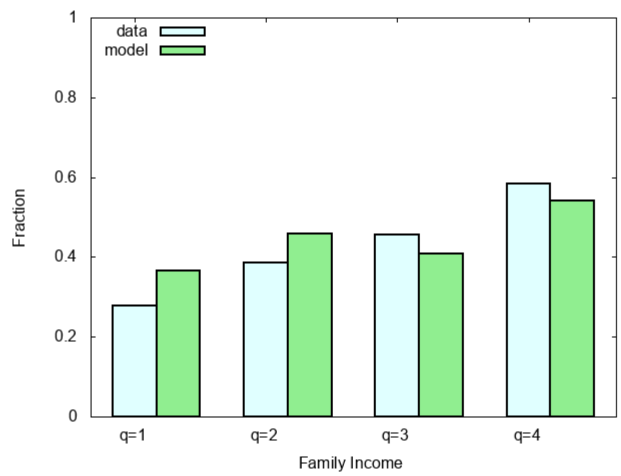
Figure 2: Model fit: enrollment and graduation rate for each ability quartile.

examined in the micro empirical literature. I simulate the partial equilibrium response of enrollment to a $1,000 increase in subsidies for all the college years and family income evenly. All the prices and the distribution of initial state are fixed at the current level and additional subsidies are given to only one generation.

The aggregate enrollment rate of the affected generation increases by 1.11 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



(a) Enrollment rate



(b) Graduation rate

Figure 3: Model fit: enrollment and graduation rates for each family income quartile.

(2001). While this literature argues that the enrollment rate of groups benefitting from an additional subsidy of $1,000 increases by between 3 to 6 percentage points, Hansen (1983) and Kane (1994) argue that there is less evidence of a rise in college enrollment of the target of the Pell Grant program (See Kane (2006) for the empirical literature). Therefore the simulation is broadly in the range of the literature. In addition, the increase in enrollment is smaller in the model than in the data, which implies this calibration is a more conservative choice. If the response of enrollment is high, the effect of changing subsidies is also high and I overestimate the effect of switching to year-dependent subsidies.

Since this paper studies the effect of subsidies on graduation, it is interesting to know how an increase in enrollment leads to changes in the shares of college graduates and dropouts separately. An increase in subsidies has two effects. Enrollment can increase due to an increase in subsidies for the first half period. Graduation can increase due to an increase in subsidies for the second half. In the simulation, the share of college graduates increases by 0.49 percentage points and that of college dropouts increases by 0.62 percentage points. Not only the people who are induced to enroll by the additional subsidy but also those who would already have enrolled without the additional subsidy have incentive to stay until graduation. This is consistent with Dynarski (2008), Castleman and Long (2016), Scott-Clayton (2011), Scott-Clayton and Zafar (2016), Denning, Marx and Turner (2018), and Sacerdote, Bettinger, Kawano, Gurantz and Stevens (2019) who all find a positive effect of subsidies on graduation.

Sluggish Increase in College Graduates

The sluggish increase in college graduates in the United States between 1980 and 2000 is a crucial factor to explain the increase in the skill premium. In this subsection, I examine how well the model can explain this sluggish increase. The benchmark calibration is targeted to the United States in 2000 and I assume 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 46.0% and the college dropout wage premium 12.4% as observed in 1980 in the United States with the other parameter values fixed. I compute the steady state with the new values and call it “1980 steady state.” The first two rows of Table 8 show that the wage premiums for college graduates and dropouts in the model and the data. By definition, the change in the model and in the data match. While I target the change

	1980	2000	change (model)	change (data)
college graduate premium	46.0%	89.7%	43.7pp	43.2pp
college dropout premium	12.4%	19.8%	7.4pp	7.4pp
share of college graduates	29.3%	34.6%	5.3pp	9.81pp
share of college dropouts	42.4%	40.8%	-1.6pp	-0.14pp

Table 8: Change in the share of college graduates and dropouts

in the skill premium, I use the change in the share of college graduates and dropouts as non-targeted moments to compare with the data.[13]

The third and fourth rows of Table 8 show the change in the share of college graduates and dropouts between "1980 steady state" and the benchmark calibration targeted to year 2000. As the college graduate premium increases by 43.7 percentage points from 1980 to 2000, the third column shows the share of college graduates increases by 5.3 percentage points. Although this is smaller than the data in the fourth column, the model can explain the sluggish increase in the share of college graduates. Interestingly, the share of college dropouts does not change in the model with the college dropout premium increasing, which is consistent with the data. The increase in the college graduate wage premium cancels out the effect of the increase in the college dropout wage premium.

4 Results

The 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 it affects enrollment, graduation, and the skill premium. In the second exercise, I keep total spending fixed but choose subsidies by

[13] I use the Current Population Survey IPUMS for the wage premiums in 1980 and the change in the shares of college graduates and dropouts between 1980 and 2000. For the shares of college graduates and dropouts, I follow the definition of Castro and Coen-Pirani (2016).

G_e	0.75 $\bar{G}_e$	$\bar{G}_e$	$1.5\bar{G}_e$	$2\bar{G}_e$
enrollment rate	73.7%	75.4%	77.9%	79.1%
share of college graduates	33.8%	34.6%	35.9%	36.9%
skill premium	94.5%	89.7%	83.0%	77.3%

Table 9: The elasticity of education: year-invariant subsidies

year to maximize the number of college graduates in the steady state. In the third exercise, I keep total spending fixed and choose subsidies to maximize the expected lifetime value of new borns in the steady state.

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 year-invariant subsidies on educational choice and the skill premium at the new stationary equilibrium. First, $\bar{G}_e$ denotes the current level government total budget or expenditure for college subsidies. I show how the share of college enrollees at age 1 and the share of college graduates at age 2 change as the government subsidies expenditure G_e changes exogenously as follows: $G_e = 0.75\bar{G}_e, \bar{G}_e, 1.5\bar{G}_e, 2\bar{G}_e$. The proportional labor income tax rate τ_l is adjusted to the changes in the total budget. The subsidies across college years and family income proportionally change with G_e fixed.

The first row of Table 9 displays the proportion of agents who enroll in college at the beginning of age 1 for each budget level, which measures enrollment in the economy. The second row of Table 9 displays the proportion of agents who graduate from college among the whole population aged 2, which measures the share of college graduates in the economy. Both enrollment and the share of college graduates increase as total budget increases. Since skilled and unskilled labor are incomplete substitutes and the supply of skilled labor increases, the skill premium decreases.

4.2 The Effect of Year-Dependent Subsidies

In this subsection, I derive the year-dependent subsidies that maximize the number of college graduates. We are interested in the relative amount of subsidies across college years rather than the level of total subsidies. Thus I fix the total spending at the current level and only allow the relative sizes of subsidies to change across college years. The maximization problem is formulated as

$$\max_{g_1, g_2, \tau_\ell} \int_{S_2^{CG}} d\mu_2^{CG} \tag{47}$$

subject to

$$\int_{S_1^c} g_1 \bar{s}(q) d\mu_1^c + \int_{S_2^c} g_2 \bar{s}(q) d\mu_2^c = G_e \tag{48}$$

and the government budget constraint. The objective of this problem is the share of college graduates in the society. 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 (note that current subsidies are independent of years in college). In this problem, the government chooses the general levels of college subsidies for each period compared to the current system, g_1 and g_2. If I increase g_1 (subsidies for the first half of college), the general level of subsidies for the second half g_2 has to decrease. Since the composition of education changes, the aggregate labor income changes and τ_ℓ needs to be adjusted to balance the government budget even though the budget for college subsidies are fixed.

Note that the space of the college subsidy system available in this problem is restricted. In particular, I do not allow to change the relative subsidies across family income and ability from the status-quo. For example, the ratio of subsidies for $q = 1$ and $q = 2$ is fixed at the current state. I want to focus on how the year-dependency of college subsidies affects educational choices and the skill premium independently from other changes such as the relative subsidies across family income levels.

$s_j(q)$	year-invariant	year-dependent
$s_1(1)$	\$13,599	\$57
$s_1(2)$	\$11,447	\$48
$s_1(3)$	\$10,922	\$46
$s_2(1)$	\$13,599	\$41,090
$s_2(2)$	\$11,447	\$34,587
$s_2(3)$	\$10,922	\$33,001

Table 10: Year-dependent subsidies maximizing the number of college graduates.

Table 10 shows the amount of annual year-invariant and year-dependent subsidies. The first column is identical to the case of year-invariant subsidies with $G_e = \bar{G}_e$ and the second column is the solution to the problem maximizing the share of college graduates. The first three rows are college subsidies at the first half across family income level $q = 1$ to 3 from the top to the bottom. The next three rows are college subsidies at the second half in the same way. In the year-dependent case, the optimal subsidies are back-loaded: subsidies are more generous for the second half than for the first half. The year-dependent subsidies for the first half is almost negligible.

The rows 1 and 2 of Table 11 display the enrollment rate the share of college graduates for each case. Year-dependent subsidies reduce the enrollment rate by 5.3 percentage points and increase the share of college graduates by 2.3 percentage points. The share of college graduates with the year-dependent subsidies is as much as the case with doubling the budget of the current year-invariant subsidies. The skill premium for college graduates is 80.2%, which is between the premiums attained by the year-invariant cases of $G_e = 1.5\bar{G}_e$ and $G_e = 2\bar{G}_e$. With the total budget for college subsidies fixed, switching from year-invariant to year-dependent subsidies is as effective in increasing college graduates as increasing the total spending of year-invariant subsidies by 50%. Changing the structure of college subsidies has as much a power effect as increasing the budget.

	year-invariant $\bar{G}_e$	year-invariant $2\bar{G}_e$	year-dependent $\bar{G}_e$
enrollment rate	75.4%	79.1%	70.1%
share of college graduates	34.6%	36.9%	36.6%
skill premium	89.7%	77.3%	80.2%

Table 11: The elasticity of education: optimal mix

The mechanism of the effect of year-dependent subsidies is the following. In the current system, already 70% of people 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 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. The higher subsidies for the second period encourages marginal dropouts to finish.

There is another mechanism of back-loaded subsidies. In the current system, the government has paid subsidies to all the people who enroll but 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 college graduates, which increases the number of college graduates. In fact, as Table 10 shows, with the back-loaded subsidies, the sum of the subsidies to college graduates for the two periods for the middle family income is \$34,635 which is higher than the case of the current system \$22,894. The back-loaded subsidies are more cost-effective from the perspective of increasing college graduates.

4.3 Welfare Analysis of Year-dependent subsidies

In this subsection, I examine how year-dependent subsidies can improve welfare. As in the previous section, I fix the total budget on college subsidies at the current level and examine how welfare improves by only varying the relative sizes of subsidies across college years. The optimization problem for the optimal policy is

$$\max_{g_1,g_2} \int_{S_0} V_0^{sp}(a,\theta_h,\eta,q,\phi)d\mu_0 \tag{49}$$

subject to

$$\int_{S_1^c} g_1\bar{s}(q)d\mu_1^c + \int_{S_2^c} g_2\bar{s}(q)d\mu_2^c = G_e \tag{50}$$

and the government budget constraint. This problem maximizes the sum of the value of newborns with an equal weight, which is the concept of utilitarian. V_0^{sp} is the value function of newborns with no bias ($\mu_c(\theta_h) = 0$ for all θ_h). This assumption implies that the government implementing the optimal policy evaluates the expected lifetime value with rational expectation, which is different from the value agents expect before enrollment.

Table 12 displays the optimal college subsidies. The optimal subsidy is back-loaded and the amount for the second half is 1.3 times the subsidy for the first half. As in the previous section, the first and second rows of Table 13 show that enrollment decreases by 0.5 percentage points and the share of college graduates increases by 0.2 percentage points by switching to the optimal policy. The skill premium decreases by 1.1 percentage points.

To examine the welfare effect of the optimal policy, I use lifetime consumption equivalence as a summary measure of welfare. Let $\tilde{V}_0^{sp}(\boldsymbol{c},\boldsymbol{h};\boldsymbol{s}_0)$ be expected lifetime utility at age 0 with the path of consumption $\boldsymbol{c}$, leisure $\boldsymbol{h}$ with the initial state $\boldsymbol{s}_0$ with

	Status-quo	Optimal
$s_1(1)$	\$13,599	\$12,529
$s_1(2)$	\$11,447	\$10,547
$s_1(3)$	\$10,922	\$10,063
$s_2(1)$	\$13,599	\$16,088
$s_2(2)$	\$11,447	\$13,542
$s_2(3)$	\$10,922	\$12,921

Table 12: Optimal policy of year-dependent subsidies.

	Status-quo	Optimal
share of college enrollees	75.4%	74.9%
share of college graduates	34.6%	34.8%
skill premium	89.7%	88.6%

Table 13: The effect of the optimal year-dependent subsidies with correcting bias.

no optimism. Then lifetime consumption equivalence is defined as ω_{tot} such that

$$\int_{S_0} \tilde{V}_0^{sp}(\boldsymbol{c}^B, \boldsymbol{h}^B; \boldsymbol{s}_0) d\mu_0^B = \int_{S_0} \tilde{V}^{sp}((1+\omega_{tot})\boldsymbol{c}^A, \boldsymbol{h}^A; \boldsymbol{s}_0) d\mu_0^A \tag{51}$$

In addition, as in Bénabou (2002), I decompose the lifetime consumption equivalence into three parts: (i) *a level effect* which measures the gain in aggregate consumption, leisure, and college utility (ii) *an uncertainty effect* which measures the effect of volatility of consumption and leisure paths on utility of risk-averse agents with incomplete markets, and (iii) *an inequality effect* which measures the distribution of initial conditions. The sum of these three effects might not necessarily be the toal welfare effect. I follow Abbott et al. (2018) in detail.

	Total	Level	Uncertainty	Inequality
Optimal	+0.12%	+0.06%	+0.05%	−0.05%

Table 14: Welfare decomposition when correcting bias.

	Status-quo	Optimal
Y	0.316	0.316
K	0.415	0.414
C	0.210	0.209
w^{CG}	0.367	0.366
w^{CD}	0.397	0.397
w^{HS}	0.397	0.397
std c	0.128	0.127
std a	0.476	0.475
std h	0.0837	0.0836
std wage	0.541	0.538

Table 15: The aggregates under the optimal year-dependent subsidies.

The welfare gain is decomposed in Table 14. First, the total welfare gain for newborns is 0.12%. The level effect is 0.06% and there is an efficiency gain while output, capital, and consumption decrease (Table 15). In the current system, individuals are over-optimistic and there is an excessively large amount of college enrollees. The optimal back-loaded subsidies screen people who enroll. By reducing subsidies for the first half, thew enrollee with low ability stop enrolling, which reduces college disutility of low ability enrollees. Optimism is a key factor for the optimal college subsidies. In the Appendix E, I calibrate the case without optimism and examine how the assumption about optimism matters.

The uncertainty effect is 0.05% as there is less uncertainty under the optimal policy. Due to a smaller skill premium, there is a less difference in wages between college graduates and dropouts. The policy can reduce the uncertainty of lifetime income from dropout decisions.

The inequality effect is -0.05% as there is more inequality across ex ante heterogeneous agents at the initial period under the optimal policy. It is counter-intuitive because the skill premium decreases by 1.1 percentage points and the standard deviations of consumption, asset, hours, and wages per hour decreas as shown in Table 15.

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

Table 16: Lifetime consumption equivalence variation for newborns.

Although inequality as of period 1 increases, cross-section inequality in the economy decreases under the optimal policy.

In order to see why it increases inequality as of the initial state, I calculate the welfare gain for each ability and family income level in Table 16[14]. Given family income, the welfare gain is greater for people with low ability. Since the price of effective labor for high school graduates and college dropouts increases as in Table 13, the welfare of agents with low ability increases more than other agents.

Given the same ability level, the welfare gain is greater for high family income, which is consistent with the negative inequality effect[15]. Enrollees from poor family 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 the consumption by agents from poor family during the first period of college. But the optimal policy does not prevent people who have enough ability to expect to graduate from enrolling for the following reasons. First, there is a correlation between ability and initial assets: high ability rich parents are more likely to have high ability children and give a larger transfer. Therefore people with high ability can use the transfer to smooth consumption during the first period of college when the credit limit is tight.

[14] The distribution of ability is different between the status-quo and the optimal case because the share of college graduates changes the mean ability of the future generation. Each ability quartile on the table is the quartile of the status-quo.

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

	% of subsidy loss
Subsidies	-100%
Labor income	+19%
(Price of hours)	+11%
(Leisure)	(-0.0018)
Transfer from parents	+0.07%
-Savings	+74%
-Tuition	+4%
Consumption	-5%

Table 17: Change in each item of the income.

In order to see how people react to the loss of college subsidies for the first period, Table 17 shows the average change in each part of the earnings and consumption for an individual with $\theta_h = 0$, $q = 1$, $\eta = \eta_H$, 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. First, the labor income increases and covers a 19% of the loss of college subsidies because the wage of unskilled college enrollees is higher due to the smaller skill premium given the working hours fixed. The third row presents that the change in the potential labor income if the agent works for the fixed hours $1 - \bar{h}$ as a ratio to the loss of college subsidies. In addition, under the optimal policy, agents work for longer hours to mitigate the loss of college subsidies. As the fourth row shows, they cut their leisure by 0.0018 out of the unit hour endowment. Second, since the college subsidies are shifted to the second half of college, savings for the second half of college are reduced. This covers a 74% of the loss of college subsidies Third, tuition of college decreases due to the lower wage of skilled labor under the optimal policy, which covers a 4% of the loss of college subsidies. In total, the agents can mitigate the loss of the college subsidies by 95%($= 100\% - 5\%$) for consumption. These results are consistent with the findings by Keane and Wolpin (1997) and Garriga and Keightley (2007).

	Total	Level	Uncertainty	Inequality
correct bias	+0.80%	−6.07%	+3.55%	−1.56%
correct bias (Optimal)	+1.07%	−5.56%	+3.39%	−1.51%

Table 18: Welfare decomposition when correcting bias.

	Status-quo	Correcting bias	Optimal
share of college enrollees	74.3%	47.4%	48.5%
share of college graduates	32.2%	28.3%	27.8%
skill premium	88.6%	120%	124%

Table 19: The elasticity of education: optimal mix

4.4 Correcting Bias

A large part of the welfare gain of year-dependent subsidies originates in optimism. 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 year-dependent subsidies. In this subsection, I show what is the welfare gain by correcting bias and compare it with year-dependent subsidies without correcting bias.

The first row of Table 18 shows the welfare gain from correcting bias with the current subsidies. As in the second column of Table 19, the enrollment rate drops significantly after correcting bias. Without optimism, enrollment is excessively small because there is a borrowing constraint and no insurance available about 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 of the first period in college. After correcting bias, the second effect is greater and enrollment is excessively small. On the other hand, less people enroll and have dropout risk, which leads to a positive uncertainty effect. Due to the loss of enrollment and college graduates, the skill premium increases and the inequality effect is negative. In total, the loss from a large skill premium and excessively small enrollment is less the gain from avoiding the excessive enrollment from

	Status-quo	Optimal
$s_1(1)$	\$19,832	\$24,089
$s_1(2)$	\$16,694	\$20,277
$s_1(3)$	\$15,928	\$19,347
$s_2(1)$	\$19,832	\$11,962
$s_2(2)$	\$16,694	\$10,069
$s_2(3)$	\$15,928	\$9,608

Table 20: Optimal subsidies with correcting bias

optimism.[16]

To examine whether 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 the second column of Table 20. The optimal subsidy is *front-loaded*: greater for the first period than for the second period. As in the third column of Table 19, by subsidizing college in the first period, the policy can mitigate uncertainty of college ability and increase enrollment. Even if the government can correct bias and do not need to implement back-loaded subsidies, there is a welfare gain by using year-dependent subsidies on the other way around: front-loaded subsidies.

Although the welfare gain for newborns is higher with correcting bias than the case of year-dependent subsidies without correcting bias, it does not imply that correcting bias is beneficial to the whole population. I use the social welfare function adding up the sum of the value of the whole population (not only newborns but also older agents) and derive the welfare measure of lifetime consumption equivalence. Table 21 displays that the welfare gain for the whole population is significantly *negative* if correcting bias. The skill premium is higher than the current state when correcting bias and it leads to an increase in wage inequality across the whole population and reduces welfare.

[16] 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 social planner who does not have optimism even before correcting bias.

	optimal	no bias	no bias optimal
Optimal	$+0.15\%$	-8.03%	-7.83%

Table 21: Welfare for the whole population.

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 decreasing 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. With back-loaded subsidies, the number of college graduates increases and the skill premium of college graduates decreases. Switching to back-loaded subsidies, with the total budget fixed, can increase the fraction of college graduates and reduce the skill premium more than doubling year-invariant subsidies. Back-loaded subsidies improve welfare without increasing the total budget of college subsidies and increasing tax.

While this paper has focused on the structure of subsidies as a policy tool to decrease inequality, redistribution through progressive taxation can also reduce consumption inequality. A future work is contrasting progressive taxation with college subsidies to combat the skill premium. While there is literature on the optimal progressive taxation and college subsidies, they often abstract from college dropout and dropout might affect the optimal taxation and college subsidies as follows. Increasing subsidies might end up with more college dropouts, which does not lead to a decrease in the skill premium. While a progressive income tax reduces the risk of dropout, it might create incentive to drop out.

Although this paper focuses on college, the mechanics is applicable to other education levels such as more than college graduation. Increasing subsidies for post-college might lead to an increase in workers of post-college, which affects the distribution of skill and wages. Changing the amount of subsidies within education before college might also have effects. More generally, age-dependent subsidies to human capital investment after finishing schooling could be beneficial under a similar mechanism to this paper. Subsidies dependent on education levels have potential to be an important policy intervention.

References

Abbott, Brant, Giovanni Gallipoli, Costas Meghir, and Giovanni L Violante, "Education Policy and Intergenerational Transfers in Equilibrium," *Cowles Foundation Discussion Paper No.1887R*, 2018.

Akyol, Ahmet and Kartik Athreya, "Risky Higher Education and Subsidies," *Journal of Economic Dynamics and Control*, 2005, *29* (6), 979–1023.

Arcidiacono, Peter, Esteban Aucejo, Arnaud Maurel, and Tyler Ransom, "College Attrition and The Dynamics of Information Revelation," *NBER Working Paper No. 22325*, 2015.

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

Bénabou, Roland, "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, 2002, *70* (2), 481–517.

Bovenberg, A. Lans and Bas Jacobs, "Redistribution and education subsidies are Siamese twins," *Journal of Public Economics*, 2005, *89* (11-12), 2005–2035.

Cameron, Stephen V and James J Heckman, "The Dynamics of Educational Attainment for Black, Hispanic, and White Males," *Journal of Political Economy*, 2001, *109* (3), 455–499.

Castex, Gonzalo, "College risk and return," *Review of Economic Dynamics*, 2017, *26*, 91–112.

Castleman, Benjamin L and Bridget Terry Long, "Looking Beyond Enrollment: the Causal Effect of Need-Based Grants on College Access, Persistence, and Graduation," *Journal of Labor Economics*, 2016, *34* (4), 1023–1073.

Castro, Rui and Daniele Coen-Pirani, "Explaining the Evolution of Educational Attaiment in the United States," *American Economic Journal: Macroeconomics*, 2016, *8* (3), 77–112.

Caucutt, Elizabeth M. and Krishna B. Kumar, "Higher education subsidies and heterogeneity: A dynamic analysis," *Journal of Economic Dynamics and Control*, 2003, *27* (8), 1459–1502.

Chatterjee, Satyajit and Felicia Ionescu, "Insuring Student Loans against the Financial Risk of Failing to Complete College," *Quantitative Economics*, 2012, *3* (3), 393–420.

Colas, Mark, Sebastian Findeisen, and Dominik Sachs, "Optimal Need-Based Financial Aid," 2018.

Cunha, Flavio, "Investments in Children When Markets Are Incomplete *," 2013, pp. 1–39.

_ **and James Heckman**, "The Technology of Skill Formation," *American Economic Review*, 2007, *97* (c), 31–47.

Daruich, Diego, "The Macroeconomic Consequences of Early Childhood Development Policies," 2017.

Denning, Jeffrey T, Benjamin Marx, and Lesley Turner, "ProPelled: The Effects of Grants on Graduation, Earnings, and Welfare," 2018.

Dynarski, Susan, "The Behavioral and Distributional Implications of Aid for College," *American Economic Review*, 2002, *92* (2), 279–285.

_, "Building the Stock of College-Edicated Labor," *The Journal of Human Resources*, 2008, *43* (3), 576–610.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren, "Integrated Public Use Microdata Series, Current Population Survey: Version 6.0 [dataset]. Minneapolis, MN: IPUMS," 2018.

Garriga, Carlos and Mark P. Keightley, *A General Equilibrium Theory of College with Education Subsidies, In-School Labor Supply, and Borrowing Constraints* 2007.

Goldin, Claudia and Lawrence F. Katz, "The Race between Education and Technology: A Review Article," *NBER Working Paper*, 2007, (March 2007), 1–44.

Hansen, W.L., "Impact of Student Financial Aid on Access," *Proceedings of the Academy of Political Science*, 1983, *35* (2), 84–96.

Hanushek, Eric A., Charles Ka Yui Leung, and Kuzey Yilmaz, "Borrowing Constraints, College Aid, and Intergenerational Mobility," *Journal of Human Capital*, 2014, *8* (1), 1–41.

Heckman, James J., Lance J. Lochner, and Petra E. Todd, "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond," *Handbook of the Economics of Education*, 2006, *1* (06), 307–458.

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

Horn, Laura and Paul Skomsvold, "Community College Student Outcomes: 1994-2009," *National Center for Education Statistics*, 2012, (November 2011), Technical Report 2012–253.

Ionescu, Felicia, "Risky human capital and alternative bankruptcy regimes for student loans," *Journal of Human Capital*, 2011, *5* (2), 153–206.

Jones, John Bailey and Fang Yang, "Skill-Biased Technical Change and the Cost of Higher Education," *Journal of Labor Economics*, 2015, *34* (3), 0.

Kane, Thomas J, "College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education," *Journal of Political Economy*, 1994, *102* (5), 878.

Kane, Thomas J., "Public Intervention in Post-Secondary Education," *Handbook of the Economics of Education*, 2006, *2* (06), 1369–1401.

_ **and Cecilia Elena Rouse**, "Labor-Market Returns to Two- and Four-Year College," *The American Economic Review*, 1995, *85* (3), 600–614.

Katz, Lawrence F. and Kevin M. Murphy, "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 1992, *107* (1), 35–78.

Keane, Michael P and Kenneth I Wolpin, "The Career Decisions of Young Men," *Journal of Political Economy*, 1997, *105* (3), 473–522.

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

Lee, Sang Yoon, Yongseok Shin, and Donghoon Lee, "The Option Value of Human Capital: Higher Education and Wage Inequality," nov 2015, (21724).

Manski, Charles F., "Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon," *Economics of Education Review*, 1989, *8* (4), 305–312.

McDaniel, Cara, "Average tax rates on consumption, investment, labor and capital in the OECD 1950-2003," 2007, (March), 1–37.

Sacerdote, Bruce, Eric Bettinger, Laura Kawano, Oded Gurantz, and Michael Stevens, "The Long-Run Impacts of Financial Aid: Evidence from California's Cal Grant," *American Economic Journal: Economic Policy*, 2019, *11* (1), 64–94.

Scott-Clayton, Judith, "On money and motivation a quasi-experimental analysis of financial incentives for college achievement.," *Journal of Human Resources*, 2011, *46* (3), 614–646.

_ **and Basit Zafar**, "Financial Aid, Debt Management, and Socioeconomic Outcomes: Post-College Effects of Merit-Based Aid," 2016.

Stange, Kevin E., "An Empirical Investigation of the Option Value of College Enrollment," *American Economic Journal: Applied Economics*, 2012, *4* (1), 49–84.

Stinebrickner, Ralph and Todd Stinebrickner, "The Effect of Credit Constraints on the College Drop-Out Decision: A Direct Approach Using a New Panel Study," *American Economic Review*, 2008, *98* (5), 2163–2184.

— **and** — , "Academic Performance and College Dropout: Using Longitudinal Expectations Data to Estimate a Learning Model," *Journal of Labor Economics*, 2014, *32* (3), 601–644.

Stinebrickner, Todd and Ralph Stinebrickner, "Learning about Academic Ability and the College Dropout Decision," *Journal of Labor Economics*, 2012, *30* (4), 707–748.

Torpey, Elka and Audrey Watson, "Education level and jobs: Opportunities by state: Career outlook," *Career Outlook, BLS*, 2013, (September), 1–12.

Trachter, Nicholas, "Stepping stone and option value in a model of postsecondary education," *Quantitative Economics*, 2015, *6* (1), 223–256.

A Computation of Stationary Equilibrium

This section describes the method of computing an equilibrium. Prices are normalized such that the average annual income of high school graduates at age 48 is \$51,933.

1. Starting from an initial vector of aggregate variables $\boldsymbol{w} = \left(\frac{K}{H}, \frac{H^S}{H}, H, \tau_l\right)$, compute prices r, w_S, w_U 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 $j = J, \dots, j_b + 1$. The number of grids for assets is 30 and that for high school ability and college ability is 5. The number of grids for college taste is 30[17]. I apply the esndogenous

[17]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 $[-1.1, 1.1]$. The range of grids for college ability is broader because of the higher variance. That of college taste is $[-2, 2]$.

grid method.

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

(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, resolve decision rules of individuals until convergence.

3. I interpolate linearly assets and ability to 80 and 25.

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 $K, H, H^S, \bar{h}$ and τ_ℓ from the government budget constraint and go back to step 2.

B Pension

The average life time income is

$$\hat{y}(e,\theta) = \frac{\sum_{j=2}^{j_r-1} w^e \varepsilon_j^e(\theta,1)\bar{h}}{j_r - 2} \tag{52}$$

The pension formula is given by

$$p(e,\theta) = \begin{cases} s_1\hat{y}(e,\theta) & \text{for } \hat{y}(e,\theta) \in [0, b_1) \\ s_1 b_1 + s_2(\hat{y}(e,\theta) - b_1) & \text{for } \hat{y}(e,\theta) \in [b_1, b_2) \\ s_1 b_1 + s_2(b_2 - b_1) + s_3(\hat{y}(e,\theta) - b_2) & \text{for } \hat{y}(e,\theta) \in [b_2, b_3) \\ s_1 b_1 + s_2(b_2 - b_1) + s_3(b_3 - b_2) & \text{for } \hat{y}(e,\theta) \in [b_3, \infty) \end{cases} \tag{53}$$

where $\bar{h} = 0.333$, $s_1 = 0.9$, $s_2 = 0.32$, $s_3 = 0.15$, $b_1 = 0.22\bar{y}$, $b_2 = 1.33\bar{y}$, $b_3 = 1.99\bar{y}$, $\bar{y} = \$28,793$ annually.

C Labor Productivity Process

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 restrict 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 in the individual (but lower than 10,000 annually). I keep only people who do not report extreme changes of 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 people with 12 years highest grade completed. College dropouts are with highest grade completed between 13 and 15. College graduates are with highest graded completed greater than 16. The estimation result is in Table 22. 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 the productivity at the first period after education is unity.

	HS	CD	CG
Age	.0530181	.0684129	.0955783
	(.0030501)	(.0040353)	(.0036997)
Age^2	-.0005314	-.0006872	-.0009521
	(.0000356)	(.0000474)	(.0000429)

Table 22: Age profile estimates of each education.

For the law of motion of residuals, I use the same 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 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.

Ability is approximated by the log of AFQT80 raw score. To estimate the coefficient on ability in effective labor, I use NLSY79 of 11,864 people. For the ability regression, I restrict samples aged between 25 and 63, which leads to 11,627 people. Since NLSY79 does not include old people, I rely on PSID to estimate the age effects. After the age effect is filtered out, I regress hourly wages on ability for each education levels (HS, CD, and CG). As in the selection of PSID, I keep only people who do not report extreme changes of 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 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 log AFQT80 and year dummies and the whole sample is people whose educations are higher than high school graduates.

Among the people who graduate high school, people with less ability are self-selected as high school graduates. For college dropouts and college graduates, I assume a linear selection equation of log AFQT80 and year dummies and the whole sample is both college dropouts and graduates. Among the people who enroll in college, people with high ability are self-selected as college graduates and people with low ability as college dropouts.

D Intergenerational Ability Transmission

To estimate the transmission of ability from parents to children, I rely on the data from NLSY79 to approximate parents' ability and "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 of test scores of the children of the women in the NLSY79 dataset. The 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 NLSY79. To focus on cognitive ability, I use the PIAT Math to approximate high school ability of children. In particular, I use the standardized PIAT Math score, which adjusts different age in 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 the 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 the respondents whose both AFQT scores and education levels are not missing. I focus on people with high school degrees. This leaves me with 6,193 children born to 2,828 mothers.

E Calibration without Optimism

In this paper, I assume that students are overoptimistic about their college abilities before enrollment and it is a key factor for the large college dropout rate in the United States. In addition, I show that this assumption is crucial to the main results. In this chapter, I examine a different approach to explain the large college dropout in the United States: a large option value of college enrollment. If the uncertainty of college ability given high school ability is large, returns to graduation can be large or small. If it turns out large after enrollment, enrollees can stay in college to earn the high returns to graduation. If it turns out low, enrollees can drop out of college to dismiss the low returns to graduation. This asymmetry of returns increases the benefit of enrollment. I assume $\mu_c(\theta_h) = 0$ for all θ_h and instead assume that the standard deviation of college ability given high school ability is

$$\sigma_c(\theta_h) = \sigma_c \exp(\sigma_c^\theta \theta_h) \tag{54}$$

Table 23 displays the calibrated values under the specification without optimism. As you 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 option value to match the high college dropout rate. The slope of the standard deviation is positive and the uncertainty of college ability is higher for higher high school ability.

Table 24 and Figures 4 and 5 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 needs 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 needs high disutility for low ability people. Then enrollment decreases and college

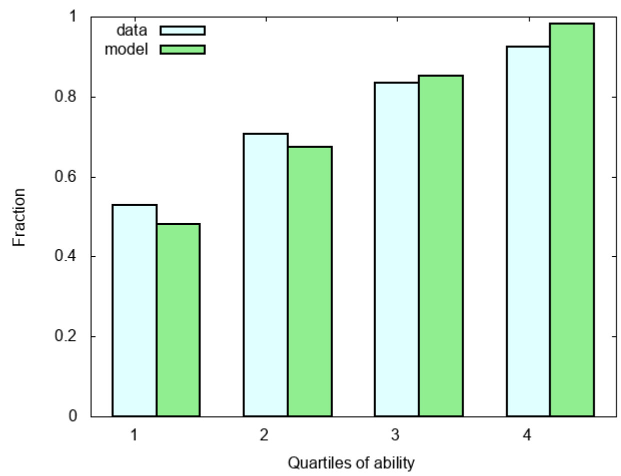
Parameter	Description	Value
λ	college utility intercept	-16.7
λ^{θ}	college utility slope	276
λ_1^{ϕ}	first period college taste	66.3
λ_2^{ϕ}	second half college taste	44.0
a^S	productivity of skilled labor	0.437
ϵ^{CD}	productivity of CD	0.983
σ_c	s.d. of college ability intercept	0.719
σ_c^{θ}	s.d. of college ability slope	0.149
κ	education cost	0.419
μ	consumption share of preference	0.415
β	time discount rate	0.930
υ	altruism	0.0625
d	lump-sum transfer ratio	0.127
ι	borrowing wedge ($r^- = r + \iota$)	18.6%
m	intergenerational ability transmission intercept	-0.0379
σ_h	intergenerational ability transmission s.d.	0.0747

Table 23: Parameters calibrated without optimism

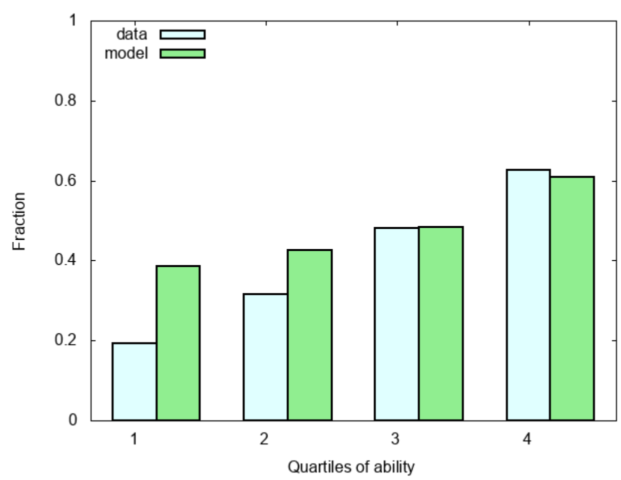
Moment	Model	Target
Enrollment rate of ability quartile	(figure)	(figure)
Graduation rate of ability quartile	(figure)	(figure)
Enrollment rate of family income quartile	(figure)	(figure)
Graduation rate of family income quartile	(figure)	(figure)
Skill premium for CG[a]	89.9%	89%
Skill premium for CD	19.6%	20%
Education cost/mean income at 48	0.314	0.33
Hours of work	32.8%	33.3%
K/Y	1.236	1.325
Transfer/mean income at 48	68.1%	66%
log pre-tax/post-tax income	61.4%	61%
Borrowers	6.62%	6.3%
Mean of AFQT	0.0879	0
Standard deviation of AFQT	0.203	0.21

Table 24: Moments matched.

[a] The skill premiums are from full-time workers in Current Population Survey (CPS) IPUMS (Flood et al. (2018))



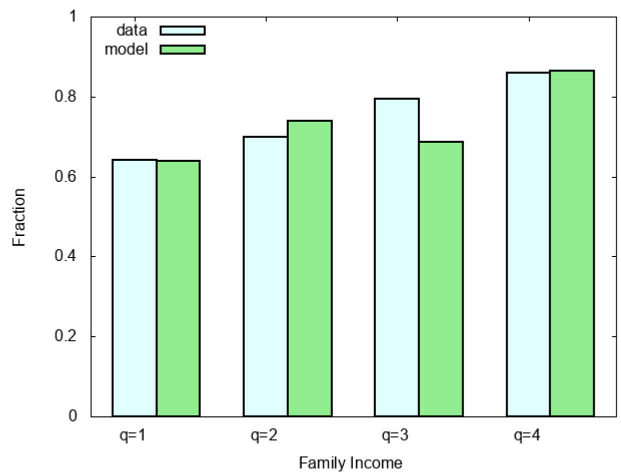
(a) Enrollment rate



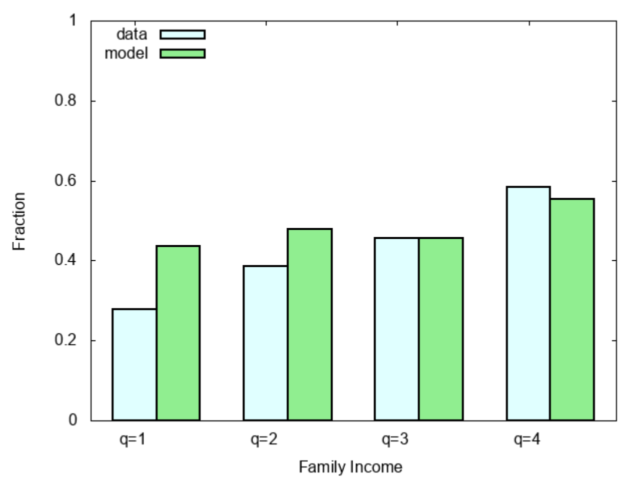
(b) Graduation rate

Figure 4: Model fit: enrollment and graduation rate for each ability quartile.

dropout also decreases. While the effect of increasing the standard deviation of college ability increases college enrollment, it also increases college disutility to match the high college dropout and decreases college enrollment at the same time, offsetting the first effect. Rather the best match requires a *low* standard deviation of college ability and low college disutility to match the high enrollment. To summarize, high option value without optimism does not explain the high dropout rate in the data as well as the model with optimism.



(a) Enrollment rate



(b) Graduation rate

Figure 5: Model fit: enrollment and graduation rates for each family income quartile.

The optimal policy with this formulation is in Table 25. The optimal policy is now *front-loaded* and the subsidies for the first period is twice as much as for the second period. Without optimism, enrollment is excessively low in the status-quo as in the case with correcting bias. This implies that the assumptions about optimism matter and one of the contributions of this paper is to calibrate this effect using available data.

	Status-quo	Optimal
$s_1(1)$	\$19,832	\$24,089
$s_1(2)$	\$16,694	\$20,277
$s_1(3)$	\$15,928	\$19,347
$s_2(1)$	\$19,832	\$11,962
$s_2(2)$	\$16,694	\$10,069
$s_2(3)$	\$15,928	\$9,608

Table 25: Optimal subsidies without optimism