Abstract

While policies seeking to increase college enrollment are often suggested as a means to reduce the rising skill premium and increase the number of college graduates, the impact of these policies is weakened by the fact that 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. The policy change is examined in a quantitative general equilibrium model of heterogenous agents who make endogenous enrollment and dropout decisions. I find that switching to back-loaded year-dependent subsidies with the same total budget increases the number of college graduates and decreases the skill premium more than the case with doubling the total budget of the current subsidies, and are welfare improving despite the fact that enrollment decreases.

1 Introduction

Wage inequality has been increasing in the United States since 1980. In particular, the skill premium—the ratio of wages of college graduates to high school graduates—has increased and contributed to wage inequality. While a college graduate worker earned 50% higher than a high school graduate in 1980, a
current college graduate earns 90% higher than a current high school graduate. A large literature (ex. Goldin and Katz (2007) and Katz and Murphy (1992)) argues that the skill premium is determined as a result of “The Race between Technology and Education.”—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, the speed of the increase in the skill premium can be reduced by increasing college graduates in the economy. Policies seeking to increase enrollment are often suggested as a means to decrease in the skill premium and increase the number of college graduates and the existing literature 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) and enrollment does not necessarily lead to graduation. It is important for us to understand how policy can affect not only enrollment but also graduation.

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. Subsidies that vary by year will have differential impacts on enrollment and graduation from the standard subsidies whose amounts are constant across years (“year-invariant subsidies”). The following examples serve to motivate my focus on year-dependent subsidies. After they graduate from high school, individuals enroll based on their high school GPA or high school ability. One’s high school GPA is not necessarily the same as his/her college GPA. After enrolling, some students learn that their college GPA or college ability is low and so choose to drop out. In this setting, consider back-loaded subsidies: increasing subsidies for the latter years of college and decreasing subsidies for the earlier 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. The existing literature has considered only the subsidies constant across years in college and has not quantifies the effect of year-dependent subsidies. The question of this paper is what timing of college subsidies will maximize the number of college graduates and what timing will maximize welfare.

To examine the effects of year-dependent subsidies on allocations and welfare, I build a life-cycle

---

1Two-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)).
Table 1: College graduation Rates for High school GPA Quartiles

<table>
<thead>
<tr>
<th>HGPA Quantile</th>
<th>% graduation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>19%</td>
</tr>
<tr>
<td>Q2</td>
<td>31%</td>
</tr>
<tr>
<td>Q3</td>
<td>48%</td>
</tr>
<tr>
<td>Q4</td>
<td>63%</td>
</tr>
<tr>
<td><strong>total</strong></td>
<td><strong>42%</strong></td>
</tr>
</tbody>
</table>

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.

general equilibrium model with endogenous enrollment and dropout decisions and credit constraints. Agents who are heterogeneous with regard to initial asset and high school ability make enrollment and dropout decisions. Importantly, agents are assumed to be optimistic with regard to the expectation of college ability before enrollment, which is consistent with the finding of Stinebrickner and Stinebrickner (2012) and Zafar (2011) whose use unique longitudinal surveys of students. At the next period after enrollment, agents learn their college ability and decide to drop out or continue until graduation. These educational decisions shape the aggregate skill in the economy and the skill premium is determined with imperfect substitution between skilled and unskilled labor. I parameterize and calibrate the model to match the US enrollment, graduation, and the skill premium given the current policy. With the model, I examine how year-dependent subsidies have differential impacts from the current constant subsidies on enrollment, graduation, and the skill premium. In addition, I examine what kind of year-dependent subsidies can maximize the expected lifetime utility of newborns with equal weights. Note that the focus of this paper is not on the total budget or the absolute size of college subsidies but on the relative sizes of subsidies across different years. Accordingly, I fix the total budget of college subsidies at the current level from now on and just change the relative sizes across years.

The main finding of this paper is the following. First, back-loaded subsidies 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 doubling the total budget of the current subsidies. This implies that just changing the structure of college subsidies with the total budget fixed has a more powerful effect than doubling the total budget. With back-loaded subsidies, subsidies are increased for latter years in college and college enrollees have incentive to stay in college and the share of college graduates increases. This increased supply of college educated labor reduces the skill premium. On the other hand, enrollment
decreases due to decreased subsidies for early years. This suggests that increasing enrollment which is
often suggested might be a misguided goal and that we should think more about how policy increases
graduation instead of enrollment. Third, back-loaded subsidies improve the social utilitarian welfare
with equal weights by more than half of a percent lifetime consumption at the steady state. The gains
come mainly from two effects. First, the back-loaded subsidies reduce the skill premium, which leads
to a decrease in the difference in wages between college graduates and college dropouts. Reducing
the difference in wages between college graduates and college dropouts can reduce the uncertainty of
wages for the people who are uncertain about college abilities. Second, the back-loaded subsidies can
reduce enrollment which was excessive in the current state due to optimism. Reducing subsidies for
early years, agents who are likely to drop out are less willing to enroll and avoid excessive enrollment.
Note that this welfare gain and the decrease in the skill premium is attainable without increasing the
budget for college subsidies and tax.

1.1 Related Literature

There is a significant amount of literature on college dropout. 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 in the models change the previous results. 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 rather
than credit constraints. Stange (2012) and Trachter (2015) quantitatively show the importance of
the option value of college dropout and that increases welfare. Hendricks and Leukhina (2017) argue
that college dropout is predictable before enrollment due to the strong correlation between high school
that college dropout increases due to a lower quality of education. This paper is based on these early
literature but introduce subsidies that can vary across college years.

In the macroeconomic literature, Bovenberg and Jacobs (2005) theoretically derive the effect of
subsidies. Abbott, Gallipoli, Meghir and Violante (2013) emphasize the effect of subsidies on parental
transfers in a quantitative overlapping generations model. Krueger and Ludwig (2016) analyze the
optimal income tax and 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. Caucutt and Kumar (2003) and Akyol and Athreya (2005) examine the normative analysis of subsidies with exogenous college dropout risk. Hanushek, Leung and Yilmaz (2014) analyze the effect of various college aid with exogenous college dropout risk. A difference between these literature and this paper is that there is college dropout as an endogenous decision of college enrollees.

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. The largest difference is that they consider only year-invariant subsidies. Chatterjee and Ionescu (2012) argue it is welfare improving to insure student loan against exogenous financial risk of dropping out with endogenous dropout decisions. This insurance is similar to front-loaded subsidies: increasing the subsidies for early years or college dropouts and reducing for later years and college graduates. They do not consider general equilibrium and imperfect substitution between skills, which is important to inequality. The wage difference between college dropouts and college graduates (the financial risk of dropping out) is determined endogenously depending on skill in the economy in this paper. I argue that year-dependent subsidies can change the supply of skill, which changes the difference in wages between college dropouts and graduates through a general equilibrium effect, which works as insurance of dropout. Considering this effect, the optimal policy is an opposite of their result and back-loaded subsidies are optimal.

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.

2 Model

The model has three main building blocks. The first is a model of college attendance featuring endogenous enrollment and graduation decisions. 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. College ability is a key factor in that it determines utility in college and the returns to education, which is consistent with Stinebrickner and Stinebrickner (2014).

The second building block is an overlapping generations life cycle with incomplete markets with
inter-generational linkage of ability and wealth. 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 ability which is inter-generationally correlated and make an endogenous transfer to their children. As Abbott et al. (2013) claims, it is important to make transfers endogenous for considering the effect of subsidies.

The third building block is a general equilibrium framework with an aggregate production function featuring imperfect substitution between skilled and unskilled labor. The educational decisions aggregate to the supply of skill in the economy, which determines the skill premium. The skill premium is a key factor of social welfare in terms of inequality. In addition, because the education sector is intensive in skilled labor, the skill premium affects tuition in equilibrium, which has a potential impact on the educational decisions.

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 overlapping generations individuals. Age is indexed by $j \in \{1, 2, \ldots, J\}$. Each individual in the economy has one offspring that lives with them before the offspring becomes independent. At the beginning of age 1, individuals become economically independent. I will identify age 1 with the graduation of high school (biological age 18), so that everyone begins their life as an independent individual as a high school graduate.

Figure 1 is the timeline. At the beginning of age 1, individuals make enrollment decisions—whether to enroll in college. Once they do not enroll in college, then 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 will observe his or her college ability and a productivity shock for working outside and then makes a decision about whether to continue in college or become a college dropout. 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 where markets for insurance and
credit are incomplete.

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 independent and 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. There are no transfers allowed at other ages.\(^2\)

Individuals survive with probability $\varphi_j \in [0, 1]$ between age $j$ and $j + 1$. Moreover, I assume $\varphi_j = 1$ for $j \in [0, j_r - 1]$. The survival rate between $j_r$ and $J - 1$ is taken from US Life Tables 2000.

### 2.2 Preferences

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

1. The expected discounted sum of instant utility:

   $$E_1 \sum_{j=1}^{J} \tilde{\beta}^{j-1} u(c_j, \ell_j)$$

   where

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

   and $c_j$ denotes consumption and $\ell_j$ is leisure at age $j$. $E_1$ is the expectation operator conditional

\(^2\) 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.
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 $\zeta$ is an adult equivalence parameter. $\beta$ is the time discount rate.

2. The expected college utility:

$$E_1 d_0(s_0) \lambda_1(\theta_c, \phi) + \beta E_1 d_1(s_1) \lambda_2(\theta_c, \phi)$$

(3)

where

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

(4)

and $d_0(s_0)$ is an indicator function which is one if the individual enrolls and $d_1(s_1)$ 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 is dependent on two components: ability $\theta$ and college taste $\phi$. $\phi$ is fixed over lifetime while the coefficient $\lambda^\phi_j$ can vary across periods (different loading). I explain ability and college taste in more detail in the individual problems section.

3. Parental altruism.

$$\beta^{j_b - 1} \nu E_1 V_0$$

(5)

where $V_0$ is the 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 $\nu$. 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:

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

(6)

$^3\bar{\beta}$ is the effective time discount rate: $\bar{\beta}_j = \beta_j \left( \prod_{k=1}^{j} \varphi_k \right)$
where aggregate labor services $H$ is a function of the inputs of two skill levels of labor: skilled labor $S$ and unskilled labor $U$.

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

(7)

where $\frac{1}{\rho}$ 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 $\delta$ the depreciation rate and hires two skills of labor at wages $w^S$ and $w^U$ respectively. I assume markets for output and inputs are competitive, so that the first order conditions for profit maximization yield:

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

(8)

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

(9)

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

(10)

Note that 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 equivalents are assumed to provide unskilled labor and college equivalents provide skilled labor. In this framework, high school equivalents are measured as the sum of high school dropout workers, high school graduate workers, and a fraction of college dropout. College equivalents are calculated as college-plus workers and a fraction of college dropout. Following this classification, I assume high school graduates provide only unskilled labor, college graduates provide only skilled labor, and college dropouts work for a $\chi$ fraction of total effective labor as skilled and $1 - \chi$ as unskilled. For convenience, I define the wage per efficiency unit for college graduates, college dropouts, and high school graduates as

$$w^{CG} = w^S$$

(11)

$$w^{CD} = \chi w^S + (1 - \chi)w^U$$

(12)

$$w^{HS} = w^U$$

(13)

Effective labor is defined as $\varepsilon_j^e(\theta, \eta)h$ where $\varepsilon_j^e(\theta, \eta)$ is labor productivity and $h$ is hours. $\varepsilon_j^e(\theta, \eta)$ is dependent on education $e$, age $j$, ability $\theta$, and idiosyncratic productivity $\eta$. The stochastic productivity shock $\eta$ is mean-reverting and follows an education-specific Markov chain $\pi^e_0(\eta'|\eta) > 0$ and
$\Pi^e_\eta$ denotes the invariant distribution function. Note that the effective labor depends on education $e$, not skilled or unskilled. Since college dropouts devote $\chi$ (resp. $1 - \chi$) fraction of effective labor as skilled (resp. unskilled), their wage per unit is $\chi w^S \varepsilon^CD_j(\theta, \eta) + (1 - \chi) w^U \varepsilon^CD_j(\theta, \eta) = w^{CD} \varepsilon^CD_j(\theta, \eta)$.

I assume that high school ability determines labor productivity of high school graduates while college ability determines labor productivity of college graduates and dropouts.

### 2.4 College

There is a representative college. To provide a student with one period of education requires $\kappa$ units of skilled labor. An interpretation of this assumption is that college enrollees obtain education from professors who are college-equivalent workers. In this formulation, education does not require any capital or unskilled labor.

The profit of college is

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

where $E$ is the measure of college enrollees and $p_e$ denotes tuition. I assume colleges are competitive and there is free entry. This implies, in equilibrium with positive units of students, $p_e = w^S \kappa$. In the United States, colleges receive subsidy from governments, which should make the sticker tuition smaller than the actual education cost for students. I reinterpret this situation as follows: colleges do not receive any subsidy while college enrollees receive subsidies instead. In both cases, enrollees pay $p_e$ less the subsidy for education.

### 2.5 Financial Markets

I assume that financial markets are incomplete. There is no insurance market against idiosyncratic risks and individuals can self-insure using only trade 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 non-arbitrage condition, the interest rate to workers is $r^- = r + \iota$. In addition, the borrowing limits for workers of education level $e$ is assumed to be $A^e$ and retired individuals have no access to loans.

---

4 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. This implies that 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.
I also assume that the cost of overseeing college enrollees is \( r + \epsilon \). With non-arbitrage condition, the interest rate to enrollees is \( r^s = r + \epsilon \). I assume these overseeing costs to match the fraction of borrowers in the economy and to imitate the actual student loan system in the United States. The borrowing limit for college enrollees is \( A^j \) at age \( j \).

2.6 Individual Problems

The lifecycle of individuals is basically composed of education, working, and retirement stages. Although college enrollees can also work in this model, 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.

\[
V_0(a, \theta_h, \eta, q, \phi) = \max[V_1^c(a, \theta_h, \eta, q, \phi), V_1(a, HS, \theta_h, \eta)]
\] (15)

An individual’s initial state is composed of initial assets \( a \), high school ability \( \theta_h \), an idiosyncratic transitory productivity \( \eta \) from \( \Pi^{HS} \), parents’ (family) income level \( q \), and unobserved education taste \( \phi \).

I assume that there are two types of ability that are distinct but related to each other: high school ability \( \theta_h \) and college ability \( \theta_c \). I assume that 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 before age 0 during high school. College ability is only observed after 1 period of college. Stinebrickner and Stinebrickner (2012) present evidence that enrollees do not have perfect foresight of their college ability before enrollment. However, college ability is correlated with high school abilities and

\[
\theta_c = \theta_h + \epsilon_c \text{ where } \epsilon_c \sim N(0, \sigma_c^2)
\] (16)

In addition, I assume that college enrollees are optimistic about their college ability, 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. They show that the expectation is higher than the actual GPA on average and that students who drop out in early years are the most optimistic and had the largest downward revision of their expectation. Given $\theta_h$, enrollees expect that

$$
\theta_c = \mu_c(\theta_h) + \theta_h + \epsilon_c \quad \text{where } \epsilon_c \sim N(0, \sigma^2_c) \quad (17)
$$

$\mu_c(\theta_h)$ is the bias between the actual mean of college ability and the one expected by them. If $\mu_c(\theta_h)$ is positive, it implies that enrollees are optimistic about their college abilities. Furthermore, the bias can depend on high school ability and I assume $\mu_c(\theta_h) = \mu_c + \mu_c \theta_h$, which implies that college enrollees with different high school ability can have different bias. I assume that the variance of the residual term is identical to the actual one.

College taste $\phi$ is observable to the individuals before enrollment but not observable to econometricians. This is necessary because people make different enrollment and dropout decisions within the same category of ability which is observable to econometrician. Within the same category of family income and high school ability, people with high college taste are more likely to enroll than those with low taste for college. Initial wealth $a$ is endogenously determined as a transfer from their parents as will be shown later. If an individual enrolls, he or she enters the first half of college and the value is $V^c_1$. If they do not enroll, they start working as high school graduates and its value is $V^1_1$.

**First half of college**

The value of being in the first half of college $V^c_1$ is

$$
V^c_1(a, \theta_h, \eta, q, \phi) = \max_{c,h,a',y} u(c, 1 - h - \bar{h}) + \mathbb{E}_{\theta_c|\theta_h} \lambda_1(\theta_c, \phi)
$$

$$
+ \beta \mathbb{E}_{\theta_c|\theta_h} \mathbb{E}_{\eta'} \max[V^c_2(a', \theta_c, \eta', q, \phi), V_2(\tilde{a}(a'), CD, \theta_c, \eta')]
$$

subject to

$$
c + a' + p_e - s_1(q) = a + y - T(c, a, y) \quad (18)
$$

$$
y = w^{HS}\varepsilon^{HS}_1(\theta_h, \eta)h, \quad a' \geq -A^c_1, \quad c \geq 0, \quad 0 \leq h \leq 1 - \bar{h} \quad (19)
$$

$$
\theta_c = \theta_h + \mu_c(\theta_h) + \epsilon_c, \quad \epsilon_c \sim N(0, \sigma^2_c) \quad \text{(perceived process)} \quad (20)
$$
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 $T(c,a,y)$ is dependent on consumption, asset, and earnings. College enrollees receive subsidies $s_j(q)$. The subsidies are need-based (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 $\theta_c$ and draw an idiosyncratic productivity as college dropout $\eta'$ from $\Pi^{CD}$. College enrollees choose whether or not to drop out of college given college ability $\theta_c$ and $\eta'$. If the individual drops out, his or her education level becomes college dropout ($e = CD$) and their value is $V_2$. Furthermore, 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\(^5\). 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

$$V_2^c(a, \theta_c, \eta, q, \phi) = \max_{c,h,a',y} u(c, 1 - h - \bar{h}) + \lambda_2(\theta_c, \phi) + \beta \mathbb{E}_{\eta'} V_3(\tilde{a}(a'), CG, \theta_c, \eta)$$  \hspace{1cm} (21)

subject to

$$c + a' + p_e - s_2(q) - y + T(c, a, y) = \begin{cases} (1 + r)a & \text{if } a \geq 0 \\ (1 + r^*)a & \text{if } a < 0 \end{cases}$$  \hspace{1cm} (22)

$$y = w^{CD} \varepsilon^{CD}(\theta_c, \eta) h, \ a' \geq -A^c_2, c \geq 0, \ 0 \leq h \leq 1 - \bar{h}$$  \hspace{1cm} (23)

They can work as college dropout during the second half of college. At the end of period, they complete college and acquire education level $e = CG$ and draw an idiosyncratic productivity as college graduates $\eta'$ from $\Pi^{CG}$. As in the case of dropout, student loan is refinanced into a single bond and the transformation is $\tilde{a}(a')$. The value of workers at age $j$ is $V_j$.

\[^5\tilde{a}(a') = a' \times \frac{r^e}{1 - (1 + r^e)^{-10}} \times \frac{1 - (1 + r^-)^{-10}}{r^-} \]
2.6.2 Working Stage

The Bellman equation for workers is:

\[
V_j(a, e, \theta, \eta) = \max_{c, h, a', y} u \left( \frac{c}{1 + I_{J_f} \zeta}, 1 - h \right) + \beta \mathbb{E}_{\eta' | \eta} V_{j+1}(a', e, \theta, \eta')
\]

subject to

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

\[
y = w^e \varepsilon_j^e(\theta, \eta)h, \quad a' \geq -\Delta^c c \geq 0, \quad 0 \leq h \leq 1
\]

where \( I_{J_f} \) 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 \( \theta = \theta_c \) for college dropout and college graduates. At each period, idiosyncratic productivity \( \eta \) transitions according to \( \pi^e_\eta \).

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:

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

subject to

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

\[
y(\theta'_h) = w^e \varepsilon_j^e(\theta, \eta)h(\theta'_h), \quad a' \geq -\Delta^c c(\theta'_h) \geq 0, \quad 0 \leq h(\theta'_h) \leq 1
\]

where

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

for all \( \theta'_h \). Before making any decisions, the individuals observe their children’s high school ability \( \theta'_h \).

The density function for the child’s ability is \( \pi_\theta(\theta'_h | \theta) \). Parents can observe neither their children’s

\[\text{After retirement, labor productivity is no longer a state variable. Thus the Bellman equation for the last period of workers is } V_{j_r-1}(a, e, \theta, \eta) = \max_{c, h, a', y} u(c, 1 - h) + \beta V_{j_r}(a', e, \theta).\]
idiosyncratic productivity $\eta'$ drawn from $\Pi^{HS}$ nor college taste $\phi$ drawn from the normal distribution $N(0, 1)$. Consumption, leisure, savings, and parental transfers can be dependent on $\theta'$. Note that the value of their children depends on family income level $q$ which reflects the potential labor income of the parental individuals.\footnote{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}

Family income determines the amount of need-based subsidies.

### 2.6.4 Retirement Stage

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

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

subject to

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

$$a' \geq 0 \quad c \geq 0$$

The sources of income are asset earnings and retirement benefits $p(e, \theta)$. In the United States, retirement benefits are determined according to the labor earnings before retirement (see Appendix B). To capture this, I assume the retirement benefits are dependent on their ability and education. The asset inflated by $\varphi_j^{-1}$ reflects that assets of expiring individuals are distributed within cohorts (perfect annuity market).

### 2.7 Government

The government collects tax revenue $T(c, a, y)$ from individuals and spends the revenues on subsidies $G_e$, other government consumption $G_c$ and retirement benefits. Government consumption $G_c$ is assumed to be exogenous and proportional to output, which follows $G_c = gY$. The total college subsidies is

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

\footnote{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}

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

where $q_1$ and $q_2$ correspond to $30,000 and $80,000.
The tax function is assumed to be

\[ T(c, a, y) = \tau_c c + \tau_k r a \mathbb{I}_{a \geq 0} + \tau_l y - \frac{dY}{N} \tag{36} \]

where the proportional consumption tax is \( \tau_c \) and the proportional capital income tax \( \tau_k \) is levied only on positive net worth. I assume the government refunds a lump-sum transfer \( dY/N \) to each individual where \( N \) is the population. This reflects the progressivity of income tax observed in the United States. \( \tau_l \) 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 \( s_j^c \in S_j^c \) be the age-specific state vector for college enrollees and \( s_j \in S_j \) for workers and retirees and \( s_0 \in S_0 \) 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 \( s_e^c \in S_e^c \). Computation is described in Appendix A.

**Definition 1** A stationary equilibrium is a list of value functions of workers and college enrollees \( \{V_j(s_j), V_j^c(s_j^c)\} \), decision rules of enrollment \( d_0(s_0) \) and graduation \( d_1(s_1^c) \), decision rules of consumption, asset holdings, labor, output, parental transfers of workers \( \{c_j(s_j), a_j^c(s_j^c), h_j(s_j), y_j(s_j), b(s_j)\} \), decision rules of college enrollees \( \{c_j^c(s_j^c), a_j^c(s_j^c), h_j^c(s_j^c), y_j^c(s_j^c)\} \), aggregate enrollees, capital, and labor inputs \( \{E, K, H^S, H^U\} \), prices \( \{r, w^S, w^U, p_e\} \), policy \( \tau_e \), measures \( \mu = \{\mu_j^c(s_j^c), \mu_j(s_j), \mu_e^c(s_j^c)\} \) such that

1. Taking prices and policies as given, value functions \( \{V_j^c(s_j^c), V_j(s_j)\} \) solve the individual Bellman equations and \( d_0(s_0), d_1(s_1^c), \{c_j(s_j), a_j^c(s_j^c), h_j(s_j), y_j(s_j), b(s_j)\}, \{c_j^c(s_j^c), a_j^c(s_j^c), h_j^c(s_j^c), y_j^c(s_j^c)\} \) are associated decision rules.

2. Taking prices and policies as given, \( K, H^S, H^CG \) solve the optimization problem of the good sector and \( E \) solves the optimization problem of the education sector.
3. The government budget is balanced.

\[ G_c + G_e + \sum_{j=1}^{J} \int_{S_j} p(e, \theta) d\mu_j = \sum_{j=1,2} \int_{S_j} T(c_j(s_j^e), a_j(s_j^e), y_j(s_j^e)) d\mu_j^e + \sum_{j} \int_{S_j} T(c_j(s_j), a_j(s_j), y_j(s_j)) d\mu_j^e \]

where

\[ G_c = gF(K,H) \] (37)

\[ G_e = \sum_{j=1,2} \int_{S_j} s_j(q) d\mu_j^e \] (38)

4. Labor, asset, and education markets clear.

\[ H^S + \kappa E = H^{CG} + \chi H^{CD} \] (39)

\[ H^U = H^{HS} + (1 - \chi)H^{CD} \] (40)

where

\[ H^{CG} = \sum_{j=3}^{J-1} \int_{S_j} e_j^{CG}(\theta, \eta) h_j(s_j) d\mu_j^{CG} \] (41)

\[ H^{CD} = \sum_{j=2}^{J-1} \int_{S_j} e_j^{CD}(\theta, \eta) h_j(s_j) d\mu_j^{CD} + \int_{S_2} e_2^{CD}(\theta, \eta) h_2^C(s_2^C) d\mu_2^C \] (42)

\[ H^{HS} = \sum_{j=1}^{J-1} \int_{S_j} e_j^{HS}(\theta, \eta) h_j(s_j) d\mu_j^{HS} + \int_{S_1} e_1^{HS}(\theta, \eta) h_1^C(s_1^C) d\mu_1^C \] (43)

and

\[ K = \sum_{j=1}^{J-1} \int_{S_j} a_j'(s_j) d\mu_j + \sum_{j=1,2} \int_{S_j} a_j^C(s_j^C) d\mu_j^C \] (44)

\[ E = \sum_{j=1,2} \int_{S_j} d\mu_j^C \] (45)

5. Measures \( \mu \) are reproduced for each period: \( \mu(S) = Q(S, \mu) \) 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 the model is parameterized and estimated. There are two sets of parameters: (1) those that are estimated outside of the model or fixed based on 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
\[
\ln \epsilon_j^e(\theta, \eta) = \ln \epsilon^e + \ln \psi_j^e + \epsilon^e \theta + \ln \eta
\]  

where \(\psi_j^e\) 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, \(\theta = \theta_h\) which is approximated by \(\ln \text{AFQT80}\). \(\eta\) is an idiosyncratic productivity shock uncorrelated with \(\theta_h\) and I can estimate the coefficient \(\theta^H_S\) using \(\ln \text{AFQT80}\). For college dropouts and college graduates, \(\theta = \theta_c\). College ability is not only a college GPA but also includes other general components represented by quality of college and college majors. Since it is hard to measure the composite, I instrument college ability using high school ability. From \(\theta_c = \theta_h + \epsilon_c\), the log labor productivity is

\[
\ln \epsilon^e + \ln \psi_j^e + \epsilon^e \theta_c + \ln \eta = \ln \epsilon^e + \ln \psi_j^e + \epsilon^e \theta_h + (\ln \eta + \epsilon^e \epsilon_c)\]

Since \(\theta_h\) is uncorrelated with \(\ln \eta + \epsilon^e \epsilon_c\), I can estimate the coefficient \(\epsilon^e_h\) using \(\ln \text{AFQT80}\) for college dropouts and college graduates in the same way as high school graduates\(^8\). Table 2 shows the coefficient on ability for each education level. As in the literature, the slope of ability is higher as the education level is higher and returns to education vary across ability.

I assume \(\pi^e(\eta'|\eta)\) is a Markov chain with two states \(\eta_H\) and \(\eta_L\), which is specific to each education

\(^8\)Since students with high \(\epsilon_c\) are self-selected as college graduates or college dropouts, I estimate using the Heckman two step estimators.

<table>
<thead>
<tr>
<th></th>
<th>HS</th>
<th>CD</th>
<th>CG</th>
</tr>
</thead>
<tbody>
<tr>
<td>log AFQT</td>
<td>.6021263</td>
<td>.7358609</td>
<td>1.306158</td>
</tr>
<tr>
<td></td>
<td>(.3102392)</td>
<td>(.3144641)</td>
<td>(.2388868)</td>
</tr>
</tbody>
</table>

Table 2: Estimated ability slope \(\epsilon^e\)
Table 3: Estimated parameters of wage process

<table>
<thead>
<tr>
<th></th>
<th>HS</th>
<th>CD</th>
<th>CG</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho^e$</td>
<td>0.9390</td>
<td>0.9545</td>
<td>0.9479</td>
</tr>
<tr>
<td>$\sigma^2_{\eta}$</td>
<td>0.0166</td>
<td>0.0208</td>
<td>0.0248</td>
</tr>
</tbody>
</table>

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

$$\ln \eta' = \rho^e \ln \eta + \epsilon^e_{\eta}, \quad \epsilon^e_{\eta} \sim N(0, \sigma^2_{\eta})$$ (48)

In addition, I assume the initial $\eta$ for each worker is drawn from the invariant distribution $\Pi^e$. After filtering out age effects, I include a fixed effect and a measurement error and employ a Minimum Distance Estimator, which 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 estimates of the parameters.

### 3.2 Intergenerational Ability Transmission

Newborns draw their high school ability $\theta^e_h$ from a normal distribution whose mean depends on the ability of their parents.

$$\theta^e_h = m^h + m^\theta \theta + \epsilon^\theta, \quad \epsilon^\theta \sim N(0, \sigma^2_{\theta})$$ (49)

The high school ability is formed partly as a result of genetics, which leads to a correlation between parents’ and childrens’ ability. Daruich (2017) argues that educational investment by their parents before high school is correlated with the parents’ earnings and education levels, which are based on parent’s ability. In addition, as Cunha and Heckman (2007), Cunha (2013), and Daruich (2017) argue, educational investment by parents has a significant effect on ability. It follows that parents’ ability also affects children’s 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 parameters $0.46^9$. A standard deviation increase in parent’s ability leads to an increase in children’s high school ability by .46 of a standard deviation.

---

*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.*
### Table 4: subsidies and family income

<table>
<thead>
<tr>
<th>$q$</th>
<th>family income</th>
<th>subsidies to students</th>
<th>subsidies to colleges</th>
<th>total $\bar{s}(q)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$-30,000$</td>
<td>$2,820$</td>
<td>$10,477$</td>
<td>$13,297$</td>
</tr>
<tr>
<td>2</td>
<td>$30,000 - 80,000$</td>
<td>$668$</td>
<td>$10,477$</td>
<td>$11,145$</td>
</tr>
<tr>
<td>3</td>
<td>$80,000 -$</td>
<td>$143$</td>
<td>$10,477$</td>
<td>$10,620$</td>
</tr>
</tbody>
</table>

### 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. (2013)), I assume subsidies are not merit-based in the status-quo.

I adopt Abbott et al. (2013) for the cost of college for enrollees and the subsidy system of the United 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 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% ($= \iota^*, \text{ annual}$). 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 Share of Skilled Labor from College Dropouts

The share of skilled labor among college dropouts $\chi$ is one of the key parameters for the effect of year-dependent subsidies. If $\chi = 0$ and all the labor provided by college dropouts is unskilled labor, encouraging enrollees to graduate increases the fraction of skilled labor in the economy and has a large effect on the skill premium. If $\chi = 1$ and all the labor provided by college dropouts is skilled labor, it is just a redistribution within skilled labor and year-dependent subsidies will have a minor effect on the skill premium.

Moreover, it is hard to calibrate $\chi$ using the skill premium of college dropouts. The fact that college dropouts earn more than high school graduates does not determine whether the labor of college dropouts is unskilled but they are more productive than high school graduates (high $\epsilon^{CD}$) or because their labor is skilled whose wage is higher than that of unskilled labor by high school graduates (high $\chi$). I need to calibrate $\chi$ using another source of data and calibrate $\epsilon^{CD}$ to match the skill premium later.

I rely on two separate sources of data. First, Torpey and Watson (2013) present the proportion of jobs in the United States that were in occupations that typically require each education level. They show that 23%, 11%, 39%, 27% of jobs require college graduation, some college, high school graduation, less than high school respectively, which is shown in the top bar of Figure 2. I interpret jobs for some college and more as skilled labor and others as unskilled labor. Then the share of skilled labor is 34% and the share of unskilled labor is 66% in the United States. The second data I use is the shares of workers of each education level at age 25 are 28%, 39%, 24%, 9% for each according to NLSY97, which is shown in the bottom bar of Figure 2. I assume college graduates provide only skilled labor and that high school graduates and high school dropouts provide only unskilled labor. Then the share of jobs of skilled labor left for college dropouts is $6\% = (34\% - 28\%)$ of all the jobs in the economy and those of unskilled labor for college dropouts is $33\% = (39\% + 27\% - 24\% - 9\%)$. This implies that the share of skilled labor by college dropouts is $\chi = 15\% = 6\%/(33\% + 6\%)$.

10 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.
Figure 2: The share of jobs and workers of each education level

Table 5: Parameters determined outside the model.

3.5 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 tertiary 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)).

3.6 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 \( (\mu_{c0}, \mu_{c1}) \), college utility \( (\lambda^0, \lambda^1, \lambda_{11}, \lambda_{22}) \), the variance of college ability
σ_c, productivity of labor \( (a^S, \epsilon^{CD}) \), education cost \( \kappa \), utility parameters \( (\mu, \beta, \nu) \), lump-sum transfer \( d \), overseeing cost \( \iota \), 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. 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 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. Note that the calibrated value of \( \mu_c \) is positive. This implies that enrollees are optimistic about their college ability on average. Since the standard deviation of college ability is 0.36\( ^{12} \), the bias for the mean ability is 48% 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). \( \lambda^0 \) is negative and agents derive disutility from college. A positive \( \lambda^1 \) 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 3 and 4. 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.

### 3.7 Validation Exercises

**Partial Equilibrium Effect of Year-Invariant subsidies**

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

\(^{11}\)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.

\(^{12}\)The square root of the sum of the variance of high school ability and \( \sigma^2_c \). \( 0.36 = \sqrt{0.215^2 + 0.287^2} \)
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_0^c$</td>
<td>college ability bias intercept</td>
<td>0.173</td>
</tr>
<tr>
<td>$\mu_1^c$</td>
<td>college ability bias slope</td>
<td>-0.311</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>college utility intercept</td>
<td>-12.3</td>
</tr>
<tr>
<td>$\lambda^\theta$</td>
<td>college utility slope</td>
<td>218</td>
</tr>
<tr>
<td>$\lambda_1^\phi$</td>
<td>first period college taste</td>
<td>48.2</td>
</tr>
<tr>
<td>$\lambda_2^\phi$</td>
<td>second half college taste</td>
<td>63.0</td>
</tr>
<tr>
<td>$\sigma_c$</td>
<td>s.d. of college ability</td>
<td>0.287</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>education cost</td>
<td>0.190</td>
</tr>
<tr>
<td>$\mu$</td>
<td>consumption share of preference</td>
<td>0.412</td>
</tr>
<tr>
<td>$\beta$</td>
<td>time discount rate</td>
<td>0.941</td>
</tr>
<tr>
<td>$\nu$</td>
<td>altruism</td>
<td>0.103</td>
</tr>
<tr>
<td>$d$</td>
<td>lump-sum transfer ratio</td>
<td>0.130</td>
</tr>
<tr>
<td>$\iota$</td>
<td>borrowing wedge ($r^- = r + \iota$)</td>
<td>17.8%</td>
</tr>
<tr>
<td>$m$</td>
<td>intergenerational ability transmission intercept</td>
<td>-0.0338</td>
</tr>
<tr>
<td>$\sigma_h$</td>
<td>intergenerational ability transmission s.d.</td>
<td>0.157</td>
</tr>
</tbody>
</table>

Table 6: Parameters calibrated.

<table>
<thead>
<tr>
<th>Moment</th>
<th>Model</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enrollment rate of ability quartile</td>
<td>(figure)</td>
<td>(figure)</td>
</tr>
<tr>
<td>Graduation rate of ability quartile</td>
<td>(figure)</td>
<td>(figure)</td>
</tr>
<tr>
<td>Enrollment rate of family income quartile</td>
<td>(figure)</td>
<td>(figure)</td>
</tr>
<tr>
<td>Graduation rate of family income quartile</td>
<td>(figure)</td>
<td>(figure)</td>
</tr>
<tr>
<td>Skill premium for CG</td>
<td>88.7%</td>
<td>89%</td>
</tr>
<tr>
<td>Skill premium for CD</td>
<td>19.7%</td>
<td>20%</td>
</tr>
<tr>
<td>Expected/Actual graduation rate $-1$</td>
<td>0.440</td>
<td>0.433</td>
</tr>
<tr>
<td>Education cost/mean income at 48</td>
<td>0.311</td>
<td>0.33</td>
</tr>
<tr>
<td>Hours of work</td>
<td>33.0%</td>
<td>33.3%</td>
</tr>
<tr>
<td>$K/Y$</td>
<td>1.318</td>
<td>1.325</td>
</tr>
<tr>
<td>Transfer/mean income at 48</td>
<td>66.8%</td>
<td>66%</td>
</tr>
<tr>
<td>log pre-tax/post-tax income</td>
<td>59.3%</td>
<td>61%</td>
</tr>
<tr>
<td>Borrowers</td>
<td>6.89%</td>
<td>6.3%</td>
</tr>
<tr>
<td>Mean of AFQT</td>
<td>-0.0603</td>
<td>0</td>
</tr>
<tr>
<td>Standard deviation of AFQT</td>
<td>0.217</td>
<td>0.21</td>
</tr>
</tbody>
</table>

Table 7: Moments matched.

*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 3: Model fit: enrollment and graduation rate for each ability quartile.

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 status-quo level and additional subsidies are given to only one generation.

The aggregate enrollment rate of the affected generation increases by 0.84 percentage points in the simulation. The micro-empirical literature has estimates of the effect of subsidies on enrollment by Dynarski (2002), Kane (1994), and Cameron and Heckman (2001). While this literature argues that the enrollment rate of groups 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.

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 fractions 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 fraction of college graduates increases by 0.35 percentage points and that of college dropouts increases by 0.49 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 complete. This is consistent with Dynarski (2008), Castleman and Long (2016),
<table>
<thead>
<tr>
<th></th>
<th>1980</th>
<th>2000</th>
<th>change (model)</th>
<th>change (data)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of college graduates</td>
<td>27.5%</td>
<td>32.2%</td>
<td>4.7pp</td>
<td>9.81pp</td>
</tr>
<tr>
<td>Share of college dropouts</td>
<td>43.6%</td>
<td>42.1%</td>
<td>-1.5pp</td>
<td>-0.14pp</td>
</tr>
<tr>
<td>College graduate premium</td>
<td>45.9%</td>
<td>88.7%</td>
<td>42.8pp</td>
<td>43.2pp</td>
</tr>
<tr>
<td>College dropout premium</td>
<td>12.4%</td>
<td>19.7%</td>
<td>7.3pp</td>
<td>7.4pp</td>
</tr>
</tbody>
</table>

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

Scott-Clayton (2011), 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 $\epsilon_{CD}$ change in the model between 1980 and 2000. In particular, I set the values of $a_S$ and $\epsilon_{CD}$ to match the college graduate wage premium 45.9% 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.” While I target the change 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.

Table 8 shows 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 43.2 percentage points from 1980 to 2000, the third column shows the share of college graduates increases by 5.44 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

In this section, I examine the effect of year-dependent subsidies. The analysis is composed of three exercises. In the first exercise, I increase overall spending without changing the structure of subsidies,
financed by higher taxes on labor income, and examine how it affects enrollment, graduation, and the skill premium. In the second exercise, I keep total spending fixed but choose subsidies by year to maximize the number of college graduates in steady state. In the third exercise, I keep total spending fixed and choose subsidies to maximize the expected lifetime value of new borns in steady state.

4.1 The Effect of Year-Invariant Subsidies on Education and the Skill Premium

As a benchmark case, I examine the general equilibrium effect of a permanent change in year-invariant subsidies on educational choice and the skill premium at the new stationary equilibrium. All the prices and the distribution are permanently changed. In particular, I examine the effect of an increase in the total government expenditure. First, I define the status-quo government total subsidies expenditure as $\bar{G}_e$. I show how the enrollment and graduation rates change as the government subsidies expenditure $G_e$ changes exogenously as follows: $G_e = 0.75 \bar{G}_e, 1.5 \bar{G}_e, 2 \bar{G}_e$. $\tau_l$ is adjusted to the changes in the subsidies expenditure. The subsidies across college years, family income, and ability proportionally change with $G_e$ fixed.

The first and second rows of Table 9 display enrollment and graduation rates for each expenditure level. While the enrollment rate increases as expenditure increases, the graduation rate is almost flat. The third and fourth rows show the fraction of college graduates and college dropouts in the economy. In the general equilibrium analysis, the skill premiums also change. As expenditure increases, college graduates and college dropouts increase and the supply of skilled labor increases. As a result, the skill premiums of college graduates and college dropouts decrease. By doubling expenditure from the current level, college graduates increase by 1.9 percentage points and the skill premium of college graduates decreases by 12.1 percentage points. College dropouts increase by 2.3 percentage points and the skill premium of college dropouts decreases by 1.7 percentage points.

<table>
<thead>
<tr>
<th>$G_e$</th>
<th>0.75 $G_e$</th>
<th>$G_e$</th>
<th>1.5$G_e$</th>
<th>2$G_e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enrollment rate</td>
<td>72.8%</td>
<td>74.3%</td>
<td>76.8%</td>
<td>78.2%</td>
</tr>
<tr>
<td>Graduation rate</td>
<td>43.4%</td>
<td>43.4%</td>
<td>43.2%</td>
<td>43.8%</td>
</tr>
<tr>
<td>Share of college graduates</td>
<td>31.6%</td>
<td>32.2%</td>
<td>33.2%</td>
<td>34.2%</td>
</tr>
<tr>
<td>Share of college dropouts</td>
<td>41.3%</td>
<td>42.1%</td>
<td>43.6%</td>
<td>44.0%</td>
</tr>
<tr>
<td>College graduate premium</td>
<td>93.1%</td>
<td>88.6%</td>
<td>82.6%</td>
<td>76.6%</td>
</tr>
<tr>
<td>College dropout premium</td>
<td>20.4%</td>
<td>19.7%</td>
<td>18.9%</td>
<td>18.1%</td>
</tr>
</tbody>
</table>

Table 9: The elasticity of education: year-invariant subsidies
4.2 The Effect of Year-Dependent Subsidies on Education and the Skill Premium

In this subsection, I find the year-dependent subsidies that maximize the number of college graduates. Note that 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

$$\max_{g_1, g_2, \tau_e} \int_{S_G^C} d\mu^C_G$$

subject to

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

and the government budget constraint. The new subsidies are \( s_1(q) = g_1 \bar{s}(q) \) and \( s_2(q) = g_2 \bar{s}(q) \) where \( \bar{s}(q) \) is the current subsidies in the status-quo (note that subsidies are independent of years in college in the status-quo). If I increase \( g_1 \) (subsidies for the first half of college), \( g_2 \) has to decrease and subsidies for the second half adjusts. Since the composition of education changes, the aggregate labor income changes and \( \tau_e \) needs to be adjusted to balance the government budget even though the subsidies spending are fixed.

The space of subsidies systems 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 \( q = 2 \) are fixed at the status-quo level. By restricting the subsidies this way, I want to focus on how year-dependent subsidies affect educational choices and the skill premiums independently from other changes such as the relative subsidies across family income levels.

Table 10 shows the implied amount of year-invariant and year-dependent subsidies. The first column is identical to the case of year-invariant subsidies with \( G_e = \bar{G}_e \) as shown in Table 9. The first three rows are subsidies of the first half across family income level \( q = 1 \) to 3 from the top to the bottom. The next three rows are subsidies of the second half in the same way. In the year-dependent case, subsidies are more generous for the second half than for the first half.

The rows 1 and 2 of Table 11 display the enrollment and graduation rates for each case. Year-dependent subsidies reduce the enrollment rate by 7.3 percentage points and increase the graduation rate by 9.6 percentage points. More people who expect to drop out refrain from enrolling due to less subsidies for the first half and generous subsidies for the second half induce more people to complete.
<table>
<thead>
<tr>
<th>( s_j(q) )</th>
<th>year-invariant</th>
<th>year-dependent</th>
</tr>
</thead>
<tbody>
<tr>
<td>( s_1(1) )</td>
<td>$13,599</td>
<td>$135</td>
</tr>
<tr>
<td>( s_1(2) )</td>
<td>$11,447</td>
<td>$114</td>
</tr>
<tr>
<td>( s_1(3) )</td>
<td>$10,922</td>
<td>$109</td>
</tr>
<tr>
<td>( s_2(1) )</td>
<td>$13,599</td>
<td>$41,509</td>
</tr>
<tr>
<td>( s_2(2) )</td>
<td>$11,447</td>
<td>$34,940</td>
</tr>
<tr>
<td>( s_2(3) )</td>
<td>$10,922</td>
<td>$33,338</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>year-invariant ( \bar{G}_e )</th>
<th>year-invariant ( 2\bar{G}_e )</th>
<th>year-dependent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enrollment rate</td>
<td>74.3%</td>
<td>78.2%</td>
<td>67.3%</td>
</tr>
<tr>
<td>Graduation rate</td>
<td>43.4%</td>
<td>43.8%</td>
<td>52.0%</td>
</tr>
<tr>
<td>Share of college graduates</td>
<td>32.2%</td>
<td>34.2%</td>
<td>35.0%</td>
</tr>
<tr>
<td>Share of college dropouts</td>
<td>42.1%</td>
<td>44.0%</td>
<td>32.3%</td>
</tr>
<tr>
<td>College graduate premium</td>
<td>88.6%</td>
<td>76.6%</td>
<td>77.8%</td>
</tr>
<tr>
<td>College dropout premium</td>
<td>19.7%</td>
<td>18.1%</td>
<td>22.4%</td>
</tr>
</tbody>
</table>

Table 11: The elasticity of education: optimal mix

As row 3 and 4 show, with year-dependent subsidies, the fraction of college graduates increases by 3.1 percentage points and that of college dropouts decreases by 9.9 percentage points. The skill premium for college graduates is 75.2%, which is less than the premium attained by the case of \( G_e = 2\bar{G}_e \) in the year-invariant case. With the expenditure fixed, switching from year-invariant to year-dependent subsidies is as effective in increasing college graduates as doubling the total spending of year-invariant subsidies.

The mechanism of the effect of year-dependent subsidies is the following. In the current system, there are lots of people who enroll in college who expect that they will not graduate. Increasing enrollment will basically encourage even 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.
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 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

\[
\begin{align*}
\max_{g_1, g_2} & \int_{S} V^{sp}_0(a, \theta, \eta, q, \phi)d\mu_0 \\
\text{subject to} & \int_{S_{c_1}} g_1 s^*(q) d\mu_{c_1} + \int_{S_{c_2}} g_2 s^*(q) d\mu_{c_2} = G_e
\end{align*}
\]  

(52)

(53)

and the government budget constraint. This problem maximizes the sum of the value of newborns with an equal weight. Note that $V^{sp}_0$ is the value function of newborns with no bias ($\mu_c(\theta_h) = 0$ for all $\theta_h$). This assumption implies that the social planner implementing the optimal policy evaluates the expected lifetime value with rational expectation, which is different from the value agents expect.

Table 12 displays the optimal college subsidies. The optimal subsidy is back-loaded and the amount for the second half is 3.8 times the subsidy for the first half. As in the previous section, the first and second rows of Table 13 show that the enrollment rate decreases by 2.9 percentage points and the graduation rate increases by 4.7 percentage points by switching to the optimal policy. The skill premium of college graduates decreases by 9.0 percentage points and that of college dropouts increases by 0.5 percentage points.

To examine the welfare effect of the optimal policy, I use lifetime consumption equivalence as a summary measure of welfare. Let $\hat{V}^{sp}_0(c, h; s_0)$ be expected lifetime utility at age 0 with the path of consumption $c$, leisure $h$ with the initial state $s_0$ with no optimism. Then lifetime consumption...
Table 13: The effect of the optimal year-dependent subsidies.

<table>
<thead>
<tr>
<th></th>
<th>Status-quo</th>
<th>Optimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enrollment rate</td>
<td>74.3%</td>
<td>72.2%</td>
</tr>
<tr>
<td>Graduation rate</td>
<td>43.4%</td>
<td>46.7%</td>
</tr>
<tr>
<td>Share of college graduates</td>
<td>32.2%</td>
<td>33.7%</td>
</tr>
<tr>
<td>Share of college dropouts</td>
<td>42.1%</td>
<td>38.4%</td>
</tr>
<tr>
<td>College graduate premium</td>
<td>88.6%</td>
<td>82.0%</td>
</tr>
<tr>
<td>College dropout premium</td>
<td>19.7%</td>
<td>20.3%</td>
</tr>
</tbody>
</table>

Table 14: Welfare decomposition.

<table>
<thead>
<tr>
<th>Total</th>
<th>Level</th>
<th>Uncertainty</th>
<th>Inequality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal</td>
<td>+0.17%</td>
<td>+0.35%</td>
<td>+0.03%</td>
</tr>
</tbody>
</table>

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. I follow Abbott et al. (2013) in detail.

The welfare gain is decomposed in Table 14. First, the total welfare gain for newborns is 0.27%. The level effect is 0.31% and there is an efficiency gain while output, capital, and consumption decrease (Table 15). In the status-quo, individuals are 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, relatively low ability enrollees stop enrolling, which reduces college disutility of low ability enrollees.

The uncertainty effect is 0.16% 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 -0.27% 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 of college graduates decreases by 10 percentage points and the standard deviations of variables decrease. This
result comes from the fact that year-dependent subsidies give less subsidies to people who drop out and are more likely to be from poor families than college graduates, which increases inequality. Although inequality as of period 1 increases, cross-section inequality in the economy decreases under the optimal policy. The standard deviations of consumption, asset, hours, and wages per hour decrease.

The welfare gain for each ability and family income level is in Table 16. 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. 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 corre-

\[\begin{array}{|c|c|c|}
\hline
& \text{Status-quo} & \text{Optimal} \\
\hline
Y & 0.291 & 0.289 \\
K & 0.383 & 0.380 \\
C & 0.193 & 0.192 \\
w^{CG} & 0.375 & 0.371 \\
w^{CD} & 0.388 & 0.391 \\
w^{HS} & 0.390 & 0.394 \\
std c & 0.117 & 0.116 \\
std a & 0.441 & 0.436 \\
std h & 0.0848 & 0.0844 \\
\text{std wage} & 0.504 & 0.497 \\
\hline
\end{array}\]

Table 15: The effect of the optimal year-dependent subsidies.

\[\begin{array}{|c|c|c|}
\hline
q & q = 1 & q = 2 & q = 3 \\
\hline
\theta = 1 & +1.7\% & +1.0\% & +1.1\% \\
\theta = 2 & +0.1\% & -0.1\% & +0.5\% \\
\theta = 3 & -1.3\% & -0.6\% & +0.1\% \\
\theta = 4 & -2.3\% & -0.6\% & -0.2\% \\
\hline
\end{array}\]

Table 16: Lifetime consumption equivalence variation for newborns.

\[\text{Note that the distribution of ability is different between the status-quo and the optimal case because the share of college graduates changes the mean ability of the future generation. Each ability quartile on the table is the quartile of the status-quo.}\]

\[\text{While the welfare loss of agents from poor family (q = 1) is large, the fraction of the poor family is only 6\% and the contribution to the social welfare is small.}\]
<table>
<thead>
<tr>
<th></th>
<th>Change</th>
<th>% of subsidy loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subsidies</td>
<td>-1.75</td>
<td>-100%</td>
</tr>
<tr>
<td>Labor income (Price of hours)</td>
<td>+0.73</td>
<td>+42%</td>
</tr>
<tr>
<td>Leisure</td>
<td>+0.27</td>
<td></td>
</tr>
<tr>
<td>Transfer from parents</td>
<td>+0.50</td>
<td>+29%</td>
</tr>
<tr>
<td>-Savings</td>
<td>+0.15</td>
<td>+8%</td>
</tr>
<tr>
<td>-Tuition</td>
<td>+0.08</td>
<td>+5%</td>
</tr>
<tr>
<td>Consumption</td>
<td>-0.30</td>
<td>-17%</td>
</tr>
</tbody>
</table>

Table 17: Change in each item of the income.

ation 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.

Second, agents change their behaviors under the optimal policy. Table [17] shows the average change in each part of the earnings and consumption for an individual with $\theta_h = 0$, $q = 2$, $\eta = \eta_H$, and $\phi = 0$ at the first half of college. The grants for the first half of college is decreased by -2.19 under the back-loaded college subsidies. The loss of subsidies for the first half of college is mitigated because (1) the labor income of college enrollees increases, (2) parents responding to this, transfer from parents increases and cancels out the loss of college subsidies, (3) since the college subsidies are shifted to the second half of college, savings for the second half of college are reduced, and (4) tuition of college decreases due to the lower wage of skilled labor under the optimal policy. In particular, the first two effects compensate the loss a lot. The labor income increases because the wage of unskilled labor college enrollees employ 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 $1 - \bar{h}$ and it increases by 0.35. 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 0.021 out of the unit hour endowment. In total, consumption decreases by 0.35 and the agents can mitigate the loss of the college subsidies by 84% (= 1 - 0.35/2.19). These results are consistent with the findings by Keane and Wolpin (1997) and Garriga and Keightley (2007).

5 Discussion

This chapter provides the optimal year-dependent subsidies and its effect in two different settings.
5.1 Optimal Year-Dependent Subsidies with Need-Based and Merit-Based System

Year-dependent subsidies affect inequality through the skill premium and screen enrollees by reducing subsidies for the first period. Policymakers can use need-based subsidies for equal access to education and merit-based subsidies for screening enrollees with high ability instead of year-dependent subsidies. It is not clear the extent to which year-dependent subsidies can be replaced by need-based and merit-based subsidies. It is helpful to understand what is the additional value of year-dependent subsidies when the policymaker can change the structure of need-based and merit-based subsidies at the same time as year-dependent subsidies.

First, as a benchmark case, I assume that the government can change the whole structure of subsidies but they have to be year-invariant. For example, the ratio of subsidies for \( q = 1 \) to that for \( q = 3 \) can be different from the status-quo. Moreover, subsidies can be different across different abilities given family income. In the second case, I allow subsidies to be year-dependent and compare the two solutions and examine the additional benefit of year-dependent subsidies.

In this subsection, I assume the function for subsidies as follows:

\[
\ln s_j(q, \theta) = \ln g_j + \ln s_q + s_\theta \theta
\] (55)

and that the subsidy for the first half depends on high school ability while that for the second half depends on college ability.

As a benchmark case, I assume subsidies are year-invariant and solve the following Ramsey problem:

\[
\max_{g_1, g_2, s_1, s_2, s_3, s_\theta} \int_{S_0} V^p_0(a, \theta_h, \eta, q, \phi) d\mu_0
\] (56)

subject to

\[
g_1 \int_{S_1^c} s(q, \theta_h) d\mu_1^c + g_2 \int_{S_2^c} s(q, \theta_c) d\mu_2^c = G_e
\] (57)

\[
g_1 = g_2
\] (58)

The solution to this problem is the second column of Table 18. The first three rows are subsidies for the first half of college for the mean ability and the next three for the second half. The first column is the status-quo. The second column is for the case of year-invariant subsidies. With the definition
Table 18: Optimal need-based and merit-based subsidies

<table>
<thead>
<tr>
<th></th>
<th>Status-quo</th>
<th>optimal year-invariant</th>
<th>optimal year-dependent</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1(1,0)$</td>
<td>$13,599$</td>
<td>$9$</td>
<td>$3$</td>
</tr>
<tr>
<td>$s_1(2,0)$</td>
<td>$11,447$</td>
<td>$18,921$</td>
<td>$15,073$</td>
</tr>
<tr>
<td>$s_1(3,0)$</td>
<td>$10,922$</td>
<td>$0$</td>
<td>$0$</td>
</tr>
<tr>
<td>$s_2(1,0)$</td>
<td>$13,599$</td>
<td>$9$</td>
<td>$6$</td>
</tr>
<tr>
<td>$s_2(2,0)$</td>
<td>$11,447$</td>
<td>$18,921$</td>
<td>$27,955$</td>
</tr>
<tr>
<td>$s_2(3,0)$</td>
<td>$10,922$</td>
<td>$0$</td>
<td>$0$</td>
</tr>
<tr>
<td>$s_\theta$</td>
<td>0</td>
<td>-1.44</td>
<td>-1.88</td>
</tr>
</tbody>
</table>

Table 19: The elasticity of education: optimal mix

<table>
<thead>
<tr>
<th></th>
<th>Status-quo</th>
<th>optimal year-invariant</th>
<th>year-dependent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enrollment rate</td>
<td>74.3%</td>
<td>74.1%</td>
<td>73.9%</td>
</tr>
<tr>
<td>Graduation rate</td>
<td>43.4%</td>
<td>43.9%</td>
<td>45.5%</td>
</tr>
<tr>
<td>Share of college graduates</td>
<td>32.2%</td>
<td>32.5%</td>
<td>33.6%</td>
</tr>
<tr>
<td>Share of college dropouts</td>
<td>42.1%</td>
<td>41.6%</td>
<td>40.3%</td>
</tr>
<tr>
<td>College graduate premium</td>
<td>88.6%</td>
<td>86.4%</td>
<td>80.9%</td>
</tr>
<tr>
<td>College dropout premium</td>
<td>19.7%</td>
<td>20.5%</td>
<td>20.7%</td>
</tr>
</tbody>
</table>

of the bins of family income $q$, the share of family income $q = 1$ is only 5%. So we can interpret $q = 2$ as the lower family income and $q = 3$ as higher family income. As you see from the table, the optimal policy is more need-based than the status-quo and the agents from rich families receive almost no subsidies. Moreover, the optimal subsidies are more generous for low-ability people. This is because high-ability people will enroll and graduate even without subsidies and there is not a large gain of subsidizing high-ability people.

The second column of Table 19 shows education and the skill premium in the economy. Since the optimal subsidies are targeting marginal enrollees more effectively by reducing subsidies for the others, both the enrollment and graduation rate increase and the skill premium of college graduates decreases. As the first row of Table 20 shows, there are large gains from less uncertainty and less inequality with the optimal year-invariant subsidies.

Next, I solve the same problem relaxing the condition $g_1 = g_2$—year-dependent subsidies. From the

Table 20: Welfare decomposition.

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Level</th>
<th>Uncertainty</th>
<th>Inequality</th>
</tr>
</thead>
<tbody>
<tr>
<td>optimal year-invariant</td>
<td>+1.68%</td>
<td>-0.60%</td>
<td>+1.63%</td>
<td>+0.66%</td>
</tr>
<tr>
<td>year-dependent</td>
<td>+2.01%</td>
<td>-0.62%</td>
<td>+1.90%</td>
<td>+0.41%</td>
</tr>
</tbody>
</table>
third column of Table 18, the general structure of need-based and merit-based subsidies does not change from the benchmark but subsidies for the second period is 51% higher than for the first period. As a result, the third column of Table 19 displays that the enrollment rate decreased and the graduation rate increases. The skill premium is reduced more than the case of year-invariant subsidies. As the second row of Table 20 shows, there is a 0.33% additional gain from year-dependent subsidies, which gain comes from the level effect of shutting down excessive enrollment and the uncertainty effect from a lower skill premium.

5.2 Correcting Bias

A large part of the welfare gain of year-dependent subsidies originates in reducing the excessive enrollment due to 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 22 shows the welfare gain from correcting bias with the current subsidies. As in the second column of Table 21, the enrollment rate drops significantly after correcting bias. Without optimism, enrollment is excessively small because of the severe credit limit for the first period in college. In the status-quo, 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
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 \( \theta_h \). The solution is the second column of Table 23. The optimal subsidy is greater for the first period than for the second period. As in the third column of Table 21, by subsidizing college in the first period, the policy can mitigate the severe borrowing limit for the first period of college and increase enrollment slightly. If the advantage of increasing the excessively small enrollment is higher than the disadvantage of increasing the skill premium, this policy is more beneficial than back-loaded subsidies. 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.

### 6 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-

\[\text{Note that correcting bias reduces the initial expected value agents have in mind even with the allocation fixed. However this is not the origin of the welfare loss of correcting bias. The welfare of the optimal policy is calculated by the social planner who does not have optimism even before correcting bias.}\]
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


A Computation of Equilibria

This section describes the method of computing equilibria. 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 $\mathbf{w} = \left( \frac{K}{H}, \frac{H^S}{H}, H, \tau_1, \bar{h} \right)$, 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 $j = J, \ldots, j_b + 1$. 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.\footnote{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 $[-2,2]$.} I apply the Endogenous grid method.

(b) Given an initial guess of the value function of newborns $V^0$, solve backward the individual problems from $j = j_b, \ldots, 0$ 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 \( \mu_0 \) and given decision rules, solve forward from \( \mu_0 \) to \( \mu_J \) and update \( \mu_0 \) until convergence.

5. Given the measures, derive the new aggregate variables \( K, H, H^S, \bar{h} \) and \( \tau_\ell \) from the government budget constraint and go back to step 2.

B Pension

The average lifetime income is

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

(59)

The pension formula is given by

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

(60)

where \( s_1 = 0.9, s_2 = 0.32, s_3 = 0.15, b_1 = 0.22\bar{y}, b_2 = 1.33\bar{y}, b_3 = 1.99\bar{y}, \bar{y} = \$28,793 \) 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 24. I take the average of the productivity of the
Table 24: Age profile estimates of each education.

<table>
<thead>
<tr>
<th></th>
<th>HS</th>
<th>CD</th>
<th>CG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.0530181</td>
<td>.0684129</td>
<td>.0955783</td>
</tr>
<tr>
<td></td>
<td>(.0030501)</td>
<td>(.0040353)</td>
<td>(.0036997)</td>
</tr>
<tr>
<td>Age²</td>
<td>-.0005314</td>
<td>-.0006872</td>
<td>-.0009521</td>
</tr>
<tr>
<td></td>
<td>(.0000356)</td>
<td>(.0000474)</td>
<td>(.0000429)</td>
</tr>
</tbody>
</table>

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 ρe, the variance of the residual σ²e, 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.