From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Economics

2019

Essays on Wage Inequality in Macroeconomics

Matsuda, Kazushige

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01w3763966h

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

Essays on Wage Inequality in Macroeconomics

Kazushige Matsuda

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Economics
Adviser: Richard Rogerson

June 2019

© Copyright by Kazushige Matsuda, 2019.

All Rights Reserved

Abstract

This collection of essays examines the effect of wage inequality on the aggregate economy and evaluates policies. In the first chapter, I evaluate the hypothesis that Skill Biased Technological Change (SBTC) leads to the decline in the firm entry rate in the United States. To the extent that people have to quit their jobs to start businesses, wages play the role of an opportunity cost of entrepreneurship. SBTC increases the skill premium and leads to an increase in the opportunity cost for college graduates to start businesses and to a decrease in the opportunity cost for high school graduates. The model can explain 21% of the actual decline in the entry rate. I conclude that SBTC is not an important factor in accounting for the decrease in the entry rate.

In the second chapter, I examine the effect of a new college subsidy scheme whose amount varies across years on enrollment, graduation, and the skill premium compared to the current system in which the subsidy is constant across years. I find that switching to back-loaded subsidies with the same total budget increases the number of college graduates, decreases college dropout and decreases the skill premium more than the case with increasing the total budget of the current subsidies by 50%, and are welfare improving despite the fact that enrollment decreases.

In the third chapter, I examine what is the optimal policy against the rising skill premium in a heterogeneous agent macroeconomic model in which agents make endogenous enrollment and dropout decisions. Some college enrollees in the model endogenously drop out after learning ability during college. Using this model, I derive the optimal progressive labor income tax and optimal college subsidies separately and compare the social welfare. While the effect of college subsidies is smaller than the case without learning ability, the optimal college subsidies improve social welfare more than the optimal progressive labor income tax.

Acknowledgements

I am extremely grateful to my advisor Richard Rogerson for his crucial advice, generous support, and encouragement. Moreover, I appreciate the helpful comments from Mark Aguiar, Titan Alon, Gregor Jarosch, Nobuhiro Kiyotaki, Ezra Oberfield, Esteban Rossi-Hansberg, Gianluca Violante, Arlene Wong, as well as participants at the Princeton Macroeconomics lunch seminar for helpful comments. In addition, I am grateful to my fellow graduate students here at Princeton. I would like to especially thank Damien Capelle, Federico Huneeus, So Kubota, Mathis Maehlum, and Julius Vutz whose friendship, support and advice has been invaluable.

I thank Dirk Krueger and Alexander Ludwig for helpful comments. I also want to thank my former advisor Kosuke Aoki for helpful comments and encouragement. I thank Susumu Cato, Tomohiro Hirano, Ryo Jinnai, Sagiri Kitao, Tomohide Mineyama, Kota Ogasawara for their academic and mental support in the completion of this dissertation.

Thank you to my parents Shigeo and Kayoko for their constant love and encouragement. Finally, I want to express my gratitude to my wife, Natsumi.

To my family

Contents

Abstract iii

Acknowledgements iv

1 Skill Biased Technological Change, the Skill Premium, and Entrepreneurship 1

1.1 Introduction 1

1.1.1 Earlier Literature 4

1.2 Data 5

1.3 Model 8

1.4 Equilibrium 10

1.5 Calibration and Computation 13

1.5.1 Calibration Procedure 13

1.5.2 Algorithm for finding a stationary equilibrium 16

1.6 Results 17

1.6.1 The effect of the increase in z_H on the stationary equilibrium . 17

1.6.2 The Benchmark Case 21

1.6.3 Increase in high skilled labor force 23

1.7 Conclusion 25

2 Optimal Timing of College Subsidies: Enrollment, Graduation, and the Skill Premium 26

2.1 Introduction 26
2.1.1 Related Literature 29
2.2 Model 30
2.2.1 Demography 31
2.2.2 Preferences 32
2.2.3 Goods Sector 35
2.2.4 College 36
2.2.5 Financial Markets 37
2.2.6 Individual Problems 37
2.2.7 Government 43
2.2.8 Equilibrium 44
2.3 Calibration 46
2.3.1 Labor Productivity Process 47
2.3.2 Intergenerational Ability Transmission 48
2.3.3 Subsidies and Loans 49
2.3.4 Government Policy 50
2.3.5 Production 51
2.3.6 The Remaining Parameters 51
2.3.7 Validation Exercises 54
2.4 Results 57
2.4.1 The Effect of Year-Invariant Subsidies 58
2.4.2 The Effect of Year-Dependent Subsidies 58
2.4.3 Welfare Analysis of Year-dependent subsidies 61
2.4.4 Correcting Bias 66
2.5 Conclusion 68
Appendix 2.A Computation of Stationary Equilibrium 69
Appendix 2.B Pension 70

Appendix 2.C Labor Productivity Process 70
Appendix 2.D Intergenerational Ability Transmission 72
Appendix 2.E Calibration without Optimism 73

3 Progressive Taxation versus College Subsidies with College Dropout 78

3.1 Introduction 78
3.1.1 Related Literature 80
3.2 Model 82
3.2.1 Demography 82
3.2.2 Preferences 83
3.2.3 Goods Sector 85
3.2.4 College 86
3.2.5 Financial Markets 86
3.2.6 Individual Problems 86
3.2.7 Government 91
3.2.8 Equilibrium 92
3.2.9 Calibration 94
3.3 Results 94
3.3.1 The Effect of an Increase in Labor Income Tax by 1% 95
3.3.2 Optimal Progressivity versus Optimal College Subsidies 97
3.3.3 Case with No Learning of College Ability 100
3.4 Conclusion 101
Appendix 3.A Computation of Equilibria 102
Appendix 3.B Pension 103
Appendix 3.C Calibration 103
3.C.1 Labor Productivity Process 104
3.C.2 Intergenerational Ability Transmission 105
3.C.3 Subsidies and Loans 105

3.C.4 The Remaining Parameters 106
Appendix 3.D Optimal Mix 107

Chapter 1

Skill Biased Technological Change, the Skill Premium, and Entrepreneurship

1.1 Introduction

Entrepreneurship is considered to be an essential driving force in the dynamics of economies. First, young firms are more likely to perform research and development and to have higher productivities.[1] Second, young firms also contribute to the dynamics of employment.[2] Nonetheless, as Figure 1.1 shows, the firm entry rate has been declining in the Unites States. The ratio of the employment by young firms to that by all firms, which I call employment entry rate from now on, has also been decreasing. Schoar (2010) argues that it is critical to distinguish between "opportunity" entrepreneurship and "necessity" entrepreneurship. The former is entrepreneurship by those who have opportunities for new businesses. The latter is entrepreneurship

[1] See Acemoglu et al. (2018), Huergo and Jaumandreu (2004) and Foster et al. (2008). See Pugsley and ahin (2018) about the effect of young firms on business cycle fluctuation.

[2] See Haltiwanger et al. (2013)

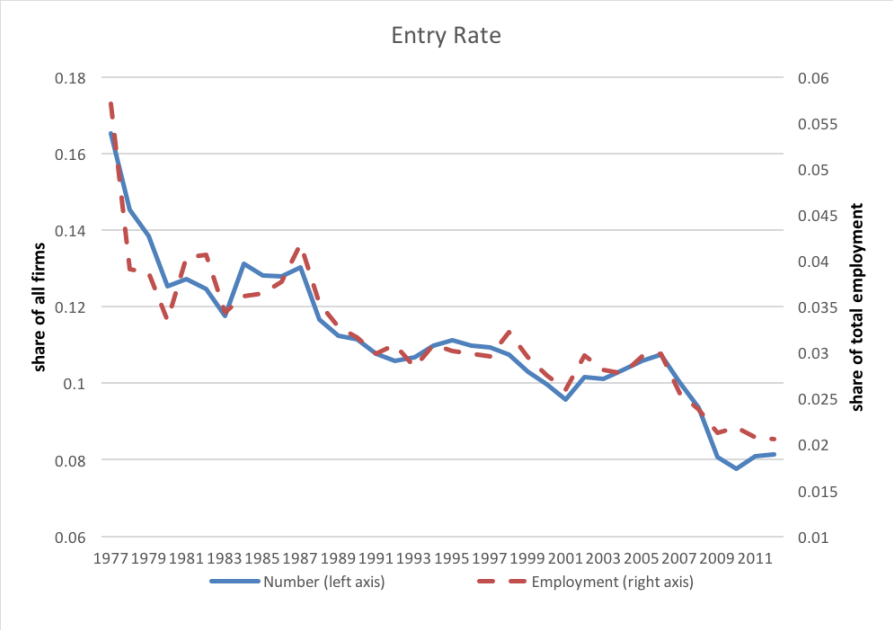


Figure 1.1: The decline in the entry rate of the United States

by those who do not have other alternatives to earn subsistence. It is difficult to separately measure the two types of entrepreneurship using data. However, Decker et al. (2014b) argue that opportunity entrepreneurship has declined since the beginning of the 1980's. In what follows, I interpret the decline in the entry rate of Figure 1.1 as a decline in opportunity entrepreneurship.

This paper examines the cause of the decline in the entry rate. In particular, I evaluate the hypothesis that Skill Biased Technological Change (SBTC) has decreased the firm entry rate in the United States. The logic of the hypothesis is the following. Suppose that entrepreneurship is more common across college graduates. To the extent that people have to quit their jobs to start businesses, their current wage plays the role of an opportunity cost that can avert people from entrepreneurship. SBTC increases the skill premium[3] and leads to an increase in the opportunity cost for college graduates to start businesses and to a decrease in the opportunity cost for high school graduates. As Figure 1.2 shows, the skill premium has been increasing since the end

[3]The skill premium is defined to be labor earnings by college graduates divided by labor earnings by high school graduates.

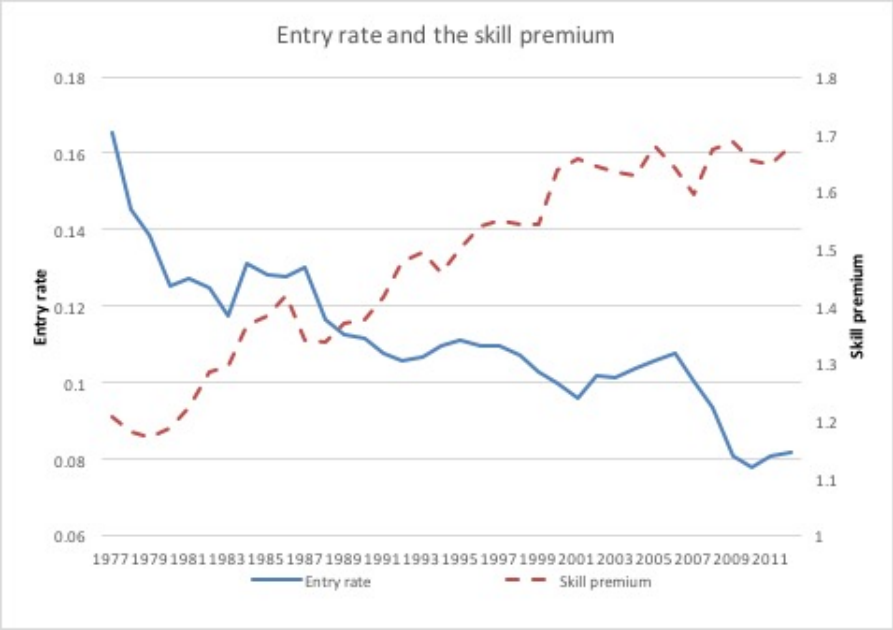


Figure 1.2: The increase in college wage premium

of the 70's.[4] If entrepreneurship is more common among college graduates, the total entry rate may decrease. I build a model integrating occupational choice with the outside option of market work in an industry dynamics framework and examine how much of the decline in the entry rate in the United States the model can account for.

My result shows that SBTC leads to a decline in the entry rate but under an unexpected mechanism. First, as expected in the hypothesis, the higher skill premium decreases the total mass of firms of college graduates and increases that of high school graduates. However, a higher skill premium also functions as a higher outside option for those operating a business and thus increases the exit of college graduates and decreases the exit of high school graduates. The effect of increasing exit of college graduates dominates that of decreasing entry and the entry rate of college graduates increases as a result. The opposite holds true for high school graduates. Since the total population of high school graduates is substantially larger than that of college graduates, the decline in the entry rate by high school graduates dominates the increase in the entry rate by college graduates and the total entry rate declines.

[4]This graph is constructed in the same manner of Katz and Murphy (1992).

SBTC in this model can account for 21% of the observed decline in the entry rate in the United States. However, the fraction of college graduates has been increasing over time in the United States. If we integrate the increase in the population of college graduates into the analysis, the net effect of SBTC is very small. I conclude that SBTC is not an important factor in accounting for the decrease in the entry rate.

This paper proceeds as follows. Section 2 discusses the dataset I use. Section 3 shows the model we employ in the following analysis. In section 4, we define an equilibrium and section 5 specifies the methodology. Section 6 shows results and section 7 concludes.

1.1.1 Earlier Literature

There is a growing body of literature with respect to firm entry. The closest predecessor of the modeling of this paper is Hopenhayn (1992), who generates firm entry and exit simultaneously in the steady state. Jovanovic (1982) argues that integrating learning into models of firm dynamics gives rise to firm entry and exit. However, the models of the earlier literature have not generated changes in the entry rate even though the decline in the entry rate in the Untied States had been observed. Hathaway and Litan (2014) and Decker et al. (2014a) empirically show the decline of business dynamism in the United States. One contribution of this paper is to examine a possible factor that changes the entry rate endogenously in a framework in which firms enter and exit simultaneously.

The hypothesis that higher opportunity cost reduces the incentive to start businesses is not brand-new as Le (1999) suggested the possibility. In addition, there are other hypotheses for the decline in the entry rate in the United States. Karahan et al. (2015) argue that the decline in the United States is caused by a decline in the growth rate of the workforce.

The skill premium is one of the key drivers of this model and is discussed in the

context of SBTC. Acemoglu and Autor (2011) survey the literature on SBTC and I simply review it here. Katz and Murphy (1992) show that the skill premium has been expanding and suggest that it had been driven by the increase in the demand for college graduates labor force that exceeded the increase in supply. Autor et al. (2003) and Krusell et al. (2000) argue that the increase in the demand for high skilled college graduates labor force is due to the advance of technology that complements high skilled work and substitutes low skilled work. This paper exploits SBTC as a key driver of the increasing skill premium and connects SBTC with firm entry.

1.2 Data

The Business Dynamics Statistics (BDS) is maintained by the U.S. census Bureau and covers the universe of employer firms in the private nonagricultural sector with at least one employee. In the dataset, a firm is defined as a composite of establishments that are under common ownership and control. BDS includes data for the age of establishments. The age of a firm is defined to be the age of the oldest establishments it owns. This feature enables us to distinguish between a new firm, a new establishment of an existing firm, and reallocation of establishments through merger and acquisition. See Foster et al. (2006) for details regarding the measurement of age. BDS also covers employment of each firm. In this paper, an entrant is defined to be a firm whose age is zero. The entry rate is defined as the number of firms of age 0 divided by the number of all firms. The employment entry rate is defined as employment of firms of age 0 divided by total employment. I focus on the span from 1978 because the observed employment entry rate in BDS is excessively high in 1977 and there might be a problem in the data because 1977 is the first year the dataset became available. In addition, to remove the effect of the financial crisis from 2007, I focus on the span from 1978 to 2006 from now on.

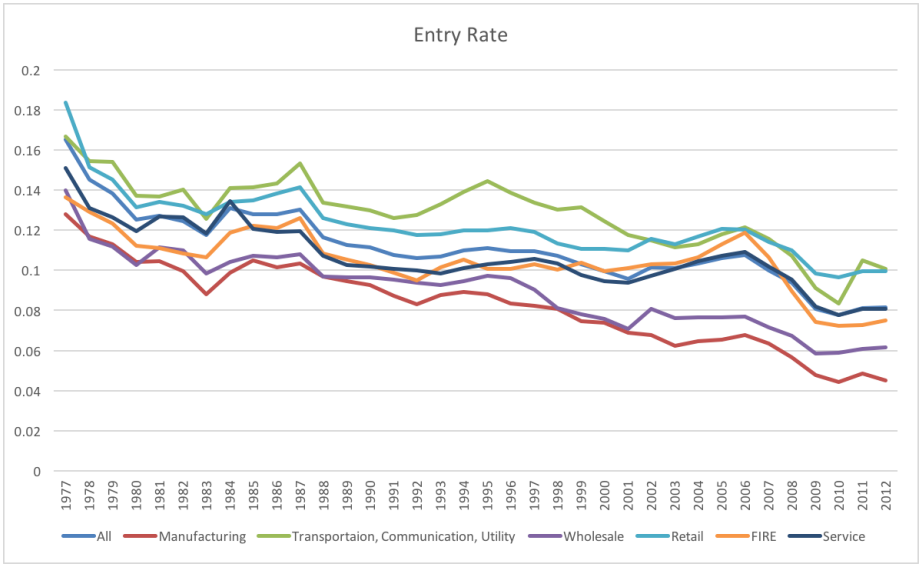


Figure 1.3: The decline in the entry rate across industries

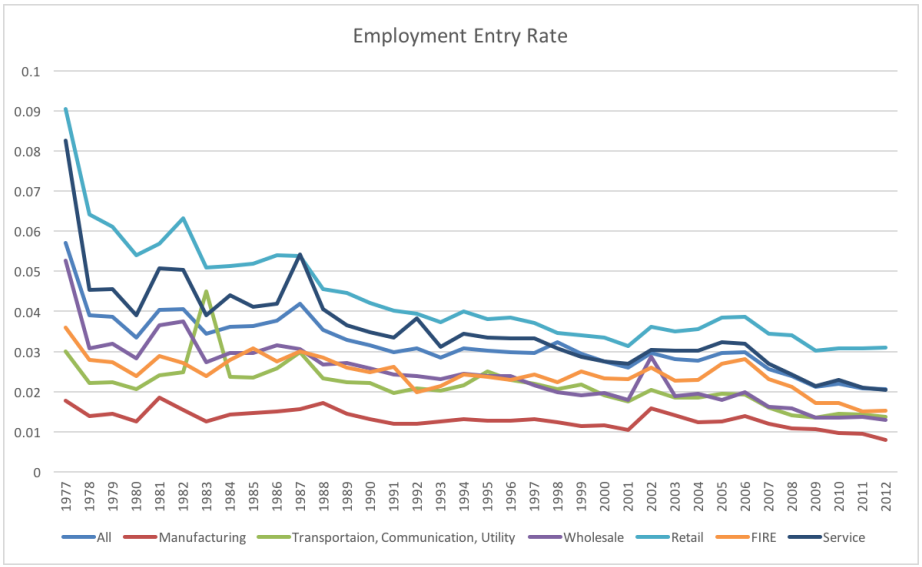


Figure 1.4: The decline in the employment entry rate across industries

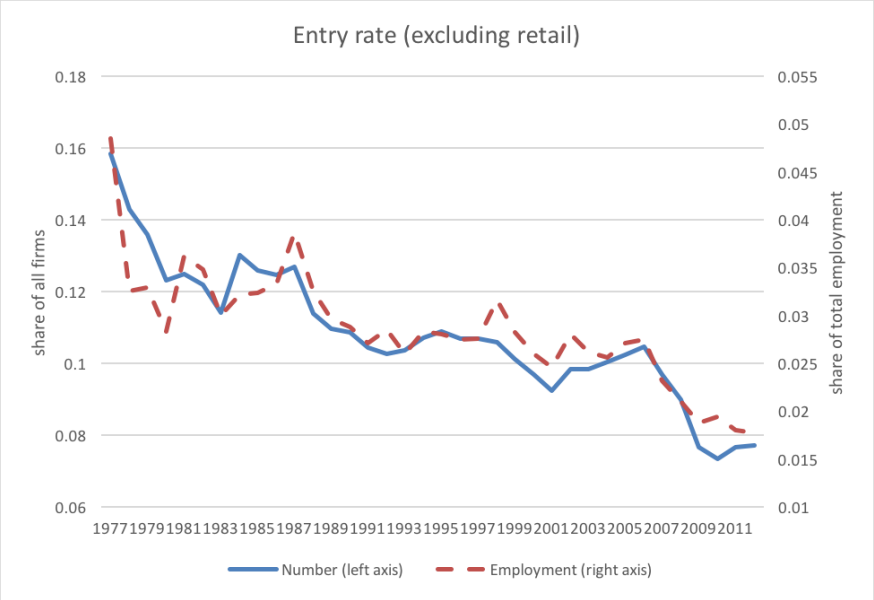


Figure 1.5: The decline in the entry rate excluding retail sector

BDS provides information by industries. Figures 1.3 and 1.4 show the entry rate and the employment entry rate respectively for each industry. Many of the industries show declines in their entry rates, so the declining total entry rate is not driven by changes in the composition from industries with high entry rates to those with low entry rates.

However, it is true that the retail sector shows a larger decline in the employment entry rate than other sectors. Foster et al. (2006) argue that the change in the business model of the retail sector due to the progress of technology such as tracking sales data and controlling inventory has replaced mom-and-pop shops with large chain stores. This argument implies that the decline in the entry rate in the retail sector is not due to higher opportunity cost for college educated entrepreneurs. Figure 1.5 displays the entry rate for non-retail sectors and shows that it has declined. The employment entry rate has also declined in the non-retail sectors, and the size of the decline in the retail sector is lower than that of the total entry rate. Thus there is a need to explain the declining entry rate among non-retail sectors.

1.3 Model

In order to examine firm entry and exit, the framework of this model is based on Hopenhayn (1992). In this paper, I focus more on the entity who actually starts a firm and examine their incentives.

Time is discrete and proceeds as $t = 0, 1, 2, \ldots$. There is mass one of identical households and each household has a continuum of agents of total mass one. Members are composed of two types that are high and low skilled, each of which is denoted as $i = H$ for high type and L for low type. The mass of agents of type i is N_i and $N_H + N_L = 1$. In the quantitative analysis, I will interpret college graduates as high skilled agents and high school graduates as low skilled agents. Each individual is either an employee or a business owner.

Firstly, I describe the preference of agents. The representative household's preferences are given by

$$\sum_{t=0}^{\infty} \beta^t u(c_t) \tag{1.1}$$

where $u(\cdot)$ is a one-period utility function, c_t is consumption at time period t, and the time discount rate is β. Households can trade claims of one unit of goods provided at the next period at interest rate r_t.

Production of goods requires low and high skilled labor. Additionally, as in the literature on the college premium, I integrate SBTC into the model. Each firm has a stochastic production function and the profit that accrues to the firm is:

$$p_t f(n_{Lt}, n_{Ht}, s_t) - w_{Lt} n_{Lt} - w_{Ht} n_{Ht} \tag{1.2}$$

where p_t is the price of goods and w_{it} is the wage for i type of labor in time period t. s_t is a stochastic idiosyncratic shock to productivity for each firm and follows a

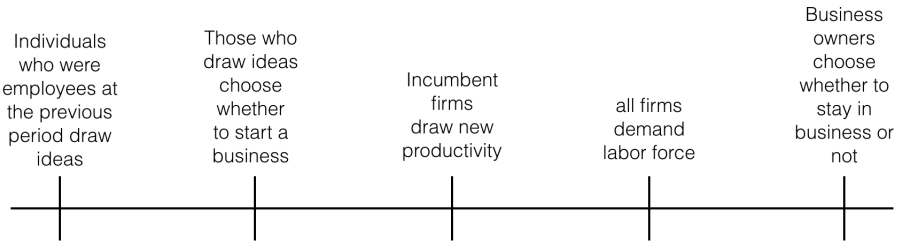


Figure 1.6: The sequence of decisions

first-order Markov process $F(s_t, s_{t+1})$, where $F(\cdot, s_t)$ is the distribution function for the next period productivity s_{t+1} given that the current productivity is s_t. n_{Lt} and n_{Ht} denote low and high skilled labor input respectively. Each firm needs to have a business owner and she can choose to exit at any period. If she chooses to stop the business, she becomes an employee.

Each individual of type i who was an employee in period $t-1$ potentially draws an entrepreneurial idea s_t at the beginning of period t with probability $\alpha_i \in [0,1]$ from a distribution $v(s_t)$. The difference between the two types of employees is not only wage rates but also the probability with which they draw business ideas. This distribution function $v(s_t)$ is assumed to be common between the high and low types. In the decision to enter, employees compare the benefit of starting the business with the value of staying as an employee.

Next, consider the timing of decisions by firms and workers. Figure 1.6 shows the sequence of decisions in period t. At the beginning of period t, each individual of type i who was an employee at the previous period draws a business idea with probability α_i and choose whether to start the business or not. If one starts the business, she enters as a business owner with the productivity of the idea she draws. Next, each incumbent firm draws a new productivity s_t from $F(s_t, s_{t-1})$. Observing s_t, each firm demands the two types of labor. At the end, incumbents choose whether to stay in business or exit. The process is repeated next period.

1.4 Equilibrium

This paper focuses on a stationary equilibrium in which prices are constant. I choose the wage of low skilled work as the numeraire so that $w_L = 1$ for all periods and abbreviate the index and use w instead of w_H. I define $\mathbf{p} = (p, w)$. In stationary equilibria, it must hold that $1 + r = 1/\beta$.

I examine the decision problems for agents and formulate the optimization problems of employees and firms in recursive form. Each employee of type i solves the following Bellman equation.

$$W_i = w_i + \beta \left(\alpha_i \int \max\{V_i(s'), W_i\} dv(s') + (1 - \alpha_i) W_i \right) \tag{1.3}$$

where $w_H = w$ and $w_L = 1$. The first term reflects current period wages of type i and the second term inside the integral compares the value of starting a business with the drawn idea which is denoted as $V_i(s')$ with that of being an employee W_i. The third term is for the case where one does not draw an idea and continues to be an employee. I denote the decision on whether to start the business as $S_i(s')$. If they start their businesses, $S_i = 1$ and if not, $S_i = 0$. In Hopenhayn (1992), there is a zero profit condition under which, comparing the benefit of entering with a fixed entry cost, firms decide whether to enter for production or not. By contrast, in this model, potential entrants compare the benefit of entering with an opportunity cost, instead of the fixed entry cost. This model generates a change in the opportunity cost of entry and it yields the change in the entry rate. As we will see later, the driving force of the change in the opportunity cost is SBTC.

The Bellman equation of a business owner of type i is:

$$V_i(s) = \max_{n_L, n_H} \{ p f(n_L, n_H, s) - n_L - w n_H + \beta \max[E_s V_i(s'), W_i] \} \tag{1.4}$$

where E_s denotes the expectation operator conditional on the current productivity s. s' denotes next period's productivity and as noted earlier is drawn from the distribution $F(s, s')$. At the end of each period, each business owner can choose to stop operating one's businesses to return to the labor market of one's type and this is reflected in the fourth term. As in the entry decision, the exit decision also takes account of the opportunity cost of staying in business, instead of the operational fixed cost in Hopenhayn (1992). SBTC also strengthens the incentive for high skilled agents to exit. The optimal choices of low and high skilled labor input are respectively denoted as $N_L(s)$ and $N_H(s)$. The exit decision is written as $X_i(s)$. If one exits, $X_i = 1$ and if not, $X_i = 0$.

The decisions described above, in combination with the dynamics of productivities, genearate an equilibrium path. Let $\mu_i(s)$ be the measure of firms of productivity s of type i in a stationary equilibrium. The mass of entrants of type i can be defined as

$$M_i = \left(N_i - \int d\mu_i(s) \right) \alpha_i \int S_i(s')v(s')ds' \tag{1.5}$$

The first term reflects the mass of employees of type i this period. Among them, the fraction α_i draw ideas. The final term reflects agents who decide to start a business given the productivity of their draw.

I write the next period measure of firms as $\mu_i'(s)$. The transition from μ_i to μ_i' is characterized as the following:

$$T_i(\mu_i)(s') = \int (1 - X_i(s))F(s, s')d\mu_i(s) + M_i \frac{S_i(s')v(s')}{\int S_i(s')v(s')ds'} \tag{1.6}$$

and $\mu_i' = T_i(\mu_i)$. The first term on the right hand side is the mass of existing firms and the second is the mass of entrants. I define the total mass of firms of productivity s as $\mu(s) = \mu_H(s) + \mu_L(s)$. The aggregate output of this economy is determined in

the following form:

$$Y = \int f(N_L(s), N_H(s), s) d\mu(s) \tag{1.7}$$

Labor demand for high and low types and profits are determined as follows.

$$L_i^d = \int N_i(s) d\mu(s) \text{ for } i = H, L \tag{1.8}$$

$$\Pi = pY - L_L^d - wL_H^d \tag{1.9}$$

Lastly, I examine the optimization problem by the representative household. In a stationary equilibrium with constant prices and the interest rate satisfying $1/(1+r) = \beta$, the optimization can be rewritten as the following static problem.

$$\max_c u(c) \tag{1.10}$$

subject to

$$pc \leq \left(N_L - \int d\mu_L(s)\right) + w\left(N_H - \int d\mu_H(s)\right) + \Pi \tag{1.11}$$

The first term on the right hand side is the labor income from low skilled employees and the second term from high skilled employees.

I define a stationary equilibrium.

Definition 1. *A stationary equilibrium consists of a list of functions* W_i, $V_i(s)$, $S_i(s')$, $N_i(s)$, $X_i(s)$, μ_i^* *for* $i = H, L$, c^* *and* $\mathbf{p}^* \geq 0$ *such that*

1. *Taking* $\mathbf{p}^*$ *as given,* c^* *is the optimal decision to the houshold's optimization problem.*

2. *For each* $i = H, L$, *taking* $\mathbf{p}^*$ *as given,* W_i *solves the Bellman equation of em-*

ployees and $S_i(s')$ is the optimal decision rule for entry.

3. *For each $i = H, L$, taking $\mathbf{p}^*$ as given, $V_i(s)$ solves the Bellman equation of business owners and $N_L(s)$, $N_H(s)$, and $X_i(s)$ are optimal decision rules.*

4. $N_i - \int d\mu_i^*(s) = \int N_i(s) d\mu^*(s)$ *for* $i = H, L$.

5. $T_i(\mu_i^*) = \mu_i^*$ *for* $i = H, L$.

Condition 4 states that the market of low and high skilled labor should clear and condition 5 states that the distribution of firms should be reproduced by the optimal decisions by agents.

1.5 Calibration and Computation

1.5.1 Calibration Procedure

In this section, I calibrate the model starting with specifying the functional forms. The production function is CES:

$$f(n_L, n_H, s) = s((z_H n_H)^{\frac{\sigma-1}{\sigma}} + (z_L n_L)^{\frac{\sigma-1}{\sigma}})^{\frac{\theta\sigma}{\sigma-1}} \tag{1.12}$$

where σ is the elasticity of substitution between low and high skilled labor force. This choice is standard in the literature on SBTC and the college premium. The Markov process $F(s, s')$ is assumed to be an AR(1) process on $\log s$:

$$\log s' = a + \rho \log s + \epsilon \tag{1.13}$$

where $\epsilon \sim N(0, \sigma_\epsilon^2)$. Later, I set a grid of s for calculation. $v(s)$ is set as the uniform distribution with the support of the range of the grid. The utility function is specified as $u(c) = \ln c$.

Parameters	Values	Source
β	0.96	Steady state interest rate = 4%
σ	1.41	Katz and Murphy (1992)
θ	0.7	Hopenhayn and Rogerson (1993)
z_L	1	Normalization
a	0	Normalization
ρ	0.97	Lee and Mukoyama (2012)
σ_ϵ	0.11	Lee and Mukoyama (2012)
α_H	0.300	Entrepreneurs among college graduates
α_L	0.136	Entrepreneurs among high school graduates
N_H	0.1751	Fraction of college graduates
N_L	0.8249	Fraction of high school graduates
z_H	0.0293	The skill premium

Table 1.1: Calibrated parameters

Calibration proceeds as follows. I interpret one period in this model as one year and the time discount rate β is set so that the annual risk free real interest rate r is 4%. I set the elasticity of substitution as $\sigma = 1.41$, which is the estimated value of the elasticity of substitution between high and low skilled labor in Katz and Murphy (1992). I use the calibrated value of θ in Hopenhayn and Rogerson (1993), which means that the labor share of production is 70% and the profit is 30% of total sales. Taking z_H/z_L as given, the price of goods, z_H, and the idiosyncratic shock enter multiplicatively in the business owner's objective function. Thus I normalize a to zero and z_L as one without loss of generality. The parameters ρ and σ_s are set as in Lee and Mukoyama (2008). They obtain the values to match the process of employment.

Using five moments, I calibrate the following parameters: α_H, α_L, N_H, N_L, and z_H. To calibrate α_H and α_L, I use data from the Global Entrepreneurship Monitor (GEM), which covers measures of entrepreneurship in several countries in a unified framework from 2001 to 2011. This dataset shows many aspects of those who start businesses. For example, it has the information on educational attainment of entrepreneurs. An advantage of this survey dataset is that it differentiates between

opportunity entrepreneurship and necessity entrepreneurship. In this model, when considering starting businesses, agents necessarily have the alternative to earn income as an employee. Thus entrepreneurship in this model is considered as opportunity entrepreneurship and I use only the data on opportunity entrepreneurs. Additionally, I'm able to obtain the share of opportunity entrepreneurs among high school graduates and college graduates from 2001 to 2011. I use the time averages of the moments in the span to calibrate α_L and α_H respectively. The obtained value of α_H is larger than that of α_L. An interpretation of this difference is that college graduates have accumulated more human capital and it helps with cognitive and creative activity, which facilitate entrepreneurship.

Next, I target the ratio of the population of college graduates to that of high school graduates of the start year 1978. I follow the methodology of Katz and Murphy (1992) except one point. The target in this model is the ratio of populations whereas they estimate the ratio of hours of work per year. In this model, agents do not supply labor while they are business owners and N_H matches with the population of college graduates. The population of low skilled agents is calibrated from $N_L = 1 - N_H$. I target the skill premium in the start year 1978 to calibrate z_H. In the model, the skill premium is determined from the relative productivity z_H/z_L and the relative labor supply of high skilled agents. Even when $z_H < z_L$, the scarcity of high skilled agents makes the skill premium larger than one.

I note here that under these calibration values the entry and exit decisions of agents have thresholds. One enters if and only if he or she draws a productivity above a threshold. This also holds true for the exit decision rule. A firm exits if and only if its productivity is below a threshold. Thus I consider the entry and exit decision rules as the thresholds.

1.5.2 Algorithm for finding a stationary equilibrium

I build a grid for $\log s$ that includes 2000 grid points which are uniformly spaced over the range of $[a/(1-\rho)-3\sigma_\epsilon/\sqrt{1-\rho^2}, a/(1-\rho)+3\sigma_\epsilon/\sqrt{1-\rho^2}]$. Using the calibrated values from the previous section, the range is $[-1.3574, 1.3574]$. Then I use Tauchen's method[5] to discretize the AR(1) of $\log s$. In addition, I discretize $v(s)$ in the following way. When $v(s)$ is uniform, the distribution function of $x = \log s$ is $g(x) = B\exp(x)$, where B is a constant to normalize the distribution function $g(x)$. Thereby I assign the value $\exp(x)$ for each grid point x and normalize it so that the sum of $g(x)$ over the range of x becomes one. The grid for s is formed transforming each value of the grid of $x = \log s$.

I describe the methodology to solve for a stationary equilibrium in three steps. Using the calibrated values and guessing $\mathbf{p}^*$, the first step finds the decision rules using conditions 2 and 3 in the definition of a stationary equilibrium. $N_L(s)$ and $N_H(s)$ are the decision rules from (1.4) and derived as

$$N_L(s) = p^* s\theta z_L^{(\sigma-1)(1-\theta)} \left(\frac{z_H^{\sigma-1}}{w^{*\sigma-1}} + z_L^{\sigma-1} \right)^{\theta\sigma/(\sigma-1)-1} \tag{1.14}$$

$$N_H(s) = \frac{(z_H/z_L)^{\sigma-1}}{w^{*\sigma}} N_L(s) \tag{1.15}$$

Then I use value function iteration to obtain W_i and $V_i(s)$ for each i. Using (1.3) and (1.4), I iterate on the value functions until they converge[6]. In deriving $S_i(s)$ ($X_i(s)$), I compare the two value functions for each grid s and assign one if the value of entry (exit) is higher and zero if not.

The second step of the algorithm determines μ_i^* for $i = H, L$ using condition 5 of the definition of an equilibrium. I build an initial guess of μ_i. With the initial guess, I iterate using $\mu_i^* = T_i(\mu_i^*)$ until μ_i converges for each i.

[5] See Tauchen (1986) for the detail.

[6] The tolerance I use for any iteration in this paper is 0.001.

The third step involves condition 4 to derive a new $\mathbf{p}^{*\prime}$. First, using the market clearing condition for low skilled labor, I derive a new $p^{*\prime}$. Specifically,

$$p^{*\prime} = \frac{N_L - \int d\mu_L^*(s)}{\int s\theta z_L^{(\sigma-1)(1-\theta)} \left(\frac{z_H^{\sigma-1}}{w^{*\sigma-1}} + z_L^{\sigma-1} \right)^{\theta\sigma/(\sigma-1)-1} d\mu^*(s)} \tag{1.16}$$

In addition, I derive a new $w^{*\prime}$ from the following market clearing condition for high skilled labor:

$$w^{*\prime} = \left\{ (z_H/z_L)^{\sigma-1} \frac{N_L - \int d\mu_L^*(s)}{N_H - \int d\mu_H^*(s)} \right\}^{1/\sigma} \tag{1.17}$$

I then reset the initial guess as $\lambda\mathbf{p}^{*\prime} + (1-\lambda)\mathbf{p}^*$[7] and repeat from the first to the third step until $\mathbf{p}^*$ converges.

1.6 Results

1.6.1 The effect of the increase in z_H on the stationary equilibrium

In this section, I examine how exogenous changes in z_H affect the properties of the stationary equilibrium. Starting from the calibrated value of z_H set to match with the skill premium in the start year 1978 in the previous section, I consider increases up to the calibrated value of z_H to match with the skill premium in the end year 2006, with the other parameters fixed. The calibrated value of z_H to match for 2006 is 0.153.

The left panel of Figure 1.7 shows that the skill premium grows from 1.18 to 1.64, which I target to calibrate z_H. It is natural that the relative enhancement of productivity on high skilled work increases the value of high skill work, which is

[7] I set $\lambda = 0.1$.

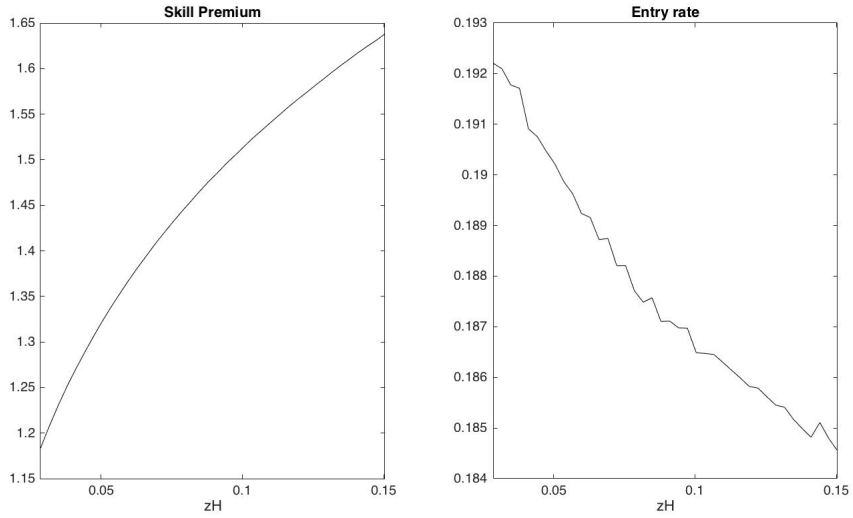


Figure 1.7: The skill premium and entry rate

consistent with the literature on the skill biased technical change. The right panel of Figure 1.7 shows that the entry rate is decreasing as z_H rises.

In order to examine the mechanism behind this result, it is helpful to decompose the entry rates into those for high skilled and low skilled agents. Figure 1.8 shows that the entry rate of high skilled agents, which is defined as the ratio of high skilled entrants to the total high skilled business owners, is increasing while the entry rate of low skilled agents, which is defined in the same way as that of high skilled agents, is decreasing. Apparently, it is inconsistent with the hypothesis that high skilled agents have a higher outside option to starting businesses and lower incentive to start businesses as the skill premium increases. As briefly mentioned in the introduction, a higher option value of being a worker has two effects on the entry rate of high skilled agents: the first one is that higher wages make it less attractive to quit one's job, which leads to a decline in entry for college graduates. Second, a higher wage also contributes to a higher outside option value against operating a business and leads to an increase in exit. Figure 1.9 shows the decline in the mass of firms started by high skilled agents and the increase in the mass of firms by low skilled agents. This

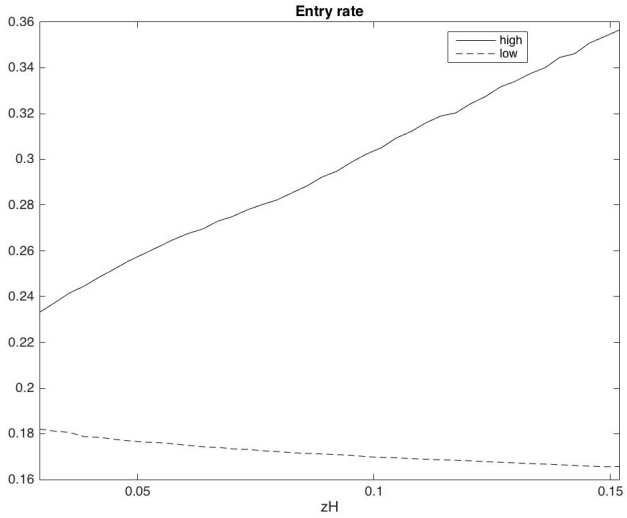


Figure 1.8: Entry rate of high and low skilled agents

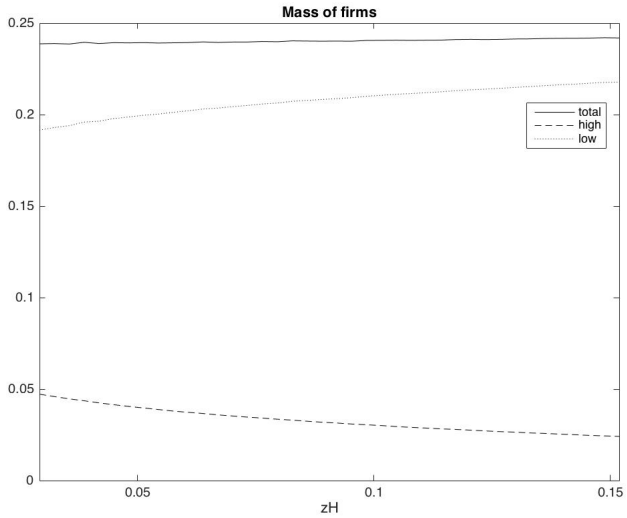


Figure 1.9: Mass of firms

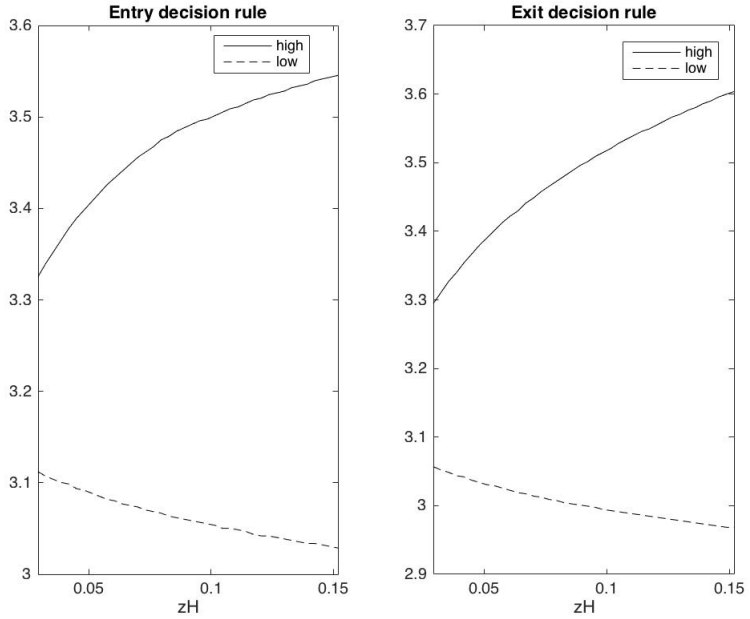


Figure 1.10: Entry and exit decision thresholds

implies that the decrease in entry by high skilled agents is dominated by the increase in exit by them. In contrast, the mass of firms by low skilled agents is increasing. The lower amount of exit dominates the increase in entry by low skilled agents and the entry rate for them is decreasing. The population of low skilled agents is significantly higher than of high skilled agents with the calibrated parameter values. Since the increase in the entry rate of high skilled agents is dominated by the decrease in the entry rate of low skilled agents, the total entry rate is decreasing as shown in the right panel of Figure 1.7.

The left panel of Figure 1.10 shows that the entry threshold is increasing for high skilled agents and decreasing for low skilled agents. This shows the first effect that a higher skill premium increases the outside option of entering for high skilled agents and decreases for low skilled agents. In contrast, the right panel of Figure 1.10 shows that the exit thresholds move in the same direction. This represents the second effect that a higher skill premium increases the value of exiting for high skilled agents and decreases for low skilled agents. These two effects contribute in opposite ways to the

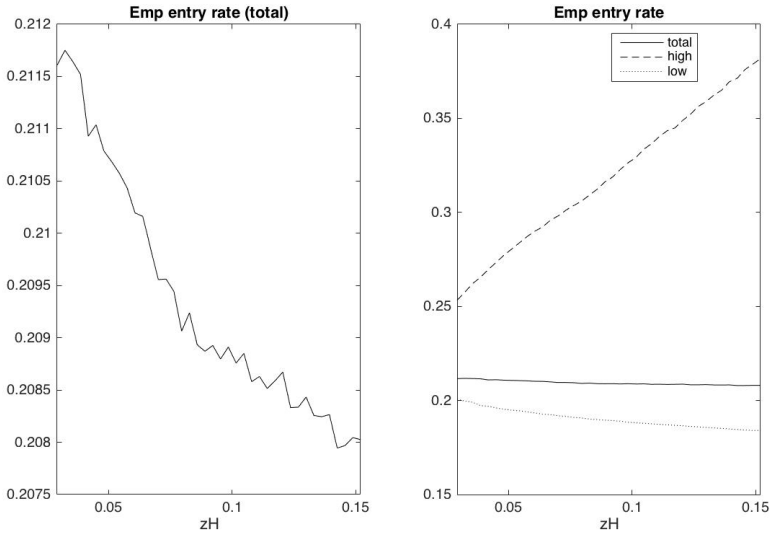


Figure 1.11: Entry rate in terms of employment

entry rate and the latter dominates the former.

Lastly, I look at the changes in the employment entry rate and average entry sizes.[8] The left panel of Figure 1.11 shows that the total employment entry rate is also declining and the right panel shows similar changes as the entry rates in high and low skilled agents. This is also because the decline in the employment entry rate by low skilled agents dominates that by high skilled agents. Figure 1.12 shows the increase in the average entry size of high skilled agents and the decrease in that of low skilled agents. Since the entry threshold is increasing for high skilled agents, the average size is also increasing and the opposite holds true for low skilled agents. The total average entry size including both high and low skilled agents is relatively flat because the changes in average entry size of high and low skilled agents cancel out.

1.6.2 The Benchmark Case

In this subsection, I quantitatively examine the differences between the stationary equilibria in 1978 and 2006. With the other parameter values fixed, I calibrate the

[8] An average entry size is defined to be the average number of employees of entrants.

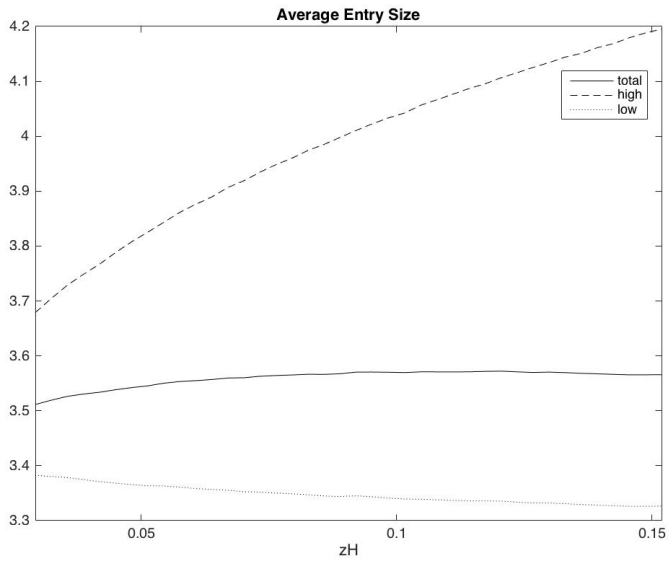


Figure 1.12: Average entry size

Variables		1978 value	2006 value
z_H/z_L		0.030	0.153
Skill Premium		1.18	1.64
Entry rate (%)	Benchmark	19.3	18.5
	data	14.5	10.8
	(w/o retail)	14.3	10.5
Employment entry rate (%)	Benchmark	21.1	20.7
	data	3.9	3.0
	(w/o retail)	3.3	2.8

Table 1.2: The change in the entry rate from the model

value of z_H to match with the skill premium in 1978 and derive the stationary equilibrium using the z_H value. Next, with the other parameter values fixed, I calibrate the value of z_H to match with the skill premium in 2006 and derive the stationary equilibrium. I compare the two stationary equilibria quantitatively. To contrast with the case in which the fraction of the population of college graduates changes, I denote this case as the benchmark case.

The results are in Table 1.2. The first row demonstrated the calibrated values of z_H for 1978 and 2006 in each column. The second row shows the observed skill

premium in the data in 1978 and in 2006 that I use to calibrate the z_H values. The third row of table 1.2 shows the decline in the entry rate by 0.8 percentage points ($= 19.3\% - 18.5\%$). In the data, the entry rate in the United States falls from 14.5% in 1978 to 10.8% in 2006, a decrease of 3.7 percentage points. Thus this model explains approximately 21% ($= 0.8/3.7$) of the change in the entry rate in the data. In this paper, I do not target the entry rate in calibration and instead target the fraction of opportunity entrepreneurs in each of high and low skilled agents for α_H and α_L. The entry rate in BDS counts both opportunity and necessity entrepreneurship that has at least one employee while entry in this model represents only opportunity entrepreneurship. Therefore the entry rate in this model is different from that in the BDS database in 1978. The fifth row of Table 1.2 is the decline in the entry rate excluding the retail sector. The size of the decline is similar to the decline in the total entry rate.

The sixth row of Table 1.2 shows the decline in the employment entry rate from 21.1% to 20.7% by the amount of 0.4 percentage points. In the data, the employment entry rate declines from 3.9% to 3.0% by the amount of 0.9 percentage points ($= 3.9\% - 3.0\%$). The employment entry rate is excessively high in this model as opposed to the data, but the model also generates the decline in the employment entry rate qualitatively.

1.6.3 Increase in high skilled labor force

The previous analysis assumed that the fraction of the population that is high skilled was the same in 1978 and 2006. In fact, this fraction increased significantly in the United States over this span. In this section, I examine how the increase in the fraction of high skilled population affects the results in the previous section. Specifically, I calibrate N_H and N_L to match with the fraction of the population that is high skilled in the data for 1978 and 2006. After that, I calibrate z_H to match w in the stationary

Variables		1978 value	2006 value
z_H/z_L		0.030	0.8567
Skill Premium		1.18	1.64
N_H		0.166	0.292
Entry rate (%)	Benchmark	19.3	18.5
	Model w/ N_H change	19.3	19.1
	data	14.5	10.8
	(w/o retail)	14.3	10.5
Employment	benchmark	21.1	20.7
entry rate (%)	Model w/ N_H change	21.1	21.8
	data	3.9	3.0
	(w/o retail)	3.3	2.8

Table 1.3: The change in the entry rate with the increase in N_H

equilibrium with the skill premium in 1978 and solve for the stationary equilibrium. In the same manner, I calibrate z_H to match w in the stationary equilibrium with the skill premium in 2006 using the N_H and N_L value for 2006 and solve for the stationary equilibrium to examine how much of the decline in the entry rate this model can explain.

The first row of Table 1.3 shows the higher increase in z_H than in the previous subsection. In this exercise, the increase in the fraction of college graduates causes more supply of high skilled work. In order to retain the same value of the skill premium, I need a larger skill biased technical change. The third row of Table 1.3 is the calibrated value of N_H for 1978 and 2006. It shows that the fraction of the population that is high skilled has approximately doubled from 17% to 30%. The fifth row demonstrates a significantly smaller decrease in the entry rate from 19.3% to 19.1% than the benchmark case. This model with the increase in the fraction of the high skilled population only explains 5.4%(=(19.3-19.1)/3.7) of the total change in the entry rate in the data. In the previous section, the result has shown that the effect of the decline in the entry rate of low skilled agents outnumbers that of high skilled agents and it generates the decline in the total entry rate. Since the

population of high skilled agents increases in this exercise, the effect by high skilled agents is becoming larger over time and the entry rate declines slightly in total. The ninth row shows the employment entry rate. The employment entry rate increases from 21.1% to 21.8% as opposed to the previous case. In the model, the average entry size of high skilled agents is larger than of low skilled agents. Therefore the effect of the decline in the entry rate of high skilled agents is amplified by the larger average entry size.

The increase in the fraction of the population that is high skilled reduces the size of the decline in the entry rate in the model and even increases the employment entry rate.

1.7 Conclusion

This paper evaluates the hypothesis that Skill Biased Technological Change (SBTC) leads to the decline in the firm entry rate in the United States. To the extent that people have to quit their jobs to start businesses, wages play the role of an opportunity cost that can avert people from entrepreneurship. SBTC increases the skill premium and leads to an increase in the opportunity cost for college graduates to start businesses and to a decrease in the opportunity cost for high school graduates. The model can explain 21% of the actual decline in the entry rate. However, allowing for the increase in the population of college graduates, the net effect of SBTC is very small. I conclude that SBTC is not an important factor in accounting for the decrease in the entry rate.

Chapter 2

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

2.1 Introduction

Wage inequality has been increasing in the United States. The skill premium—the wage premium of college graduates to high school graduates—has increased from 50% in 1980 to 90% now. A large literature (ex. Goldin and Katz (2009) and Katz and Murphy (1992)) argues that the skill premium rises because the increase in the supply of college graduates does not catch up with the increase in the demand for skilled labor. In this framework, we can reduce the skill premium by increasing college graduates in the economy. With this in mind, the existing literature often suggests policies seeking to increase college enrollment to increase college graduates. However, in the United States, while over 70% of high school graduates enroll in college, more than half of college enrollees drop out before earning a bachelor's degree

HGPA Quantile	% graduation
Q1	19%
Q2	32%
Q3	49%
Q4	63%
total	44%

Table 2.1: College graduation Rates for High school GPA Quartiles

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

(See Table 2.1)[1]. Enrollment does not necessarily lead to graduation and we should treat enrollment and graduation as two different margins. It is important for us to understand how policy can affect graduation separately as well as enrollment.

In this paper, I propose a new college subsidy scheme in which the amount of subsidies vary with years of college ("year-dependent subsidies"), i.e., subsidies that differ for freshmen, sophomores, and so on. The existing literature has only considered subsidies that are constant across years in college. Subsidies that vary by year will have differential impacts on enrollment and graduation unlike constant subsidies, as the following example suggests. After they graduate from high school, individuals decide to enroll in college based on their high school grade point average (GPA) or high school ability. People with high ability want to enroll but one's high school GPA is not necessarily the same as her college GPA. After enrolling, some students learn that their college GPA or college ability is low and drop out. Consider back-loaded subsidies in this setting: increasing subsidies for juniors and seniors and decreasing subsidies for freshmen and sophomores. People who expect to drop out before earning the increased subsidies for latter years stop enrolling due to the decreased subsidies

[1]Two-year college graduates who do not transfer are counted as dropout. According to a report from the National Center for Education Statistics for 1994-2009, more than 80 percent of community college freshmen say that their ultimate goal is a bachelors or higher degree (Horn and Skomsvold (2011)). 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)).

for early years in college. 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. It is vice versa for front-loaded subsidies. Year-dependent subsidies can affect enrollment and graduation in different ways unlike increasing or decreasing constant subsidies. The question of this paper is how year-dependent subsidies affect enrollment and graduation and what timing of college subsidies will maximize the number of college graduates and welfare.

I build a life-cycle general equilibrium model with credit constraints, endogenous enrollment and dropout decisions. Agents are heterogeneous with regard to initial asset, high school ability, and college ability. College ability affects utility in college and earnings after graduation. Agents are over-optimistic with regard to college ability before enrollment in order to be consistent with the existing empirical findings of Stinebrickner and Stinebrickner (2012)[2]. Agents learn their college abilities after enrollment and decide to drop out or not. These educational decisions shape the aggregate skill in the economy and the skill premium through imperfect substitution between skilled and unskilled labor. I calibrate the model to match the enrollment, graduation, and the skill premium of the United States given the current policy. Using the model, I examine how year-dependent subsidies have differential impacts from the current constant subsidies on enrollment, graduation, the skill premium, and the expected utilitarian social welfare function. The focus of this paper is on relative sizes across years and I fix the total budget of college subsidies at the current level.

The main finding of this paper is the following. First, it is back-loaded subsidies that maximize the number of college graduates and welfare. Second, by switching to the back-loaded subsidies with the same total budget, the number of college graduates increases and the skill premium decreases more than the case with increasing the total budget of constant subsidies by 50%. College enrollees have incentive to stay in

[2]Zafar (2011) argues optimism matters to the college major decisions

college and the share of college graduates increases, which reduces the skill premium. On the other hand, enrollment decreases due to decreased subsidies for early years in college. Third, the back-loaded subsidies improve the social welfare by 0.12% of life-time consumption at the steady state, without increasing tax. Back-loaded subsidies reduce enrollment which was excessive due to over-optimism and prevent low-ability people from enrolling and deriving disutility, which improves welfare. Second, the reduced skill premium leads to a decrease in the difference in wages between college graduates and college dropouts, which also reduces the uncertainty of wages from uncertain college ability, which is beneficial to risk-averse agents.

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

2.1.1 Related Literature

In the macroeconomic literature, Bovenberg and Jacobs (2005) theoretically derive the effect of subsidies. Abbott et al. (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 and show that the less progressive labor income tax and a large amount of subsidies than the current state are optimal for a utilitarian social welfare function.

One of the early papers of a model with college dropout is Manski (1989), who show that dropout has an option value and college enrollees can experiment on the real value of college going. Arcidiacono et al. (2016), Athreya and Eberly (2016), Lee et al. (2015), and Castex (2017) analyze how introducing college dropout into models change the allocation of human capital and returns to education. Stange (2012) and Trachter (2015) quantitatively show the importance of the option value of college

dropout quantitatively. Hendricks and Leukhina (2017) argue that college dropout is predictable before enrollment due to the strong correlation between high school GPA and the dropout rate.

Caucutt and Kumar (2003) and Akyol and Athreya (2005) show the normative implication of subsidies with exogenous college dropout risk. Hanushek et al. (2004) analyze the effect of various college aid with exogenous college dropout risk. There are some papers who regard dropout as endogenous. 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 dropout decisions and labor supply in a general equilibrium framework. Chatterjee and Ionescu (2012) argue it is welfare improving to insure student loan against risk of dropping out with endogenous dropout decisions. Findeisen and Sachs (2016) show that the optimal college subsidies are more need-based than the current system. A difference between these literature and this paper is that they have not considered subsidies that can vary across years in college.

The sequential papers of Stinebrickner and Stinebrickner (2008), Stinebrickner and Stinebrickner (2012), Stinebrickner and Stinebrickner (2014) show that borrowing constraints do not matter but learning academic ability during college is a main driver of college dropout. The reason of dropout in the model of this paper is based on their empirical findings.

2.2 Model

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

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

The third building block is an overlapping generations life cycle with incomplete markets with inter-generational linkage of ability and wealth, based on Abbott et al. (2013) and Daruich (2018). Individuals in the model face uncertainty with regard to college ability and labor productivity over life cycle with no insurance available. Individuals give birth to children with inter-generationally correlated ability and make an endogenous wealth transfer to their children. College subsidies can crowd-out endogenous wealth transfers from parents to their children.

The fourth building block is a general equilibrium framework with an aggregate production function featuring imperfect substitution between skilled and unskilled labor consistent with the findings of Goldin and Katz (2009) and Katz and Murphy (1992). The educational decisions aggregate to the supply of skill, which determines the skill premium.

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

2.2.1 Demography

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

Figure 3.1 is the timeline. At the beginning of age 1, individuals make enrollment decisions. Once they do not enroll in college, they cannot enroll later. One period in the model corresponds to two years. Consistent with college typically requiring four years in reality, college graduation requires two periods in the model. At the beginning of age 2, a college enrollee makes a decision about whether to stay in college until graduation or drop out. Once an individual finishes their schooling, they will be one of three types: high school graduates ($e = HS$) for those who do not enroll at age 1, college dropouts ($e = CD$) for those who do not continue college at the beginning of age 2, and college graduates ($e = CG$) for those who finish two periods of college. After that, they face a standard life cycle problem with income risk.

Individuals give birth to children at age $j_f = 7$ which is biological age 30 ($j_f = (30 - 18)/2 + 1 = 7$). At age $j_b = 16$ (at biological age 48), their children leave and become economically independent and parents transfer wealth to their children. There are no transfers allowed at other ages[3]. Individuals retire at age $j_r = 25$ (at biological age 66) and the maximum age is $J = 42$ (at biological age 100).

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

2.2.2 Preferences

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

[3]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 change 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 allowing transfers from parents at age 2 changes the outcome.

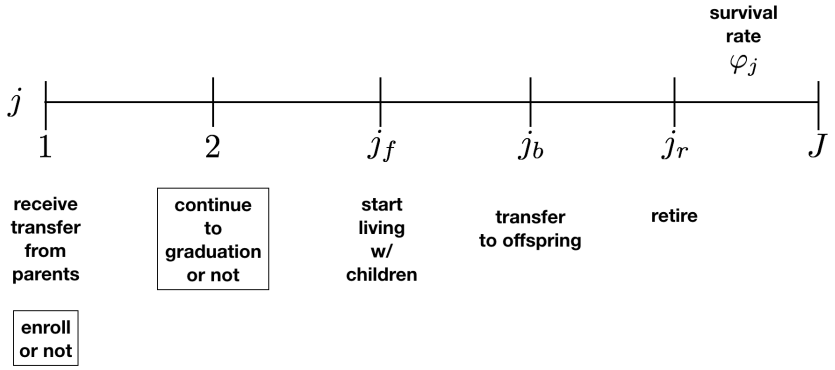


Figure 2.1: Timeline.

1. The expected discounted sum of instant utility:

$$\mathbb{E}_1 \sum_{j=1}^{J} \tilde{\beta}^{j-1} u(c_j, \ell_j) \tag{2.1}$$

where

$$u(c, \ell) = \frac{\left(c^{\mu} \ell^{1-\mu}\right)^{1-\gamma}}{1-\gamma} \tag{2.2}$$

and c_j is consumption and ℓ_j is leisure at age j. $\mathbb{E}_1$ is the expectation operator conditional on the information at the beginning of age 1. The individuals are endowed with one unit of time each period. At age $j \in [j_f, j_b - 1]$, individuals live with their children and consumption is discounted by $1 + \zeta$ where ζ is an adult equivalence parameter. β is the time discount rate.[4]

2. The expected utility of college attendance:

$$\mathbb{E}_1 d_0(\boldsymbol{s}_0) \lambda_1(\theta_c, \phi) + \beta \mathbb{E}_1 d_1(\boldsymbol{s}_1^c) \lambda_2(\theta_c, \phi) \tag{2.3}$$

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

where

$$\lambda_j(\theta_c, \phi) = \lambda + \lambda^{\theta}\theta_c + \lambda_j^{\phi}\phi \tag{2.4}$$

and $d_1(\boldsymbol{s}_0)$ is an indicator function which is one if the individual enrolls at age 1 and $d_2(\boldsymbol{s}_1^c)$ is an indicator function for enrolling at age 2. As in Heckman et al. (2006), the psychic cost of education is an important factor determining educational choice. I define $\lambda_j(\theta_c, \phi)$ not as disutility but as utility without loss of generality. Utility of college attendance depends on two components: college ability θ_c and college taste ϕ. ϕ is fixed over lifetime while the coefficient λ_j^{ϕ} can vary across periods (different loading).

I need the two different factors θ_c and ϕ for utility of college attendance to match the data. In the data, within a category of high school ability and family income, there is heterogeneity in terms of enrollment decisions (some enroll but others don not). To explain it, I need college taste ϕ which is unobservable to econometricians. Individuals observe their college tastes ϕ before enrollment. I explain ability and college taste in more detail in the section of the individual problems.

3. Parental altruism.

$$\beta^{j_b - 1}\nu\mathbb{E}V_0 \tag{2.5}$$

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

2.2.3 Goods Sector

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

$$Y = F(K, H) = K^{\alpha} H^{1-\alpha} \tag{2.6}$$

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

$$H = (a^S (H^S)^{\rho} + (1 - a^S)(H^U)^{\rho})^{\frac{1}{\rho}} \tag{2.7}$$

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

$$r = \alpha \left(\frac{K}{H}\right)^{\alpha-1} - \delta \tag{2.8}$$

$$w^S = (1 - \alpha) a^S \left(\frac{K}{H}\right)^{\alpha} \left(\frac{H}{H^S}\right)^{1-\rho} \tag{2.9}$$

$$w^U = (1 - \alpha)(1 - a^S) \left(\frac{K}{H}\right)^{\alpha} \left(\frac{H}{H^U}\right)^{1-\rho} \tag{2.10}$$

There are two types of skill in production while there exist three levels of education. In the literature on the skill premium as in Katz and Murphy (1992), high school graduates are assumed to provide unskilled labor and college graduates provide skilled labor. I assume college dropouts provide unskilled labor. Torpey and Watson (2014)[5]

[5]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. Bu-

present the proportion of jobs in the United States by required education level such as "Bachelor's degree", "Associate's degree", "Some college, no degree", "High school diploma or equivalent", or "Less than high school". They show that only 5% of jobs require "Some college, no degree" and "Associate's degree", which implies that most college dropouts take jobs requiring education level "High school diploma or equivalent". Thus I assume college dropouts provide unskilled labor. For convenience, I define the price of effective labor by college graduates, college dropouts, and high school graduates as $w^{CG} = w^S$ and $w^{HS} = w^{CD} = w^U$.

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

2.2.4 College

There is a representative college. To provide a college enrollee with one period of education requires κ units of skilled labor, which means college enrollees receive education from professors who are college graduates. I assume education does not require any capital or unskilled labor.[6]

The profit of college is

$$p_e E - w^S \kappa E \tag{2.11}$$

where E is the measure of college enrollees and p_e denotes tuition. Colleges are competitive and there is free entry. This implies, in equilibrium with positive units

reau of Labor Statistics.

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

of students, that $p_e = w^S \kappa$. In the United States, colleges receive subsidies from governments, which enables the sticker tuition smaller than the actual education cost. I reinterpret this situation as follows: colleges do not receive any subsidy while college enrollees receive subsidies instead. At the same time, they have to pay the full education cost. In both cases, enrollees pay p_e less the subsidy to colleges.

2.2.5 Financial Markets

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

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

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

2.2.6 Individual Problems

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

Education Stage

Enrollment

At the beginning of $j = 1$, individuals become independent as high school graduates and their first decision is whether to enroll in college or not. I define V_0 to be

the value function.

$$V_0(a, \theta_h, \eta, q, \phi) = \max\{\underbrace{V_1^c(a, \theta_h, \eta, q, \phi)}_{\text{enrolling}}, \underbrace{V_1(a, HS, \theta_h, \eta)}_{\text{not enrolling}}\} \tag{2.12}$$

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

There are two types of ability that are distinct but correlated to each other: high school ability θ_h and college ability θ_c. Individuals observe high school ability but do not observe college ability before the enrollment decision. Individuals observe their high school abilities through high school GPA or test scores during high school. College ability is only observed at the beginning of age 2. Stinebrickner and Stinebrickner (2012) present evidence that college enrollees do not have perfect foresight of their college abilities before enrollment. However, college abilities are correlated with high school abilities and

$$\theta_c = \theta_h + \epsilon_c \text{ where } \epsilon_c \sim N(0, \sigma_c^2) \tag{2.13}$$

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

downward revisions of their expectations. Given θ_h, college enrollees expect that

$$\theta_c = \underbrace{\mu_c(\theta_h)}_{\text{bias}} + \underbrace{\theta_h + \epsilon_c}_{\text{actual ability}} \quad \text{where } \epsilon_c \sim N(0, \sigma_c^2) \tag{2.14}$$

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

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

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

First half of college

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

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

subject to

$$c + a' + p_e - s_1(q) = a + y - T(c, a, y) \tag{2.15}$$

$$y = w^{HS} \varepsilon_1^{HS}(\theta_h, \eta) h, \;\; a' \geq -\underline{A}_1^c \;\; c \geq 0, \;\; 0 \leq h \leq 1 - \bar{h} \tag{2.16}$$

$$\theta_c = \theta_h + \mu_c(\theta_h) + \epsilon_c, \;\; \epsilon_c \sim N(0, \sigma_c^2) \text{ (perceived process)}, \;\; \eta \sim \Pi^{CD} \tag{2.17}$$

Going to college requires a fraction $\bar{h}$ of time, tuition p_e and additive utility $\lambda_j(\theta, \phi)$

for each enrolling period. c is consumption, ℓ is labor hours, y is labor earnings and a' is the next period assets. The total tax $T(c, a, y)$ depends on consumption, asset holdings, and earnings. College enrollees receive subsidies $s_j(q)$ dependent on family income q. They can work as high school graduates during the first half of college.

At the end of the period, college enrollees observe their college abilities θ_c and a new idiosyncratic productivity η' drawn from Π^{CD}. College enrollees choose whether or not to drop out of college after this. If the individual drops out, her education level becomes college dropout ($e = CD$) and their value is V_2. Her labor productivity depends on θ_h in the working stage. After dropping out, all the student loan is refinanced into a single bond that carries interest rate r^-. $\tilde{a}(a)$ is the transformation from the asset position during college to the position after college so that the total payment is identical. When making this calculation I assume that fixed payments would have been made for 20 years (10 periods) after dropout and

$$\tilde{a}(a') = a' \times \frac{r^s}{1-(1+r^s)^{-10}} \times \frac{1-(1+r^-)^{-10}}{r^-} \tag{2.18}$$

If the individual does not drop out, she proceeds 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

$$V_2^c(a, \theta_c, \eta, q, \phi) = \max_{c,h,a',y} u(c, 1-h-\bar{h}) + \lambda_2(\theta_c, \phi) + \beta \mathbb{E}_{\eta'} V_3(\tilde{a}(a'), CG, \theta_c, \eta' s) \tag{2.19}$$

subject to

$$c + a' + p_e - s_2(q) - y + T(c, a, y) = \begin{cases} (1+r)a & \text{if } a \geq 0 \\ (1+r^s)a & \text{if } a < 0 \end{cases} \tag{2.20}$$

$$y = w^{CD}\varepsilon_2^{CD}(\theta_c, \eta)h, \;\; a' \geq -\underline{A}_2^c \;\; c \geq 0, \;\; 0 \leq h \leq 1 - \bar{h}, \;\; \eta' \sim \Pi^{CG} \tag{2.21}$$

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

Working Stage

The Bellman equation for workers is[7]

$$V_j(a, e, \theta, \eta) = \max_{c,h,a',y} u\left(\frac{c}{1 + \mathbf{1}_{\mathcal{J}_f}\zeta}, 1 - h\right) + \beta \mathbb{E}_{\eta'|\eta} V_{j+1}(a', e, \theta, \eta') \tag{2.22}$$

subject to

$$c + a' - y + T(c, a, y) = \begin{cases} (1 + r)a & \text{if } a \geq 0 \\ (1 + r^-)a & \text{if } a < 0 \end{cases} \tag{2.23}$$

$$y = w^e \varepsilon_j^e(\theta, \eta)h, \;\; a' \geq -\underline{A}^e \;\; c \geq 0, \;\; 0 \leq h \leq 1, \;\; \eta' \sim \pi^e(\cdot|\eta) \tag{2.24}$$

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

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

Transfer

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

$$V_{j_b}(a,e,\theta,\eta) = \mathbb{E}_{\theta'_h|e,\theta}\left\{\max_{b(\theta_h)\in[0,a]} \tilde{V}_{j_b}(a-b,e,\theta,\theta'_h,\eta) + \nu\mathbb{E}_{\eta'',\phi}V_0(b,\theta'_h,\eta'',\tilde{q}(w^e\varepsilon^e_j(\theta,\eta)),\phi)\right\} \tag{2.25}$$

subject to

$$\theta'_h \sim \pi_\theta(\cdot|\theta),\ \ \eta'' \sim \Pi^{HS},\ \ \phi \sim N(0.1) \tag{2.26}$$

where

$$\tilde{V}_{j_b}(a,e,\theta,\theta'_h,\eta) = \max_{c,h,a',y} u(c,1-h) + \beta\mathbb{E}_{\eta'|\eta}V_{j_b+1}(a',e,\theta,\eta') \tag{2.27}$$

subject to

$$c + a' - y + T(c,a,y) = \begin{cases} (1+r)a & \text{if } a \geq 0 \\ (1+r^-)a & \text{if } a < 0 \end{cases} \tag{2.28}$$

$$y = w^e\varepsilon^e_j(\theta,\eta)h,\ \ a' \geq -\underline{A}^e\ \ c \geq 0,\ \ 0 \leq h \leq 1,\ \ \eta' \sim \pi^e(\cdot|\eta) \tag{2.29}$$

At the beginning of the period, parents choose their transfer of wealth to their children b. Before making any decisions, parents observe their children's high school ability θ'_h whose density function is $\pi_\theta(\theta'_h|\theta)$ but neither their children's initial idiosyncratic productivity η'' drawn from Π^{HS} nor college taste ϕ drawn from the normal distribution $N(0,1)$. Consumption, leisure, asset holdings, and parental transfers can depend on θ'_h. The lifetime utility of their children depends on family income level q

which is a function of the potential labor income of the parents.[8]

Retirement Stage

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

$$V_j(a, e, \theta) = \max_{c,a'} u(c, 1) + \beta\varphi_{j+1}V_{j+1}(a', e, \theta) \tag{2.31}$$

subject to

$$c + a' = (1 + r)\varphi_j^{-1}a + p(e, \theta) - T(c, \varphi_j^{-1}a, 0) \tag{2.32}$$

$$a' \geq 0 \;\; c \geq 0 \tag{2.33}$$

The sources of income are interest payments and retirement benefits $p(e, \theta)$. In the United States, retirement benefits are determined by the labor earnings before retirement (see Appendix 3.B). To capture this, the retirement benefits depend on their abilities and education. The asset inflated by φ_j^{-1} reflects that assets of expiring individuals are distributed within cohorts (perfect annuity market).

2.2.7 Government

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

[8]Note that the parental income is not the actual labor income. The parents can control the actual labor income by adjusting their working hours. In this setting, this manipulation of parental income is not allowed and parental income is a function of "potential" income which is labor earnings if they spend 35% working. Thus the family income mapping is

$$\tilde{q}(w^e\varepsilon_j^e(\theta, \eta)) = \begin{cases} 1 & \text{if } w^e\varepsilon_j^e(\theta, \eta) \times 0.35 \in [0, q_1] \\ 2 & \text{if } w^e\varepsilon_j^e(\theta, \eta) \times 0.35 \in [q_1, q_2] \\ 3 & \text{else} \end{cases} \tag{2.30}$$

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

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

$$G_e = \sum_{j=1,2} \int_{S_j^c} s_j(q) d\mu_j^c \tag{2.34}$$

The tax function is

$$T(c, a, y) = \tau_c c + \tau_k r a \mathbf{1}_{a \geq 0} + \tau_l y - d\frac{Y}{N} \tag{2.35}$$

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

2.2.8 Equilibrium

The model includes J overlapping generations and is solved numerically to characterize a stationary equilibrium in which the cross-sectional allocation is invariant. In equilibrium, individuals maximize expected lifetime utility, the representative firm and college maximize profits, the government budget is balanced each period, and prices clear all the markets. Let $\boldsymbol{s}_j^c \in S_j^c$ be the age-specific state vector for college enrollees and $\boldsymbol{s}_j \in S_j$ for workers and retirees and $\boldsymbol{s}_0 \in S_0$ for individuals at the beginning of age 1. I also define the age-specific state vector for workers and retirees conditional on education e as $\boldsymbol{s}_j^e \in S_j^e$. Computation is described in Appendix 3.A.

Definition 2. *A stationary equilibrium is a list of value functions of workers and college enrollees* $\{V_j(\boldsymbol{s}_j), V_j^c(\boldsymbol{s}_j^c)\}$*, decision rules of enrollment* $d_1(\boldsymbol{s}_0)$ *and graduation* $d_2(\boldsymbol{s}_1^c)$*, decision rules of consumption, asset holdings, labor hours, output, parental transfers of workers* $\{c_j(\boldsymbol{s}_j), a'_j(\boldsymbol{s}_j), h_j(\boldsymbol{s}_j), y_j(\boldsymbol{s}_j), b(\boldsymbol{s}_j)\}$*, decision rules of college enrollees* $\{c_j^c(\boldsymbol{s}_j^c), a'^c_j(\boldsymbol{s}_j^c), h_j^c(\boldsymbol{s}_j^c), y_j^c(\boldsymbol{s}_j^c)\}$*, aggregate enrollees, capital, and labor inputs*

$\{E, K, H^S, H^U\}$, *prices* $\{r, w^S, w^U, p_e\}$, *policy* τ_ℓ, *measures* $\boldsymbol{\mu} = \{\mu_j^c(\boldsymbol{s}_j^c), \mu_j(\boldsymbol{s}_j), \mu_j^e(\boldsymbol{s}_j^e)\}$ *such that*

1. *Taking prices and policies as given, value functions* $\{V_j^c(\boldsymbol{s}_j^c), V_j(\boldsymbol{s}_j)\}$ *solve the individual Bellman equations and* $d_1(\boldsymbol{s}_0)$, $d_2(\boldsymbol{s}_1^c)$, $\{c_j(\boldsymbol{s}_j), a_j'(\boldsymbol{s}_j), h_j(\boldsymbol{s}_j), y_j(\boldsymbol{s}_j), b(\boldsymbol{s}_j)\}$, $\{c_j^c(\boldsymbol{s}_j^c), a_j'^c(\boldsymbol{s}_j^c), h_j^c(\boldsymbol{s}_j^c), y_j^c(\boldsymbol{s}_j^c)\}$ *are associated decision rules.*

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

3. *The government budget is balanced.*

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

where

$$G_c = gF(K, H) \tag{2.36}$$

$$G_e = \sum_{j=1,2} \int_{S_j^c} s_j(q) d\mu_j^c \tag{2.37}$$

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

$$H^S + \kappa E = H^{CG} \tag{2.38}$$

$$H^U = H^{HS} + H^{CD} \tag{2.39}$$

where

$$H^{CG} = \sum_{j=3}^{j_r - 1} \int_{S_j^{CG}} \epsilon_j^{CG}(\theta, \eta) h_j(\boldsymbol{s}_j) d\mu_j^{CG} \tag{2.40}$$

$$H^{CD} = \sum_{j=2}^{j_r-1} \int_{S_j^{CD}} \epsilon_j^{CD}(\theta,\eta) h_j(\boldsymbol{s}_j) d\mu_j^{CD} + \int_{S_2^c} \epsilon_2^{CD}(\theta,\eta) h_2^c(\boldsymbol{s}_2^c) d\mu_2^c \tag{2.41}$$

$$H^{HS} = \sum_{j=1}^{j_r-1} \int_{S_j^{HS}} \epsilon_j^{HS}(\theta,\eta) h_j(\boldsymbol{s}_j) d\mu_j^{HS} + \int_{S_1^c} \epsilon_1^{HS}(\theta,\eta) h_1^c(\boldsymbol{s}_1^c) d\mu_1^c \tag{2.42}$$

and

$$K = \sum_{j=1}^{J} \int_{S_j} a_j'(\boldsymbol{s}_j) d\mu_j + \sum_{j=1,2} \int_{S_j^c} a_j'^c(\boldsymbol{s}_j^c) d\mu_j^c \tag{2.43}$$

$$E = \sum_{j=1,2} \int_{S_j^c} d\mu_j^c \tag{2.44}$$

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

2.3 Calibration

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

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

2.3.1 Labor Productivity Process

I assume labor productivity

$$\ln \epsilon_j^e(\theta, \eta) = \ln \epsilon^e + \ln \psi_j^e + \epsilon_\theta^e \theta + \ln \eta \tag{2.45}$$

where ψ_j^e is the age profile of workers at education level e estimated from the PSID (See Appendix 2.C). The coefficients can vary across education levels.

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

$$\ln \epsilon^e + \ln \psi_j^e + \epsilon_\theta^e \theta_c + \ln \eta = \ln \epsilon^e + \ln \psi_j^e + \epsilon_\theta^e \theta_h + (\ln \eta + \epsilon_\theta^e \epsilon_c) \tag{2.46}$$

Since θ_h is uncorrelated with $\ln \eta + \epsilon_\theta^e \epsilon_c$, I can estimate the coefficient ϵ_θ^{CG} using ln AFQT80 [9]. Table 3.8 shows the estimated coefficients on ability for each education level. As in the literature, returns to education are higher for high ability.

ϵ^e is the intercepts of log labor productivity. I normalize $\epsilon^{HS} = \epsilon^{CG} = 0$ and calibrate ϵ^{CD} to match the wage premium of college dropouts as explained later.

I assume $\pi_\eta^e(\eta'|\eta)$ is a Markov chain with two states η_H and η_L specific to each education level e. It has exactly the same persistence and conditional variance as the

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

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

Table 2.2: Estimated ability slope ϵ^e_θ of labor productivity

Source: NLSY79. See the Data Appendix for the detail.

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

Table 2.3: Estimated parameters of the residual labor productivity process

Source: NLSY79. See the Data Appendix for the detail.

AR(1) process:

$$\ln \eta' = \rho^e \ln \eta + \epsilon^e_\eta, \quad \epsilon^e_\eta \sim N(0, \sigma^{e2}_\eta) \tag{2.47}$$

After filtering out age effects using the PSID, I employ a Minimum Distance Estimator with a fixed effect and a measurement error. I use as moments the covariances of the wage residuals at different lags and age groups, separately for each education level. In Appendix 2.C, I discuss sample selections and the detail of the estimation procedures. Table 3.9 is the estimated parameters.

2.3.2 Intergenerational Ability Transmission

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

$$\theta'_h = m + m_\theta \theta + \epsilon_h, \quad \epsilon_h \sim N(0, \sigma^2_h) \tag{2.48}$$

High school ability is formed partly as a result of genetics, which leads to a correlation between parents' and childrens' ability. In addition, as Cunha and Heckman (2007), Cunha (2013), and Daruich (2018) suggest, high ability parents earn higher income, which increases early educational investment and improves their children's high school ability.

In order to estimate the conditional mean of the inter-generational ability transmission, I regressed children's ability on parents' ability in NLSY79 to obtain the estimated value of m_θ 0.46[10].

2.3.3 Subsidies and Loans

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

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

I adopt Abbott et al. (2013) for the cost for college enrollees and the subsidy system of the Unites States (see Table 3.10 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 full cost of education above and the cost for enrollees, $\$17,489 - \$6,710 = \$10,779$. In the model, the subsidies for college enrollees $\bar{s}(q)$ are the sum of this subsidy and the subsidies as in Table 3.10. In the current system, college subsidies are constant across periods in college and $s_1(q) = s_2(q) = \bar{s}(q)$.

[10]For college dropouts and college graduates, θ is college ability but I use ln AFQT80 as an

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

Table 2.4: subsidies and family income

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

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 case so I focus on unsubsidized loan. Students' interest rate is the prime rate plus 2.3% ($= \iota^s$, annual). I assume students face a borrowing limit $\bar{A}_j$ 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.

2.3.4 Government Policy

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

instrument as in the estimation of the labor productivity process.

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

Table 2.5: Parameters determined outside the model.

2.3.5 Production

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

2.3.6 The Remaining Parameters

Given the parameter values set outside the model in Table 3.11, there are 16 remaining parameters: the bias of expectation of college ability (μ_c^0, μ_c^1), college utility $(\lambda, \lambda^\theta, \lambda_1^\epsilon, \lambda_2^\epsilon)$, the standard deviation of college ability σ_c, productivity of labor (a^S, ϵ^{CD}), education cost κ, utility parameters (μ, β, v), lump-sum transfer d, overseeing cost ι, and inter-generational ability parameters (m, σ_h).

I choose 27 moments in Table 3.13[11] and minimize the average Euclidean percent-

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

Moment	Model	Target
Enrollment rate of ability quartile	(figure)	(figure)
Graduation rate of ability quartile	(figure)	(figure)
Enrollment rate of family income quartile	(figure)	(figure)
Graduation rate of family income quartile	(figure)	(figure)
Skill premium for CG	90.7%	90.2%
Skill premium for CD	19.7%	19.9%
Expected/Actual graduation rate -1	0.429	0.433
Education cost/mean income at 48	0.318	0.33
Hours of work	34.0%	33.3%
K/Y	1.305	1.325
Transfer/mean income at 48	67.1%	66%
log pre-tax/post-tax income	61.3%	61%
Borrowers	6.78%	6.3%
Mean of AFQT	-0.0135	0
Standard deviation of AFQT	0.217	0.213

Table 2.6: Moments matched and the model fit.

age deviation of the model from the data[12]. 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 before enrollment 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%($= 0.86/0.60 - 1$). The mean and standard deviation of high school ability are also important to replicate the distribution of the ability in the United States.

The third column of Table 3.12 presents the calibrated values. The calibrated value of μ_c^0 is positive and enrollees are optimistic about their college ability on average. Since the standard deviation of college ability is 0.40[13], the bias for the mean ability is 48% of the standard deviation of college ability. In addition, μ_c^1 is negative and

[12]For the mean of high school ability, I chose 5.04, which is the mean of ln AFQT80 in the data, for the denominator of the percent deviation. I do not take the percent deviation for the enrollment and graduation rates.

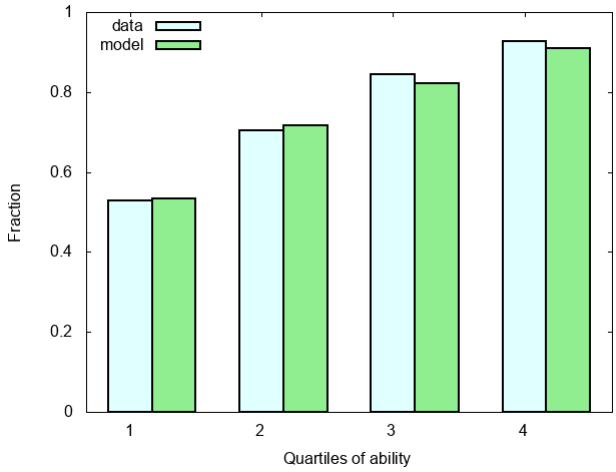
[13]The square root of the sum of the variance of high school ability and σ_c^2. $0.40 = \sqrt{0.213^2 + 0.341^2}$

Parameter	Description	Value
μ_c^0	college ability bias intercept	0.190
μ_c^1	college ability bias slope	-0.413
λ	college utility intercept	-22.2
λ^θ	college utility slope	239
λ_1^ϕ	first period college taste	64.0
λ_2^ϕ	second half college taste	40.5
a^S	productivity of skilled labor	0.446
ϵ^{CD}	productivity of CD	1.03
σ_c	s.d. of college ability	0.341
κ	education cost	0.226
μ	consumption share of preference	0.419
β	time discount rate	0.939
v	altruism	0.0952
d	lump-sum transfer ratio	0.125
ι	borrowing wedge ($r^- = r + \iota$)	18.1%
m	intergenerational ability transmission intercept	-0.0471
σ_h	intergenerational ability transmission s.d.	0.171

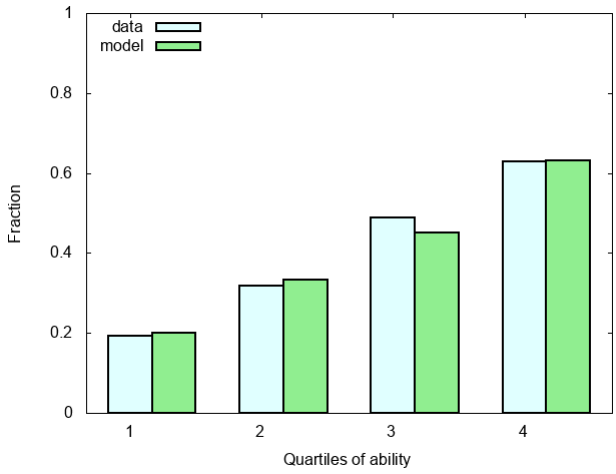
Table 2.7: The remaining parameters.

enrollees with lower high school ability are more optimistic than enrollees with higher high school ability. These characteristics are consistent with the pattern of the bias of college GPA observed in Stinebrickner and Stinebrickner (2012). λ^0 is negative and agents derive disutility from college. A positive λ^1 implies that the disutility is smaller for agents with high ability.

The model fit is presented in Table 3.13 and Figures 3.2 and 3.3. In general, the model fits well considering over-identification of 16 parameters against 27 moments. In the data, ability is correlated with enrollment and graduation more than family income and the model captures this pattern. Although the graduation rates across family income are somewhat flatter than the data, they capture the key pattern. The enrollment and graduation rates are higher for the second quartile than for the third quartile in the model because there are only three bins for family income q and there is a jump of college subsidies when people cross over the threshold of family income.



(a) Enrollment rate



(b) Graduation rate

Figure 2.2: Model fit: enrollment and graduation rate for each ability quartile.

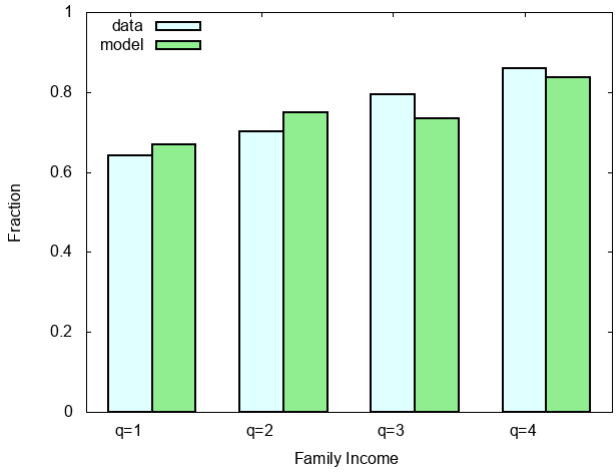
Data Source: NLSY97. I use the sample of only 25 year old high school graduates. Ability is the log of AFQT score using the definition from NLSY79. The scores are adjusted by age as in Altonji et al. (2012) and Castex and Kogan Dechter (2014).

2.3.7 Validation Exercises

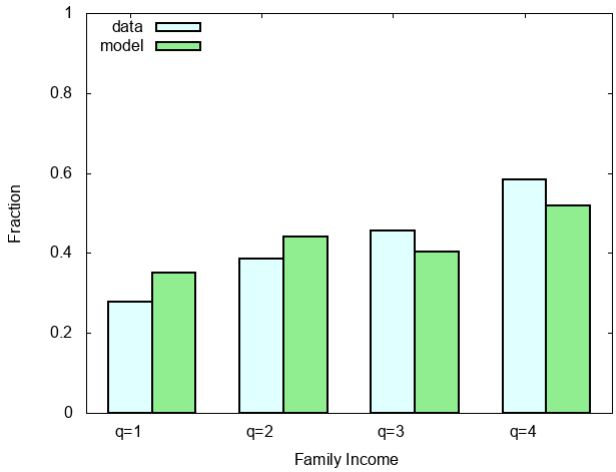
In this subsection, I compare the model simulation with the data for the moments I do not target when calibrating.

The Partial Equilibrium Effect of Year-Invariant subsidies

The elasticity of enrollment with regard to tuition or subsidies has been examined in the micro empirical literature. I simulate the partial equilibrium response of enroll-



(a) Enrollment rate



(b) Graduation rate

Figure 2.3: Model fit: enrollment and graduation rates for each family income quartile.

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

ment to a $1,000 increase in subsidies for all the years in college and family income evenly and compare with the estimates of the empircial literature. All the prices and the distribution at age 1 are fixed at the current level and these additional subsidies are given to only one generation.

The aggregate enrollment rate of the affected generation increases by 1.13 percentage points in the simulation. The micro-emprirical literature has estimates of

the effect of subsidies on enrollment by Dynarski (2002), Kane (1994), and Cameron and Heckman (2001). While this literature argues that the enrollment rate of groups 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 survey of the literature). Therefore the simulation is broadly in the range of the literature. In addition, the increase in enrollment is smaller in the model than in the data, which implies this calibration is a more conservative choice. If the response of enrollment is high, the effect of changing college subsidies is also high and I overestimate the effect of switching to year-dependent subsidies.

In the simulation, the share of college graduates increases by 0.50 percentage points and that of college dropouts increases by 0.62 percentage points. Not only the people who are induced to enroll by the additional subsidy but also those who would already have enrolled without the additional subsidy have incentive to stay until graduation. This is consistent with Dynarski (2008), Castleman and Long (2016), Scott-Clayton (2011), Scott-Clayton and Zafar (2019), Denning et al. (2017), and Bettinger et al. (2019) who all find a positive effect of subsidies on graduation.

The Sluggish Increase in College Graduates

The increase in college graduates has been sluggish in the United States since 1980 despite the increase in the skill premium during the same period. In this subsection, I examine how well the model can explain this sluggish increase by targeting the skill premium of 1980 and 2000 in the United States. The benchmark calibration is targeted to the United States in 2000 and I assume only the productivity of skilled labor a^S and productivity of college dropouts ϵ^{CD} change in the model between 1980 and 2000. In particular, I set the values of a^S and ϵ^{CD} to match the college graduate wage premium 46.2% and the college dropout wage premium 12.1% as observed in 1980 in the United States with the other parameter values fixed. I compute the steady

	1980	2000	change (model)	change (data)
college graduate premium	46.2%	90.7%	44.5pp	43.2pp
college dropout premium	12.1%	19.7%	7.6pp	7.4pp
share of college graduates	28.3%	32.9%	4.6pp	4.98pp
share of college dropouts	42.8%	41.4%	-1.4pp	2.41pp

Table 2.8: Change in the share of college graduates and dropouts

state with the new values to replicate 1980 in the United States. The first two rows of Table 2.8 show that the wage premiums for college graduates and dropouts in the model and the data. By definition, the change in the model and in the data are the same. I compare the change in the share of college graduates and dropouts with the data.[14]

The third and fourth rows of Table 2.8 show the change in the share of college graduates and dropouts between 1980 and 2000. As the college graduate premium increases by 43.7 percentage points from 1980 to 2000, the third column shows the share of college graduates increases by 5.3 percentage points. The model can explain the sluggish increase in the share of college graduates. 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.

2.4 Results

The section is composed of three exercises. In the first exercise, I increase overall spending without changing the structure of subsidies, financed by the proportional labor income tax, and examine how it affects enrollment, graduation, and the skill premium. In the second exercise, I keep total spending fixed but choose subsidies by

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

G_e	0.75 $\bar{G}_e$	$\bar{G}_e$	$1.5\bar{G}_e$	$2\bar{G}_e$
enrollment rate	72.6%	74.2%	77.2%	78.0%
share of college graduates	32.1%	32.9%	34.3%	35.1%
skill premium	94.9%	90.7%	82.5%	77.7%

Table 2.9: Education and the skill premium for the current college subsidy scheme.

year to maximize the number of college graduates in the steady state and compare its effect with the first exercise. In the third exercise, I keep total spending fixed and choose subsidies by year to maximize the utilitarian social welfare function in the steady state as a normative analysis.

2.4.1 The Effect of Year-Invariant Subsidies

As a benchmark case, I examine the general equilibrium effect of a permanent change in the total budget of the current college subsidy scheme which I call year-invariant subsidies. Table 2.9 shows how the enrollment rate defined as the number of college enrollees, the number of college graduates, and the skill premium change as the government subsidies budget increases. $\bar{G}_e$ denotes the current level government total budget for college subsidies. I consider an increase in the budget of college subsidies G_e from $0.75\bar{G}_e$ to $2\bar{G}_e$. The proportional labor income tax rate τ_l is adjusted to the change in the budget. The subsidies across college years and family income proportionally change with G_e fixed.

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

2.4.2 The Effect of Year-Dependent Subsidies

In this subsection, I derive the year-dependent subsidies that maximize the number of college graduates and show how it affects enrollment, graduation, and the skill pre-

mium. I fix the total spending at the current level $\bar{G}_e$ and only allow the relative sizes of subsidies to differ across college years. The maximization problem is formulated as

$$\max_{g_1>0,g_2>0,\tau_\ell} \int_{S_2^c} d\mu_2^{CG} \tag{2.49}$$

subject to

$$\int_{S_1^c} g_1 \bar{s}(q) d\mu_1^c + \int_{S_2^c} g_2 \bar{s}(q) d\mu_2^c = \bar{G}_e \tag{2.50}$$

and the government budget constraint. The new subsidies are $s_1(q) = g_1\bar{s}(q)$ and $s_2(q) = g_2\bar{s}(q)$ where $\bar{s}(q)$ is the current college subsidy system. In this problem, the government chooses the general levels of college subsidies g_j for each period j compared to the current system. If I increase subsidies for the first half of college g_1, the general level of subsidies for the second half g_2 has to decrease.[15]s

Note that I do not allow to change the relative subsidies across different family income and ability within a year in college from the current system. For example, the ratio of subsidies for $q = 1$ and $q = 2$ in the first half of college is fixed at the current state. This paper focuses on how the year-dependency of college subsidies affects educational choices and the skill premium and abstracts from the analysis of the optimal need-based and merit-based subsidies[16].

Table 2.10 shows the annual amount of the subsidies of the two cases. The first column is identical to the case of the year-invariant subsidies with $G_e = \bar{G}_e$. The second column is the college subsidies that maximize the number of college graduates. The first three rows are college subsidies at the first half of college across family income level $q = 1$ to 3 from the top to the bottom. The next three rows are college

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

[16]See Findeisen and Sachs (2016).

$s_j(q)$	year-invariant $\bar{G}_e$	year-dependent $\bar{G}_e$
$s_1(1)$	\$13,599	\$20
$s_1(2)$	\$11,447	\$17
$s_1(3)$	\$10,922	\$16
$s_2(1)$	\$13,599	\$42,376
$s_2(2)$	\$11,447	\$35,670
$s_2(3)$	\$10,922	\$34,034

Table 2.10: Year-dependent subsidies maximizing the number of college graduates.

	year-invariant $\bar{G}_e$	year-dependent $\bar{G}_e$	year-invariant $1.5\bar{G}_e$
enrollment rate	74.2%	68.8%	77.2%
share of college graduates	32.9%	34.6%	34.3%
skill premium	90.7%	82.0%	82.5%

Table 2.11: Education and the skill premium for the optimal subsidy scheme.

subsidies at the second half. The optimal year-dependent subsidies are back-loaded: subsidies are more generous for the second half than for the first half. The optimal year-dependent subsidies for the first half is negligible[17].

Table 2.11 display the enrollment rate, the share of college graduates for each case. Year-dependent subsidies reduce the enrollment rate by 5.4 percentage points and increase the share of college graduates by 1.7 percentage points. To contrast, the third column presents the year-invariant case of $G_e = 1.5\bar{G}_e$. The year-dependent subsidies increases the share of college graduates and reduces the skill premium more than the case with increasing the total budget of year-invariant subsidies by 50%. Changing the structure of college subsidies is as effective as increasing the budget by 50%.

The mechanism of the effect of year-dependent subsidies is the following. In the current system, already 70% who graduate from high school enroll in college. Increasing enrollment will basically encourage more people to enroll who are likely to drop out. This means that the enrollment margin is not so important from the

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

perspective of getting people to graduate. The marginal person who drops out is better able to benefit from college than the marginal person who does not enroll. It is easier to create incentives for the marginal dropout to finish than to create incentives for the marginal non-enrollee to enroll and finish. Decreasing subsidies for the first period serves mainly to discourage people who are unlikely to graduate from enrolling. The higher subsidies for the second period encourages marginal dropouts to finish.

There is another mechanism of the back-loaded subsidies. In the current system, the government has paid subsidies to all the people who enroll but drop out. With the back-loaded subsidies, the government does not need to pay high subsidies to people who drop out before the second period and can give more subsidies to the people who graduate, which increases the number of college graduates. In fact, as Table 2.10 shows, the sum of the subsidies to college graduates for the two periods for the middle family income is \$35,687($= \$17 + \$35,670$) which is higher than the case of the current system \$22,894($= \$11,447 + \$11,447$). The back-loaded subsidies are more cost-effective from the perspective of increasing the number of college graduates.

2.4.3 Welfare Analysis of Year-dependent subsidies

In this subsection, I examine how year-dependent subsidies can improve welfare. I fix the total budget on college subsidies at the current level and examine how the utilitarian social welfare function improves by only varying the relative sizes of subsidies across college years.[18] The optimization problem is

$$\sum_j N_j \left(\int V_j(\boldsymbol{s}_j) d\bar{\mu}_j(\boldsymbol{s}_j) + \int V_j^c(\boldsymbol{s}_j^c) d\bar{\mu}_j(\boldsymbol{s}_j^c) \right) \tag{2.51}$$

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

	Status-quo	Optimal
$s_1(1)$	\$13,599	\$10,558
$s_1(2)$	\$11,447	\$8,887
$s_1(3)$	\$10,922	\$8,480
$s_2(1)$	\$13,599	\$20,259
$s_2(2)$	\$11,447	\$17,053
$s_2(3)$	\$10,922	\$16,271

Table 2.12: Year-dependent subsidies maximizing social welfare.

	Status-quo	Optimal
share of college enrollees	74.2%	73.9%
share of college graduates	32.9%	33.6%
skill premium	90.7%	86.9%
welfare gain		+0.17%

Table 2.13: The effect of the optimal year-dependent subsidies with correcting bias.

subject to

$$\int_{S_1^c} g_1 \bar{s}(q) d\mu_1^c + \int_{S_2^c} g_2 \bar{s}(q) d\mu_2^c = G_e \tag{2.52}$$

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

Table 2.12 displays the optimal college subsidies are back-loaded and the amount for the second half is twice the subsidy for the first half. The first and second rows of Table 3.3 show that enrollment decreases by 0.3 percentage points and the share of college graduates increases by 0.7 percentage points by switching to the optimal policy. The skill premium decreases by 3 percentage points.

To examine the welfare gain of the optimal policy, I use lifetime consumption equivalence as a summary measure of welfare. Let $\tilde{V}_j(\boldsymbol{c}, \boldsymbol{h}; \boldsymbol{s}_j)$ be expected lifetime

	Total	Level	Uncertainty	Inequality
Optimal	+0.07%	+0.17%	+0.04%	−0.11%

Table 2.14: Welfare decomposition.

utility at age j with the path of consumption $\boldsymbol{c}$, leisure $\boldsymbol{h}$ with the state $\boldsymbol{s}_j$ in which agents in the initial period has no optimism when calculating the lifetime value. Then lifetime consumption equivalence is defined as ω_{tot} such that

$$\sum_j N_j \int_{S_0} \tilde{V}_j(\boldsymbol{c}^B, \boldsymbol{h}^B; \boldsymbol{s}_j) d\mu_j^B = \sum_j N_j \int_{S_0} \tilde{V}_j((1+\omega_{tot})\boldsymbol{c}^A, \boldsymbol{h}^A; \boldsymbol{s}_j) d\mu_j^A \tag{2.53}$$

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

The fourth row of Table 3.3 shows that the lifetime consumption equivalence of the optimal year-dependent subsidies is 0.17%. We do not need to increase the labor income tax rate and the budget of the government to get the welfare gain.

To examine why welfare improves, I define the lifetime consumption equivalence only of the newborns in the economy and decompose it into three parts as in Benabou (2002): (i) *a level effect* which measures the gain in aggregate consumption, leisure, and college utility (ii) *an uncertainty effect* which measures the effect of the volatility of consumption and leisure paths across states and over time on utility of risk-averse agents, and (iii) *an inequality effect* which measures the distribution at the beginning of age 1.[19] I follow the decomposing process in Abbott et al. (2013).

Table 2.14 shows that the total welfare gain for newborns is 0.07%. There is a positive level effect of 0.17%. While output, capital, and consumption decrease (See the first three rows of Table 2.15), optimism is a key factor. In the current system, individuals are over-optimistic and there is an excessively large amount of college enrollees. The optimal back-loaded subsidies screen people who enroll and reduce

[19]The sum of these three effects might not necessarily be the total welfare effect.

	Status-quo	Optimal
Y	0.320	0.320
K	0.418	0.418
C	0.213	0.213
w^S	0.355	0.353
w^U	0.407	0.409
std c	0.130	0.129
std a	0.482	0.480
std h	0.0837	0.0835
std wage	0.546	0.542

Table 2.15: The aggregates under the optimal year-dependent subsidies.

enrollment. By reducing subsidies for the first half, the marginal enrollee with low ability stop enrolling, which reduces their college disutility. Optimism is a key factor for the optimal college subsidies. In the Appendix 2.E, I calibrate the case without optimism and examine how the assumption about optimism matters to the optimal policy.

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

The inequality effect is -0.11% as there is more inequality across heterogeneous agents at age 1 under the optimal policy. It is counter-intuitive because the difference in wages of skilled and unskilled labor decreases and the standard deviations of consumption, asset, hours, and wages per hour decreases as shown in Table 2.15. Although inequality as of period 1 increases, cross-section inequality in the economy decreases under the optimal policy.

In order to see why inequality at age 1 increases, I calculate the welfare gain for each ability and family income level in Table 2.16[20].

[20]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.

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

Table 2.16: Lifetime consumption equivalence variation for newborns.

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 3.3, the welfare of agents with low ability increases more than other agents.

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

Given the same ability level, the welfare gain is greater for high family income, which is consistent with the negative inequality effect[21]. College 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.

In order to see how people from poor family adjust to the loss of college subsidies for the first period, Table 2.17 shows the average change in each part of income and consumption for an individual with $\theta_h=0$, $q=1$, $\eta=\eta_H$, and $\phi=0$ at the first half of college. The loss of subsidies for the first half of college does not lead to the

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

	% of subsidy loss
Subsidies	-100%
Labor income	+24%
(Price of an hour of working)	+13%
(Leisure)	(-0.063)
Transfer from parents	+0.04%
Reducing Savings	+65%
Less tuition	+4%
Consumption	-7%

Table 2.17: Change in each item of income.

same amount of loss of consumption. The labor income increases and covers a 24% of the loss of college subsidies. First, they provide unskilled labor during the first half of college and the price of unskilled labor increases. Second, agents work for longer hours to mitigate the loss of college subsidies. As the fourth row shows, they cut their leisure by 0.063 out of the unit hour endowment. These results are consistent with the findings by Keane and Wolpin (1997) and Garriga and Keightley (2007).

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

2.4.4 Correcting Bias

A large part of the welfare gain of year-dependent subsidies originates in over-optimism of college ability. If the government can provide information to students to correct the bias on college ability before the enrollment decision, it can improve welfare and we might not need to rely on year-dependent subsidies. In this subsection, I show what is the welfare gain by correcting bias and compare it with the back-loaded subsidies

	Status-quo	Correcting bias	Optimal
share of college enrollees	74.3%	45.6%	45.9%
share of college graduates	32.2%	26.2%	26.1%
skill premium	88.6%	124%	125%
CEV		-9.06%	-8.99%

Table 2.18: Education and the skill premium when correcting bias.

without correcting bias.

The second column of Table 2.18 shows a welfare loss from correcting bias with the current subsidies. The enrollment rate drops significantly. Without optimism, enrollment is excessively small because there is a borrowing constraint and no insurance available about the risk of college ability. In the current system, while optimism leads to excessive enrollment, optimism also cancels out the effect of the tight credit limit or the abcense of insurance for the risk of college ability. In total, the loss from a large skill premium and excessively small enrollment is higher the gain from avoiding the excessive enrollment from optimism.[22]

To examine whether combining year-dependent subsidies with correcting bias, I solve the optimal policy problem in Section 2.4.3 without bias, that is $\mu_c(\theta_h) = 0$ for all θ_h. The solution is the second column of Table 2.19. The optimal subsidy is *front-loaded*: greater for the first period than for the second period. As shown in the third column of Table 2.18, by subsidizing college in the first period, the policy can increase enrollment. However, the welfare gain is still negative compared to the current case. Using the optimal back-loaded subsidies is more beneficial than correcting bias.

[22]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.

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

Table 2.19: Optimal subsidies with correcting bias

2.5 Conclusion

The skill premium has been expanding in the United States and policymakers often consider educational subsidies as a tool to increase college enrollment and decreasing inequality. However, enrollment does not necessarily lead to graduation and it is important to understand how policy can affect graduation. This paper quantitatively assesses the effects of year-dependent subsidies on enrollment, graduation, and the skill premium compared to year-invariant subsidies. Switching to back-loaded subsidies, with the total budget fixed, can increase the fraction of college graduates and reduce the skill premium more than the case with increasing year-invariant subsidies by 50%. Back-loaded subsidies improve welfare without increasing the total budget of college subsidies and increasing tax.

Although this paper focuses on college, the mechanics is applicable to other education levels such as higher education. Increasing subsidies for post-college might lead to an increase in workers with higher education, which affects the distribution of skill and wages. Changing the amount of subsidies by year before college might also have effects. Furthermore, 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 and age have potential to be an important policy tool.

Appendix 2.A Computation of Stationary Equilibrium

This section describes the method of computing an equilibrium.

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

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

   (a) Solve backward the Bellman equations for age $j = J, \ldots, j_b + 1$. The number of grids for assets is 30 and that for high school ability and college ability is 5[23]. The number of grids for college taste is 30. I apply the esndogenous grid method.

   (b) Given an initial guess of the value function of newborns V^0, solve backward the individual problems from $j = j_b, \cdots, 1$ for value functions and policy functions. 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, solve decision rules of individuals until convergence.

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

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

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

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

Appendix 2.B Pension

The average life time income is

$$\hat{y}(e,\theta) = \frac{\sum_{j=2}^{j_r-1} w^e \varepsilon_j^e(\theta,1)\bar{h}}{j_r - 2} \tag{2.54}$$

where $\bar{h} = 0.333$.

The pension formula is given by

$$p(e,\theta) = \begin{cases} s_1\hat{y}(e,\theta) & \text{for } \hat{y}(e,\theta) \in [0, b_1) \\ s_1 b_1 + s_2(\hat{y}(e,\theta) - b_1) & \text{for } \hat{y}(e,\theta) \in [b_1, b_2) \\ s_1 b_1 + s_2(b_2 - b_1) + s_3(\hat{y}(e,\theta) - b_2) & \text{for } \hat{y}(e,\theta) \in [b_2, b_3) \\ s_1 b_1 + s_2(b_2 - b_1) + s_3(b_3 - b_2) & \text{for } \hat{y}(e,\theta) \in [b_3, \infty) \end{cases} \tag{2.55}$$

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.

Appendix 2.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 use the SRC sample of heads whose age is between 25 and 63, which leads to 11,512 samples. I restrict observations to those with positive hours of labor 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

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

Table 2.20: Age profile estimates of each education.

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

(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 grade completed greater than 16. The estimation result is in Table 2.20. I take the average of the productivity of the corresponding two years for the productivity of j in the model and normalize the process so that the productivity at the first period after the education stage is unity.

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

Ability is approximated by the log of AFQT80 raw score. To estimate the coefficient on ability in effective labor, I use NLSY79 of 11,864 people. For the ability regression, I restrict samples aged between 25 and 63, which leads to 11,627 people. Since NLSY79 does not include old people, I rely on PSID to estimate the age effects.

After the age effect is filtered out, I regress hourly wages on ability for each education levels (HS, CD, and CG). As in the selection of PSID, I keep only people who do not report extreme changes of hourly wages (changes in log earnings larger than 4 or less than -2) or extreme hourly wages (less than $1 or larger than $400). I keep only people with 8 or more year observations, which leads to 3,851 people. I exclude enrolled students and hours worked per week less than 20. I also control dummies for each year.

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

Appendix 2.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.

Appendix 2.E Calibration without Optimism

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

$$\sigma_c(\theta_h) = \sigma_c \exp(\sigma_c^{\theta} \theta_h) \tag{2.56}$$

Table 2.21 displays the calibrated values under the specification without optimism. As you see, the intercept of the standard deviation of college ability σ_c is larger than

Parameter	Description	Value
λ	college utility intercept	-16.6
λ^{θ}	college utility slope	277
λ_1^{ϕ}	first period college taste	68.6
λ_2^{ϕ}	second half college taste	40.9
a^S	productivity of skilled labor	0.435
ϵ^{CD}	productivity of CD	0.984
σ_c	s.d. of college ability intercept	0.719
σ_c^{θ}	s.d. of college ability slope	0.155
κ	education cost	0.420
μ	consumption share of preference	0.421
β	time discount rate	0.931
υ	altruism	0.0632
d	lump-sum transfer ratio	0.130
ι	borrowing wedge ($r^- = r + \iota$)	18.7%
m	intergenerational ability transmission intercept	-0.0380
σ_h	intergenerational ability transmission s.d.	0.0761

Table 2.21: Parameters calibrated without optimism

the case with optimism. I need a high standard deviation and a high option value to match the high college dropout rate. The slope of the standard deviation is positive and the uncertainty of college ability is higher for higher high school ability.

Table 2.22 and Figures 2.4 and 2.5 display moments. The graduation rate of low high school ability is excessively high in this formulation without optimism. In order to match the high dropout rate of low ability people, the model needs a high standard deviation of college ability. Then too many people draw high college ability enough to stay and graduate. To match the low graduation rate of high ability people, the model needs high disutility for low ability people. Then enrollment decreases and college dropout also decreases. While the effect of increasing the standard deviation of college ability increases college enrollment, it also increases college disutility to match the high college dropout and decreases college enrollment at the same time, offsetting the first effect. Rather the best match requires a *low* standard deviation of college ability and low college disutility to match the high enrollment. To summarize,

Moment	Model	Target
Enrollment rate of ability quartile	(figure)	(figure)
Graduation rate of ability quartile	(figure)	(figure)
Enrollment rate of family income quartile	(figure)	(figure)
Graduation rate of family income quartile	(figure)	(figure)
Skill premium for CG[a]	90.4%	90.2%
Skill premium for CD	19.9%	19.9%
Education cost/mean income at 48	0.309	0.33
Hours of work	33.2%	33.3%
K/Y	1.240	1.325
Transfer/mean income at 48	67.0%	66%
log pre-tax/post-tax income	60.8%	61%
Borrowers	6.64%	6.3%
Mean of AFQT	0.0873	0
Standard deviation of AFQT	0.203	0.213

Table 2.22: Moments matched without optimism.

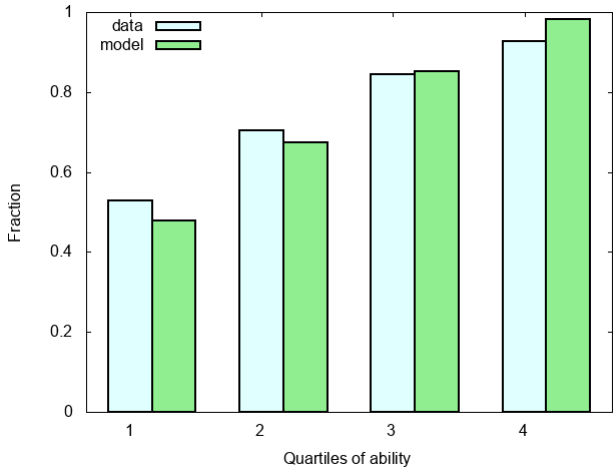
[a]The skill premiums are from full-time workers in Current Population Survey (CPS) IPUMS (Flood et al. (2018))

high option value without optimism does not explain the high dropout rate in the data as well as the model with optimism.

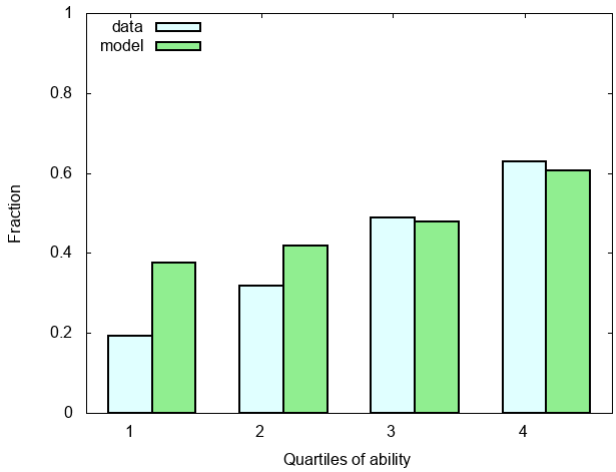
The optimal policy with this formulation is in Table 2.23. The optimal policy is now *front-loaded* and the subsidies for the first period is twice as much as for the second period. Without optimism, enrollment is excessively low in the status-quo as in the case with correcting bias. This implies that the assumption about optimism matters and one of the contributions of this paper is to calibrate this effect using

	Status-quo	Optimal
$s_1(1)$	\$13,599	\$14,152
$s_1(2)$	\$11,447	\$11,912
$s_1(3)$	\$10,922	\$11,366
$s_2(1)$	\$13,559	\$12,476
$s_2(2)$	\$11,447	\$10,502
$s_2(3)$	\$10,922	\$10,020

Table 2.23: Optimal subsidies without optimism.



(a) Enrollment rate

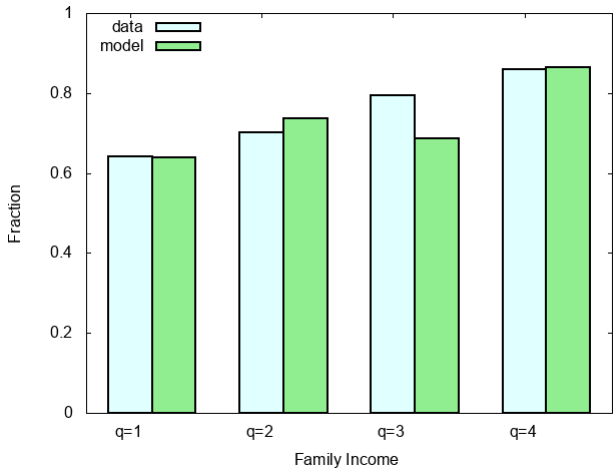


(b) Graduation rate

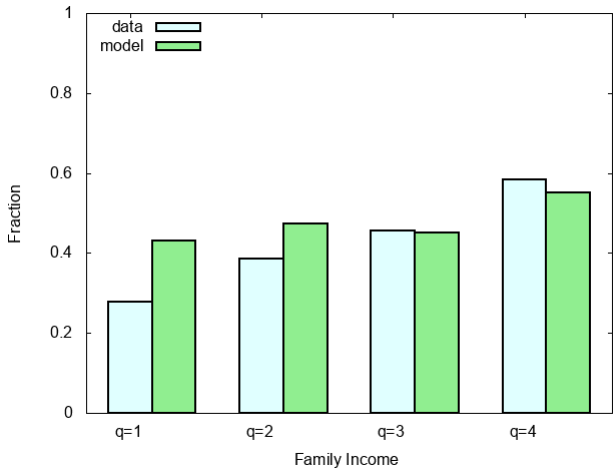
Figure 2.4: Model fit: enrollment and graduation rate for each ability quartile.

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

available data.



(a) Enrollment rate



(b) Graduation rate

Figure 2.5: Model fit: enrollment and graduation rates for each family income quartile.

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

Chapter 3

Progressive Taxation versus College Subsidies with College Dropout

3.1 Introduction

Wage inequality has been rising in the United States. Autor et al. (2008) document that the log real weekly wages by the 90th percentile of earners rise by approximately more than 55% relative to the 10th percentile earners. Wage differences by education and age increased substantially at the same time. In particular, the skill premium—the ratio of wages of college graduates to high school graduates—has risen from 1.5 in 1980 to 1.9 now. These increases in wage inequality contribute to an increase in consumption inequality (Cutler and Katz (1992)). It is important to understand how policy can alleviate wage inequality. There are two branches of policies that have been discussed in the public.

First, progressive labor income tax offers social insurance in many countries (Guvenen et al. (2013) for OECD countries) and reduces inequality in disposable income.

Second, college subsidies are considered as a policy to increase college enrollment and reduce the skill premium. A large literature (ex. Goldin and Katz (2009) 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. The question of this paper is which policy is better in terms of welfare and reducing inequality.

Although there is a long tradition of research on optimal tax and college subsidies, they abstract from college dropout. College dropout is an important factor when considering progressive income tax and college subsidies. In the United States, more than 50% of college enrollees drop put before earning a bachelor's degree and this fact draws attention to the policy circle. The existence of college dropout changes the effects of college subsidies on inequality as follows. In the current system, many people have already enrolled but drop out. It follows that increasing college subsidies just encourages people who are likely to drop out to enroll. Increasing college subsidies does not lead to an increase in the supply of college graduates and the skill premium might not decrease.

In this paper, I examine which policy is better for reducing inequality and social welfare, progressive labor income tax or college subsidies with college dropout. I build a life-cycle general equilibrium model using the same model as Matsuda (2019) where agents heterogeneous with regard to initial asset and high school ability—ability as of the graduation of high school—make enrollment and dropout decisions. Agents are assumed to be optimistic with regard to college ability before enrollment, which is consistent with the findings of Stinebrickner and Stinebrickner (2012) and Zafar (2011). Agents observe college ability only after enrollment and decide to drop out or not. Learning ability during college is a key factor of college dropout. The skill pre-

mium is determined by the aggregate skill in the economy with imperfect substitution between skilled and unskilled labor. With the model, I solve for the optimal progressivity of labor income tax and the optimal college subsidies maximizing a utilitarian social welfare function.

The main finding of this paper is the following. First, the optimal labor income tax is more progressive and the optimal college subsidies are higher than the current level. The optimal labor income tax is less progressive than the current one which is consistent with Krueger and Ludwig (2016). Higher college subsidies increase the number of college graduates, reduce the skill premium, and reduce consumption inequality. Second, college dropout lowers the effect of college subsidies. The welfare gain of college subsidies with college dropout is smaller than the case without college dropout. This implies that college dropout is a key factor for evaluating the welfare effect of college subsidies. Third, the optimal college subsidies are more powerful than the optimal progressive labor income tax even with college dropout.

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

3.1.1 Related Literature

The model and calibration is the same as Matsuda (2019). The paper focuses on the effect of college subsidies that vary across years in college but does not compare with progressive income tax.

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 college enrollees can experiment on the real value of enrollment. Stange (2012) and Trachter (2015) quantitatively examine the relevance of the option value of college dropout. Arcidiacono et al. (2016), Athreya and Eberly (2016), Lee et al. (2015), and Castex

(2017) examine the effect of introducing college dropout in the models. Stinebrickner and Stinebrickner (2008), Stinebrickner and Stinebrickner (2012), Stinebrickner and Stinebrickner (2014) show that learning academic ability during college is a key factor of college dropout and that credit constraints are not. Hendricks and Leukhina (2017) argue that the risk of college dropout given the information of ability before enrollment is small. Bound and Turner (2007) and Bound et al. (2010) examine the cause of the increase in college dropout in the United States.

Bovenberg and Jacobs (2005) theoretically derive the effect of subsidies. Abbott et al. (2013) examine the effect of college subsidies with parental transfers. Caucutt and Kumar (2003), Abbott et al. (2013), Ionescu (2011),Chatterjee and Ionescu (2012), and Hanushek et al. (2004) analyze the effect of various college aid policy with college dropout risk. Garriga and Keightley (2007) show the effect of an increase in subsidies on dropout decisions. The largest difference is that they do not compare the effect of college subsidies with progressive labor income tax.

The optimal tax and human capital investment subsidies has been analyzed in a dynamic public finance. For example, Stantcheva (2017) derives optimal income tax and human capital subsidies with risky human capital in a Mirrleesian style. In contrast, this paper derives optimal policy in a quantitative macroeconomic model in a Ramsey style. The closest paper to this paper is Krueger and Ludwig (2016) who 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. The main differnce from this paper is that they abstract from college dropout.

3.2 Model

The model is the same as in Matsuda (2019) and is composed of three main building blocks. The first is a model of endogenous enrollment and graduation decisions. Individuals observe their initial asset and high school ability at the first period and decide to enroll in college or not. Once an individual enrolls in college, she observes her college ability and makes a dropout decision. Utility from college attendance and returns to college graduation are determined by college ability. The second is an overlapping generations life cycle with no insurance markets. There is uncertainty with regard to labor productivity after they start working. In addition, the ability of parents and that of their children are correlated and parents choose how much wealth to transfer to their children. The third is imperfect substitution between skilled and unskilled labor. The skill premium is determined by aggregate supply of skilled and unskilled labor.

Since I focus on a stationary equilibrium in which the cross-sectional allocation and prices are constant, I do not include any time subscript in the description of the model.

3.2.1 Demography

There exists a continuum of over-lapping generations individuals each of whom has one offspring. Age is discrete and $j \in \{1, 2, \ldots, J\}$ and one period is two years. Individuals start independent economic decisions at the beginning of age 1 (biological age 18) as a high school graduate.

Figure 3.1 is the timeline. At the beginning of age 1, individuals make an enrollment decisions. Once they enroll, they choose whether to drop out or not at the beginning of age 2. If they choose not to drop out, they graduate from college at the end of age 2. After finishing schooling, each individual is categorized as one of the

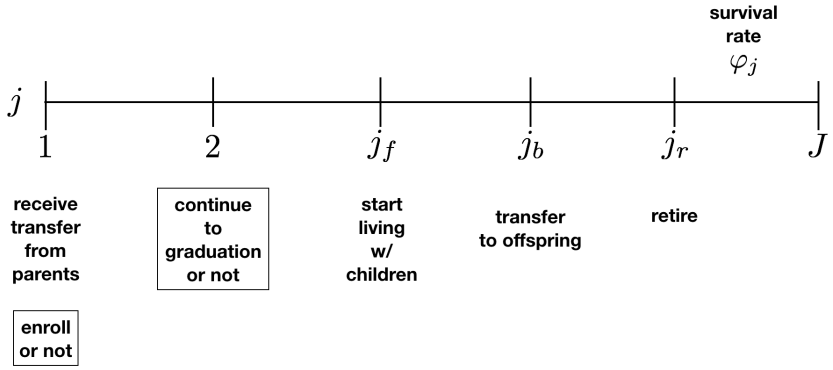


Figure 3.1: Timeline.

Source: Figure 1 in Matsuda (2019).

three education levels: a high school graduate ($e = HS$), a college dropout ($e = CD$), or a college graduate ($e = CD$). Individuals start living with their children at age j_f (biological age 30) and the children leave and become independent at their parents' age j_b (biological age 48). Parents are allowed to transfer wealth to their children only at this age. Individuals retire at age j_r (biological age 66) and the maximum age is J (biological age 100). Each individual at age j survives until the next period with probability $\varphi_j \in [0, 1]$. For simplicity, I assume $\varphi_j = 1$ for $j \in [1, j_r - 1]$ but the survival rate between j_r and $J - 1$ is from the US Life Tables 2000.

3.2.2 Preferences

At the beginning of age 1, individuals have preferences composed of the three additive parts.

1. The expected discounted sum of utility from consumption and leisure:

$$\mathbb{E}_1 \sum_{j=1}^{J} \tilde{\beta}^{j-1} u(c_j, \ell_j) \tag{3.1}$$

where

$$u\left(c,\ell\right)=\frac{\left(c^{\mu}\ell^{1-\mu}\right)^{1-\gamma}}{1-\gamma} \tag{3.2}$$

where c_j denotes consumption and ℓ_j is leisure at age j and β is the time discount rate.[1]. At age $j \in [j_f, j_b - 1]$, individuals discount consumption by $1+\zeta$ when they live with their children.

2. The expected utility of college attendance. The utility at age j is

$$\lambda_j(\theta_c,\phi)=\lambda+\lambda^{\theta}\theta_c+\lambda_j^{\phi}\phi \tag{3.3}$$

Individuals get this utility of college attendance only while in college. It depends on two factors: college ability θ_c and college taste ϕ. College ability θ_c is one of the key parameters in this model. If college ability is high, enrollees enjoy studying more and are more likely to enroll than enrollees with low ability. This is important to explain the observed positive correlation between ability and enrollment. However, within a category of ability and family income, there is heterogeneity in enrollment and graduation decisions. To explain this, this model has college taste ϕ unobservable to econometricians but observable for individuals in the model before enrollment.

3. Parents value their children's expected lifetime utility V_0 with a weight ν.

$$\beta^{j_b-1}\nu\mathbb{E}_1V_0 \tag{3.4}$$

[1] $\tilde{\beta}_j=\beta^j\left(\prod_{k=1}^{j}\varphi_k\right)$

3.2.3 Goods Sector

A representative firm produces final good from capital K and labor H. The production function is assumed to be Cobb-Douglas and

$$Y = F(K, H) = K^{\alpha} H^{1-\alpha} \tag{3.5}$$

where aggregate labor H is a CES function of skilled labor H^S and unskilled labor H^U.

$$H = (a^S (H^S)^{\rho} + (1 - a^S)(H^U)^{\rho})^{\frac{1}{\rho}} \tag{3.6}$$

where $\frac{1}{1-\rho}$ is the elasticity of substitution. The rent of capital and wages for skilled and unskilled labor are $r + \delta$, w^S, and w^U. I assume competitive markets for production inputs. The first order condition for profit maximization implies:

$$r = \alpha \left(\frac{K}{H}\right)^{\alpha - 1} - \delta \tag{3.7}$$

$$w^S = (1 - \alpha) a^S \left(\frac{K}{H}\right)^{\alpha} \left(\frac{H}{H^S}\right)^{1-\rho} \tag{3.8}$$

$$w^U = (1 - \alpha)(1 - a^S) \left(\frac{K}{H}\right)^{\alpha} \left(\frac{H}{H^U}\right)^{1-\rho} \tag{3.9}$$

I assume college graduates provide only skilled labor and that college dropouts and high school graduates provide only unskilled labor. The wages of each education are $w^{CG} = w^S$ and $w^{HS} = w^{CD} = w^U$.

Effective labor per hour is denoted by $\varepsilon_j^e(\theta, \eta)$. η is a mean-reverting labor productivity shock and follows an education-specific Markov chain $\pi^e(\eta'|\eta) > 0$. Π^e is its invariant distribution function. Effective labor depends on high school ability for high school graduates and college dropouts while it depends on college ability for college

graduates.

3.2.4 College

A representative college provides an individual with education of one period using κ units of skilled labor. The profit is

$$p_e E - w^S \kappa E \tag{3.10}$$

where E is the number of college enrollees and p_e is the price of education or tuition. I assume colleges are competitive and that entry is free, which implies that $p_e = w^S \kappa$.

3.2.5 Financial Markets

There is no insurance market against idiosyncratic risks but individuals have access to risk-free assets with interest r.

There is a cost of overseeing borrowers at the working stage $\iota > 0$ per unit of capital. With the non-arbitrage condition, the borrowing interest rate is $r^- = r + \iota$. In addition, there is a borrowing limit $\underline{A}^e$ for workers with education e. Retired individuals have no access to loans. The overseeing cost for college enrollees is $\iota + \iota^s$. Likewise, the interest rate to enrollees is $r^s = r + \iota + \iota^s = r^- + \iota^s$ and the borrowing limit for college enrollees is $\underline{A}^c_j$ for age j.

3.2.6 Individual Problems

There are three stages in one's lifecycle: education, working, and retirement. Individuals in the education stage can also work but I call the individuals not in college and before retirement "workers" and its stage "working stage".

Education Stage

Enrollment

After becoming economically independent, individuals make enrollment decisions. The value function V_0 is

$$V_0(a, \theta_h, \eta, q, \phi) = \max \left\{ \underbrace{V_1^c(a, \theta_h, \eta, q, \phi)}_{\text{enrolling}}, \underbrace{V_1(a, HS, \theta_h, \eta)}_{\text{not enrolling}} \right\} \tag{3.11}$$

The state variables are initial assets a, high school ability θ_h, an idiosyncratic transitory productivity η from Π^{HS}, parents' (family) income level q, and education taste ϕ.

In this model, there are two different ability: high school ability θ_h and college ability θ_c. Individuals observe their high school abilities before the enrollment decision but enrollees observe their college abilities at the beginning of age 2. College abilities are correlated with high school abilities and

$$\theta_c = \theta_h + \epsilon_c \text{ where } \epsilon_c \sim N(0, \sigma_c^2) \tag{3.12}$$

I assume that college enrollees have optimism about their college abilities. This assumption is to be consistent with an empirical finding of Stinebrickner and Stinebrickner (2012) that optimism is a key factor of college dropout. Given θ_h, the expectation before the enrollment decision is

$$\theta_c = \underbrace{\mu_c(\theta_h)}_{\text{bias}} + \underbrace{\theta_h + \epsilon_c}_{\text{actual ability}} \text{ where } \epsilon_c \sim N(0, \sigma_c^2) \tag{3.13}$$

where $\mu_c(\theta_h)$ is the bias. A positive $\mu_c(\theta_h)$ is associated with optimism about college ability. I also assume that the bias depends on high school ability $\mu_c(\theta_h) = \mu_{c0} + \mu_{c1}\theta_h$. The variance of the residual term is the same as the actual variance.

Initial wealth a is endogenous transfer from their parents. A high idiosyncratic productivity η implies that there is a good outside option of working if they do not choose to enroll. Family income level q affects college subsidies during college. The value V_1^c the value of entering the first period of college. V_1 is the value of entering the working stage as a high school graduate.

First half of college

The Bellman equation for the first period of college is

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

subject to

$$c + a' + p_e - s_1(q) = a + y - T(c,a,y) \tag{3.14}$$

$$y = w^{HS}\varepsilon_1^{HS}(\theta_h,\eta)h, \;\; a' \geq -\underline{A}_1^c \;\; c \geq 0, \;\; 0 \leq h \leq 1-\bar{h} \tag{3.15}$$

$$\theta_c = \theta_h + \mu_c(\theta_h) + \epsilon_c, \;\; \epsilon_c \sim N(0,\sigma_c^2) \text{ (perceived process)}, \;\; \eta' \sim \Pi^{CD} \tag{3.16}$$

College enrollees spend exogenous $\bar{h}$ of time studying, pay tuition p_e, receive subsidies $s_j(q)$ and derive utility of college attendance $\lambda_j(\theta,\phi)$. They can work as high school graduates during the first period of college. c is consumption, y is labor earnings and a' is the next period assets. The tax $T(c,a,y)$ depends on consumption, asset holdings, and earnings.

One draws a college ability θ_c and a new idiosyncratic productivity η' drawn from Π^{CD} at the end of the first period of college. Based on this, enrollees choose to stay in college for the value V_2^c or drop out for the value V_2.

All the student loan is refinanced into a bond with interest rate r^- for economizing

state variables where $\tilde{a}(a)$ is the transformation so that the total payment is the same[2].

Second half of college

The Bellman equation for the second half of college is

$$V_2^c(a,\theta_c,\eta,q,\phi) = \max_{c,h,a',y} u(c,1-h-\bar{h}) + \lambda_2(\theta_c,\phi) + \beta \mathbb{E}_{\eta'} V_3(\tilde{a}(a'),CG,\theta_c,\eta') \quad (3.17)$$

subject to

$$c + a' + p_e - s_2(q) - y + T(c,a,y) = \begin{cases} (1+r)a & \text{if } a \geq 0 \\ (1+r^s)a & \text{if } a < 0 \end{cases} \quad (3.18)$$

$$y = w^{CD}\varepsilon_2^{CD}(\theta_c,\eta)h, \;\; a' \geq -\underline{A}_2^c \;\; c \geq 0, \;\; 0 \leq h \leq 1-\bar{h}, \;\; \eta' \sim \Pi^{CG} \quad (3.19)$$

They can work as college dropouts. They finish college at the end of the second period and observe a new idiosyncratic productivity η' from Π^{CG} to get value V_j. Student loan is refinanced with transformation $\tilde{a}(a')$.

Working Stage

The Bellman equation of the working stage is

$$V_j(a,e,\theta,\eta) = \max_{c,h,a',y} u\left(\frac{c}{1+\mathbf{1}_{\mathcal{J}_f}\zeta}, 1-h\right) + \beta \mathbb{E}_{\eta'|\eta} V_{j+1}(a',e,\theta,\eta') \quad (3.20)$$

subject to

$$c + a' - y + T(c,a,y) = \begin{cases} (1+r)a & \text{if } a \geq 0 \\ (1+r^-)a & \text{if } a < 0 \end{cases} \quad (3.21)$$

$$y = w^e \varepsilon_j^e(\theta,\eta)h, \;\; a' \geq -\underline{A}^e \;\; c \geq 0, \;\; 0 \leq h \leq 1, \;\; \eta' \sim \pi^e(\cdot|\eta) \quad (3.22)$$

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

where $\mathbf{1}_{\mathcal{J}_f}$ is an indicator function for living with children ($j \in [j_f, j_b - 1]$). For high school graduates and college dropouts, ability θ is high school ability. For college graduates, ability is college ability. Idiosyncratic productivity η transitions with Markov chain π^e_η.

Transfer

At the beginning of age j_b, individuals choose the amount of transfer to their children b. The value at the beginning of age j_b is

$$V_{j_b}(a, e, \theta, \eta) = \mathbb{E}_{\theta'_h|\theta} \max_{b(\theta'_h)\in[0,a]} \{\tilde{V}_{j_b}(a-b, e, \theta, \theta'_h, \eta) + \nu \mathbb{E}_{\eta'',\phi} V_0(b, \theta'_h, \eta'', \tilde{q}(w^e \varepsilon^e_j(\theta, \eta)), \phi)\} \tag{3.23}$$

$$\eta'' \sim \Pi^{HS}, \;\; \phi \sim N(0,1), \;\; \theta_h \sim \pi_\theta(\cdot|\theta) \tag{3.24}$$

where

$$\tilde{V}_{j_b}(a, e, \theta, \theta'_h, \eta) = \max_{c,h,a',y} u(c, 1-h) + \beta \mathbb{E}_{\eta'|\eta} V_{j_b+1}(a', e, \theta, \eta') \tag{3.25}$$

subject to

$$c + a' - y + T(c, a, y) = \begin{cases} (1+r)a & \text{if } a \geq 0 \\ (1+r^-)a & \text{if } a < 0 \end{cases} \tag{3.26}$$

$$y = w^e \varepsilon^e_j(\theta, \eta) h, \;\; a' \geq -\underline{A}^e \;\; c \geq 0, \;\; 0 \leq h \leq 1 \tag{3.27}$$

Before the decisions of transfer, individuals observe their children's high school ability θ'_h with its density function $\pi_\theta(\theta'_h|\theta)$ but do not observe the children's initial productivity η'' drawn from Π^{HS} or taste ϕ from the normal distribution $N(0,1)$. The potential labor income $w^e \varepsilon^e_j(\theta, \eta)$ of the individuals affect the value of their children

through college subsidies.[3]

Retirement Stage

The Bellman equation after retirement of age j_r is

$$V_j(a, e, \theta) = \max_{c,a'} u(c, 1) + \beta\varphi_{j+1}V_{j+1}(a', e, \theta) \tag{3.29}$$

subject to

$$c + a' = (1 + r)\varphi_j^{-1}a + p(e, \theta) - T(c, \varphi_j^{-1}a, 0) \tag{3.30}$$

$$a' \geq 0 \;\; c \geq 0 \tag{3.31}$$

Retired individuals are assumed to provide no labor but receive interest payments of savings and retirement benefits $p(e, \theta)$. Retirement benefits depend on ability and education to capture the social security system of the United States (see Appendix 3.B). There is a perfect annuity market for savings of retired individuals so that the asset holdings at the next period is inflated by φ_j^{-1} (assets of expiring individuals are distributed within cohorts).

3.2.7 Government

The government spends the tax revenues on college subsidies G_e, other government consumption proportional to the aggregate output Y, $G_c = gY$, and retirement ben-

[3]The potential parental income is defined as labor earnings if they spend 35% working. There are three categories and

$$\tilde{q}(w^e\varepsilon_j^e(\theta, \eta)) = \begin{cases} 1 & \text{if } w^e\varepsilon_j^e(\theta, \eta) \times 0.35 \in [0, q_1] \\ 2 & \text{if } w^e\varepsilon_j^e(\theta, \eta) \times 0.35 \in [q_1, q_2] \\ 3 & \text{else} \end{cases} \tag{3.28}$$

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

efits. The total spending of college subsidies is

$$G_e = \sum_{j=1,2} \int_{S_j^c} s_j(q) d\mu_j^c \tag{3.32}$$

In the main exercises, I change the value of $s_j(q)$ for the optimal college subsidies.

The tax function is

$$T(c, a, y) = \tau_c c + \tau_k r a \mathbf{1}_{a \geq 0} + \tau_l y - d\frac{Y}{N} \tag{3.33}$$

The consumption tax is proportional with rate τ_c and the capital tax rate is τ_k. The government provides each individual with a lump-sum transfer $d\frac{Y}{N}$ where N is the measure of all the individuals in the economy. If the lump-sum transfer is greater, the tax system is more progressive. In the main exercises, I change the value of d for the optimal progressivity. τ_l is the proportional labor income tax rate.

3.2.8 Equilibrium

The model is overlapping generations and is numerically solved to characterize a stationary equilibrium. The cross-sectional allocation within each cohort j is invariant. In equilibrium, individuals maximize expected lifetime utility, a representative firm and a representative colleges maximize profits, the government budget is balanced for each period, and all the markets clear. Let $\boldsymbol{s}_j^c \in S_j^c$ denote the age-specific state vector for college enrollees and $\boldsymbol{s}_j \in S_j$ for workers and retirees and $\boldsymbol{s}_0 \in S_0$ for individuals at the beginning of age 1. The age-specific state vector for workers and retirees conditional on education e is defined as $\boldsymbol{s}_j^e \in S_j^e$.

Definition 3. *A stationary equilibrium is a list of value functions of workers and college enrollees* $\{V_j(\boldsymbol{s}_j), V_j^c(\boldsymbol{s}_j^c)\}$*, decision rules of enrollment* $d_1(\boldsymbol{s}_0)$ *and graduation* $d_2(\boldsymbol{s}_1^c)$*, decision rules of consumption, asset holdings, labor, output, parental*

transfers of workers $\{c_j(\boldsymbol{s}_j), a'_j(\boldsymbol{s}_j), h_j(\boldsymbol{s}_j), y_j(\boldsymbol{s}_j), b(\boldsymbol{s}_j)\}$*, decision rules of college enrollees* $\{c^c_j(\boldsymbol{s}^c_j), a'^c_j(\boldsymbol{s}^c_j), h^c_j(\boldsymbol{s}^c_j), y^c_j(\boldsymbol{s}^c_j)\}$*, aggregate enrollees, capital, and labor inputs* $\{E, K, H^S, H^U\}$*, prices* $\{r, w^S, w^U, p_e\}$*, policy* τ_ℓ*, measures* $\boldsymbol{\mu} = \{\mu^c_j(\boldsymbol{s}^c_j), \mu_j(\boldsymbol{s}_j), \mu^e_j(\boldsymbol{s}^e_j)\}$ *such that*

1. *Taking prices and policies as given, value functions* $\{V^c_j(\boldsymbol{s}^c_j), V_j(\boldsymbol{s}_j)\}$ *solve the individual Bellman equations and* $d_1(\boldsymbol{s}_0)$*,* $d_2(\boldsymbol{s}^c_1)$*,* $\{c_j(\boldsymbol{s}_j), a'_j(\boldsymbol{s}_j), h_j(\boldsymbol{s}_j), y_j(\boldsymbol{s}_j), b(\boldsymbol{s}_j)\}$*,* $\{c^c_j(\boldsymbol{s}^c_j), a'^c_j(\boldsymbol{s}^c_j), h^c_j(\boldsymbol{s}^c_j), y^c_j(\boldsymbol{s}^c_j)\}$ *are associated decision rules.*

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

3. *The government budget is balanced.*

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

where

$$G_c = gF(K, H) \tag{3.34}$$

$$G_e = \sum_{j=1,2} \int_{S^c_j} s_j(q) d\mu^c_j \tag{3.35}$$

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

$$H^S + \kappa E = H^{CG} \tag{3.36}$$

$$H^U = H^{HS} + H^{CD} \tag{3.37}$$

where

$$H^{CG} = \sum_{j=3}^{j_r-1} \int_{S_j^{CG}} \epsilon_j^{CG}(\theta,\eta) h_j(\boldsymbol{s}_j) d\mu_j^{CG} \tag{3.38}$$

$$H^{CD} = \sum_{j=2}^{j_r-1} \int_{S_j^{CD}} \epsilon_j^{CD}(\theta,\eta) h_j(\boldsymbol{s}_j) d\mu_j^{CD} + \int_{S_2^c} \epsilon_2^{CD}(\theta,\eta) h_2^c(\boldsymbol{s}_2^c) d\mu_2^c \tag{3.39}$$

$$H^{HS} = \sum_{j=1}^{j_r-1} \int_{S_j^{HS}} \epsilon_j^{HS}(\theta,\eta) h_j(\boldsymbol{s}_j) d\mu_j^{HS} + \int_{S_1^c} \epsilon_1^{HS}(\theta,\eta) h_1^c(\boldsymbol{s}_1^c) d\mu_1^c \tag{3.40}$$

and

$$K = \sum_{j=1}^{J} \int_{S_j} a_j'(\boldsymbol{s}_j) d\mu_j + \sum_{j=1,2} \int_{S_j^c} a_j'^c(\boldsymbol{s}_j^c) d\mu_j^c \tag{3.41}$$

$$E = \sum_{j=1,2} \int_{S_j^c} d\mu_j^c \tag{3.42}$$

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

3.2.9 Calibration

The calibration of this model is the same as Matsuda (2019). I will explain it in the detail in Appendix 3.C.

3.3 Results

In this chapter, I examine which is more effective in terms of reducing inequality and improving social welfare, progressive labor income tax or college subsidies.

	status-quo	optimal transfers	optimal subsidies
enrollment rate	74.2%	74.2%	78.6%
share of college graduates	32.9%	32.9%	35.6%
skill premium	90.7%	90.3%	75.0%
d	12.5%	13.1%	(12.5%)
$\bar{s}$	1	(1)	2.14
CEV		-0.05%	+1.50%

Table 3.1: The effect of an increase in labor income tax rate by 1%.

3.3.1 The Effect of an Increase in Labor Income Tax by 1%

First, I examine the effect of the changes in progressive tax and college subsidies corresponding to an increase in labor income tax rate τ_ℓ by 1 percentage point. By increasing τ_ℓ, tax revenue increases and the government can do either of the following: increasing lump-sum transfer or increasing college subsidies. I examine these two cases in turn.

The second column of Table 3.1 presents the case where the lump-sum transfer d increases. d determines the progressivity of the labor income tax. If d is high, low labor income earners receive more transfers than tax and it leads to a more progressive tax system. Financed by an increase in τ_ℓ by 1%, d increases 0.6 percentage points and the tax system is more progressive. By assumption, college subsidies are fixed. The increase in d does not affect the enrollment rate and the share of college graduates. The skill premium decreases because the increase in the lump-sum transfers increases leisure by the income effect. The income effect is greater for workers providing unskilled labor in this case and the supply of unskilled labor decreases more than the supply of skilled labor.

For normative analysis, I assess the effect of policy changes on the following social welfare function.

$$\sum_j N_j \left(\int V_j(\boldsymbol{s}_j) d\bar{\mu}_j(\boldsymbol{s}_j) + \int V_j^c(\boldsymbol{s}_j^c) d\bar{\mu}_j(\boldsymbol{s}_j^c) \right) \tag{3.43}$$

which takes the sum of the lifetime utility of all the existing individuals at the steady state. N_j is the relative population of age j calculated from survival rates φ_j. Let $\tilde{V}_j(\boldsymbol{c}, \boldsymbol{h}; \boldsymbol{s}_j)$ be expected lifetime utility at age j with the path of consumption $\boldsymbol{c}$, leisure $\boldsymbol{h}$ with the state $\boldsymbol{s}_j$. I assume the government implementing this optimal policy maximizes the lifetime utility without optimism about college ability. Then lifetime consumption equivalence is defined as ω_{tot} such that

$$\sum_j N_j \int_{S_0} \tilde{V}_j(\boldsymbol{c}^B, \boldsymbol{h}^B; \boldsymbol{s}_j) d\mu_j^B = \sum_j N_j \int_{S_0} \tilde{V}_j((1+\omega_{tot})\boldsymbol{c}^A, \boldsymbol{h}^A; \boldsymbol{s}_j) d\mu_j^A \tag{3.44}$$

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

This welfare measure decreases with the increase in d. and increasing transfers through an increase in the labor income tax is not beneficial. The loss of efficiency from increasing the labor income tax outweighs the gain from reducing inequality.

The third column of Table 3.1 presents the case where college subsidies increase. In the current state, college subsidies are denoted by $s_j(q)$ and now I assume college subsidies are $\bar{s}s_j(q)$, which implies $\bar{s} = 1$ in the current state. College subsidies are multiplied by 2.14 financed by an increase in τ_ℓ by 1%. The enrollment rate and the share of college graduates increase and the skill premium decreases. Lifetime consumption equivalence increases by 1.5% and increasing college subsidies are beneficial.

	status-quo	optimal tax	optimal subsidies
d	12.5%	11.1%	(12.5%)
$\bar{s}$	1	(1)	12.4
τ_ℓ	36.2%	34.1%	46.7%

Table 3.2: Optimal policy.

3.3.2 Optimal Progressivity versus Optimal College Subsidies

Second, I derive the optimal progressivity of the labor income tax given the current college subsidies. The optimal policy problem is defined as

$$\max_{d,\tau_\ell} W = \sum_j N_j \left(\int V_j(\boldsymbol{s}_j) d\bar{\mu}_j(\boldsymbol{s}_j) + \int V_j^c(\boldsymbol{s}_j^c) d\bar{\mu}_j(\boldsymbol{s}_j^c) \right) \tag{3.45}$$

subject to the government budget constraint. If you increase lump-sum transfer, then you have to adjust labor income tax rate τ_ℓ to balance the budget and this is the main tradeoff.

Next, I derive the optimal college subsidies given the current progressivity of labor income tax. The optimal policy problem is defined as

$$\max_{\bar{s},\tau_\ell} W = \sum_j N_j \left(\int V_j(\boldsymbol{s}_j) d\bar{\mu}_j(\boldsymbol{s}_j) + \int V_j^c(\boldsymbol{s}_j^c) d\bar{\mu}_j(\boldsymbol{s}_j^c) \right) \tag{3.46}$$

subject to the government budget constraint. If you increase college subsidies, then you have to adjust the labor income tax rate τ_ℓ to balance the budget.

The result of the optimal progressivity of labor income tax and college subsidies are in Table 3.2. The second column of the table shows the optimal tax is less progressive than the current policy. With this policy, the second column of Table 3.3 shows that the share of college enrollees increases by 0.2 percentage points and that the share of college graduates increases by 0.1 percentage points. Less transfers lead

	status-quo	optimal tax	optimal subsidies
enrollment rate	74.2%	74.4%	88.9%
share of college graduates	32.9%	33.0%	44.0%
skill premium	90.7%	91.3%	43.1%
Y	0.320	0.329	0.310
K	0.418	0.435	0.452
C	0.213	0.218	0.200
w^S	0.355	0.358	0.336
w^U	0.407	0.409	0.473
std c	0.130	0.134	0.110
std a	0.482	0.499	0.398
std h	0.0836	0.0812	0.111
std wage	0.546	0.551	0.545
CEV (consumption equivalence value)		+0.08%	+6.19%

Table 3.3: Aggregates.

to a decrease in the proportional labor income tax rate and increases the returns to education. While the share of skilled workers increase, the skill premium increases a little because unskilled workers decrease leisure more than skilled workers and the supply of effective unskilled labor increases. Due to an increase in the skill premium and a decreases in lump-sum transfers, the standard deviations of consumption, asset holdings, and wages increase. Since unskilled workers increase labor hours more than skilled workers, the standard deviation of hours decreases. The lifetime consumption equivalence with this policy increases by 0.10%.

The third column of Table 3.2 shows that multiplying the current college subsidies by 12 is optimal. With this policy, the second column of Table 3.3 shows that the share of college enrollees increases by 14.4 percentage points and the share of college graduates increases by 11.1 percentage points. With the increase in college subsidies, more people enroll and graduate and the skill premium decreases by 47.6 percentage points. The wage of college graduates decreases while the wages of college dropouts and high school graduates increase. The lifetime consumption equivalence increases by 6.18%.

	Total	Level	Uncertainty	Inequality
Optimal tax	+0.08%	+1.00%	−0.26%	−0.12%
Optimal subsidies	+19.2%	−4.66%	+21.1%	+1.20%

Table 3.4: Welfare decomposition.

To summarize, the welfare gain is higher for the optimal college subsidies than for the optimal progressive labor income tax. In addition, the optimal progressivity is less than the current progressivity.

As in Benabou (2002), I decompose the gain of lifetime consumption equivalence of newborns 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 the decomposing process in Abbott et al. (2013).

Table 3.4 shows the welfare gain for newborns and its decomposition. Note that the total welfare is different from Table 3.3 because this is the welfare gain only for newborns and does not include old people. As shown in Table 3.4, the level effect is significantly positive for the optimal progressive labor income tax because of less distortion of labor hours. In addition, the standard deviation of consumption increases in the optimal progressive labor income tax and increases uncertainty and inequality.

On the other hand, the case of the optimal subsidies is associated with a decrease in the level effect due to the high labor income tax. With the large decrease in the skill premium, inequality and uncertainty decrease, which is beneficial to the utilitarian social welfare measure.

The optimal mix of progressive income tax and college subsidies is very similar to the mix of the optimal progressive income tax and the optimal college subsidies. I show the results in Appendix.

Parameter	Description	Value
λ	college utility intercept	10.5
λ^{θ}	college utility slope	123
λ_1^{ϕ}	first period college taste	71.8
a^S	productivity of skilled labor	0.517

Table 3.5: Parameters re-calibrated for the case without learning ability.

3.3.3 Case with No Learning of College Ability

I examine how taking college dropout into account affects the optimal policy. The main reason why agents in the model drop out is learning college ability. In this subsection, I consider a hypothetical economy in which there is no optimism and no uncertainty of college ability, which implies $\mu_c(\theta_h) = 0$ for any θ_h and $\sigma_c^2 = 0$.

Given the values of the other parameters as the benchmark case, I re-calibrate college utility parameters λ, λ^{θ}, λ_1^{ϕ}, and productivity of skilled labor a^S to match the share of college graduates for the quartiles of ability and family income and the skill premium from the data (I assume $\lambda_2^{\phi} = \lambda_1^{\phi}$). The parameters are in Table 3.5. I solve again the same optimal problems for this economy and compare with the previous subsection.

Table 3.6 shows the optimal policy in this hypothetical economy. Although the optimal progressivity does not significantly change from the previous case, the optimal college subsidies are much higher than the previous case. In this economy, enrollment necessarily leads to graduation and college subsidies are a more powerful tool to increase college graduates and decrease the skill premium. The second column of Table 3.7 shows the effect of the same optimal college subsidies as the previous subsection. The share of college graduates increases and the skill premium decreases considerably more than the previous case. Accordingly, the welfare gain is also much greater. Including leaning ability in the model changes the effect of college subsidies on the share of college graduates, the skill premium, and welfare gain.

	status-quo	optimal tax	optimal subsidies
d	12.5%	11.6%	(12.5%)
$\bar{s}$	1	(1)	22.4
τ_ℓ	36.4%	36.4%	48.7%

Table 3.6: Optimal policy without learning ability.

	status-quo	prev optimal subsidies	optimal subsidies
share of college graduates	32.8%	53.3%	58.4%
skill premium	92.4%	18.7%	9.4%
CEV		+12.5%	+14.3%

Table 3.7: Aggregates without learning ability.

With the new optimal policy for this economy, college graduates increase by 25.8 percentage points and the skill premium decreases from 92.3% to 9.2%. The welfare gain is 14.4% and it has a large effect. This implies that the optimal college subsidies are smaller if we consider leaning of college ability in the model but college subsidies are still more powerful than progressive labor income tax with learning college ability.

3.4 Conclusion

This paper examines the effect of two policies on inequality and welfare: progressive labor income tax and college subsidies. College dropout is a key factor considering the welfare effect. With leaning of college ability and college dropout, the effect of college subsidies becomes small but still more effective than the optimal progressive labor income tax.

Appendix 3.A Computation of Equilibria

This section describes the method of computing equilibria, which is the same as Matsuda (2019).

1. Starting from an initial vector of aggregate variables $\boldsymbol{w} = \left(\frac{K}{H}, \frac{H^S}{H}, H, \tau_l, \bar{h}\right)$, 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.[4] The number of grids for college taste is 30. I apply the Endogenous grid method.

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

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

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

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

4. Starting from an initial measure μ_1 and given decision rules, solve forward from μ_1 to μ_J and update μ_1 until convergence.

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

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

Appendix 3.B Pension

The average life time income is

$$\hat{y}(e,\theta) = \frac{\sum_{j=2}^{j_r-1} w^e \varepsilon_j^e(\theta,1)\bar{h}}{j_r - 2} \tag{3.47}$$

where $\bar{h} = 0.333$ and the pension formula is given by

$$p(e,\theta) = \begin{cases} s_1\hat{y}(e,\theta) & \text{for } \hat{y}(e,\theta) \in [0, b_1) \\ s_1 b_1 + s_2(\hat{y}(e,\theta) - b_1) & \text{for } \hat{y}(e,\theta) \in [b_1, b_2) \\ s_1 b_1 + s_2(b_2 - b_1) + s_3(\hat{y}(e,\theta) - b_2) & \text{for } \hat{y}(e,\theta) \in [b_2, b_3) \\ s_1 b_1 + s_2(b_2 - b_1) + s_3(b_3 - b_2) & \text{for } \hat{y}(e,\theta) \in [b_3, \infty) \end{cases} \tag{3.48}$$

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.

Appendix 3.C Calibration

See Matsuda (2019) for the detail and the sample criteria. There are two sets of parameters: (1) those that are estimated outside of the model or fixed based on the literature and (2) the remaining parameters to match key moments given the first set of parameter values. Prices are normalized such that the average annual income of high school graduates at age 48 is \$51,933.

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

Table 3.8: Estimated ability slope ϵ_θ^e

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

Table 3.9: Estimated parameters of wage process

3.C.1 Labor Productivity Process

I assume labor productivity

$$\ln \epsilon_j^e(\theta, \eta) = \ln \epsilon^e + \ln \psi_j^e + \epsilon_\theta^e \theta + \ln \eta \tag{3.49}$$

where ψ_j^e is the age profile of workers at age j at education level e estimated from the PSID (See Appendix of Matsuda (2019)). The coefficients can vary across education levels. Ability is approximated by ln AFQT80 and I regress hourly wages on ability after filtering out the age effects. Table 3.8 shows the estimated coefficients on ability for each education level.

I assume $\pi^e(\eta'|\eta)$ is a Markov chain with two states η_H and η_L specific to each education level, which has exactly the same persistence and conditional variance as the AR(1) process:

$$\ln \eta' = \rho^e \ln \eta + \epsilon_\eta^e, \quad \epsilon_\eta^e \sim N(0, \sigma_\eta^{e2}) \tag{3.50}$$

After filtering out age effects, I employ a Minimum Distance Estimator and Table 3.9 is the estimated parameters.

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

Table 3.10: subsidies and family income

3.C.2 Intergenerational Ability Transmission

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

$$\theta'_h = m + m_\theta \theta + \epsilon_\theta, \;\; \epsilon_\theta \sim N(0, \sigma_h^2) \tag{3.51}$$

In order to estimate the conditional mean of inter-generational ability transmission, I regressed children's ability on parents' ability in NLSY79 to obtain the parameter 0.46.

3.C.3 Subsidies and Loans

I adopt Abbott et al. (2013) for the cost of college for enrollees and the subsidy system of the Unites States (see Table 3.10 for federal and state subsidies). The cost of college for each enrollee is $\$6,710$ and the education cost per student is \$17,489 in 2000. The difference $\$10,779$ is additional subsidies to enrollees in the model.

Students' interest rate is the prime rate plus 2.3% ($= \iota^s$, annual). The loan limit for the first period is set to be \$6,125 ($= \$2,625 + \$3,500$) and the loan limit for the second period is \$23,000 to imitate the Stafford loan limits. The borrowing limits for workers are based on self-reported maximimum amount of borrowing on unsecured credit by education level from the 2001 Survey of Consumer Finances.

Parameters	Interpretation	Value
γ	Coef of relative risk aversion	4
$\bar{h}$	Study time	0.25
ζ	Adult equivalence scale	0.3
α	Capital share	33.3%
δ	Depreciation (annual)	7.55%
ρ	Elasticity of substitution in production 1.41	0.2908
ι^s	Stafford interest premium (annual)	2.3%
$\underline{A}_1^c$	Borrowing constraint for 1st half (Stafford loan)	\$6,125
$\underline{A}_2^c$	Borrowing constraint for 2nd half (Stafford loan)	\$23,000
$\underline{A}^{HS}$	Borrowing constraint, HS (SCF)	\$17,000
$\underline{A}^{CD}$	Borrowing constraint, CD (SCF)	\$20,000
$\underline{A}^{CG}$	Borrowing constraint, CG (SCF)	\$34,000
τ_c	Consumption tax rate (McDaniel (2007))	7%
τ_k	Capital income tax rate (McDaniel (2007))	27%
g	Gov cons to GDP ratio	17.1%

Table 3.11: Parameters determined outside the model.

3.C.4 The Remaining Parameters

Given the parameter values set outside the model in Table 3.11, there are 16 remaining parameters: bias of expectation of college ability (μ_c^0, μ_c^1), college utility $(\lambda, \lambda^\theta, \lambda_1^\epsilon, \lambda_2^\epsilon)$, the variance of college ability σ_c, productivity of labor (a^S, ϵ^{CD}), education cost κ, utility parameters (μ, β, v), lump-sum transfer d, overseeing cost ι, and inter-generational ability parameters (m, σ_h).

I choose 27 moments in Table 3.13 and minimize the average Euclidean percentage deviation of the model from the data[5].

Optimism is the key driver of college dropouts. I try to match the difference between the expected and the actual graduation rates. 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 $43\%(= 0.86/0.60 - 1)$.

[5]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.

Parameter	Description	Value
μ_c^0	college ability bias intercept	0.190
μ_c^1	college ability bias slope	-0.413
λ	college utility intercept	-22.2
λ^θ	college utility slope	239
λ_1^ϕ	first period college taste	64.0
λ_2^ϕ	second half college taste	40.5
a^S	productivity of skilled labor	0.446
ϵ^{CD}	productivity of CD	1.03
σ_c	s.d. of college ability	0.341
κ	education cost	0.226
μ	consumption share of preference	0.419
β	time discount rate	0.939
v	altruism	0.0952
d	lump-sum transfer ratio	0.125
ι	borrowing wedge ($r^- = r + \iota$)	18.1%
m	intergenerational ability transmission intercept	-0.0471
σ_h	intergenerational ability transmission s.d.	0.171

Table 3.12: Parameters calibrated.

The third column of Table 3.12 presents the calibrated values. The calibrated value of μ_c is positive, which implies that enrollees are optimistic about their college ability. The model fit is in Table 3.13 and Figures 3.2 and 3.3.

Appendix 3.D Optimal Mix

The optimal policy problem is defined as

$$\max_{d,s,\tau_\ell} W = \sum_j N_j \left(\int V_j(\boldsymbol{s}_j) d\mu_j(\boldsymbol{s}_j) + \int V_j^c(\boldsymbol{s}_j^c) d\mu_j(\boldsymbol{s}_j^c) \right) \tag{3.52}$$

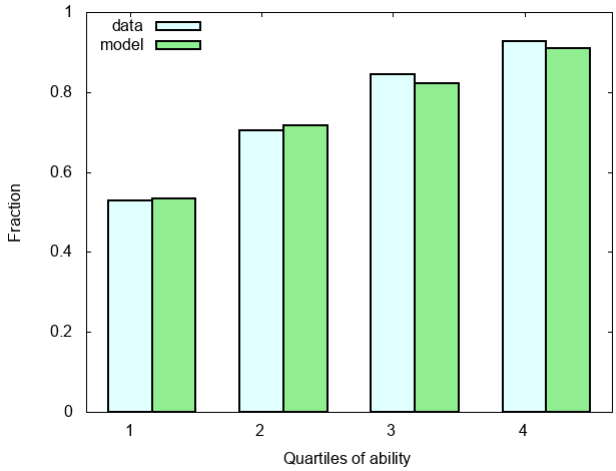
subject to the government budget constraint. Table 3.14 shows the optimal mix of progressivity of labor income tax and college subsidies.

Moment	Model	Target
Enrollment rate of ability quartile	(figure)	(figure)
Graduation rate of ability quartile	(figure)	(figure)
Enrollment rate of family income quartile	(figure)	(figure)
Graduation rate of family income quartile	(figure)	(figure)
Skill premium for CG	90.7%	90.2%
Skill premium for CD	19.7%	19.9%
Expected/Actual graduation rate -1	0.429	0.433
Education cost/mean income at 48	0.318	0.33
Hours of work	34.0%	33.3%
K/Y	1.305	1.325
Transfer/mean income at 48	67.1%	66%
log pre-tax/post-tax income	61.3%	61%
Borrowers	6.78%	6.3%
Mean of AFQT	-0.0135	0
Standard deviation of AFQT	0.217	0.213

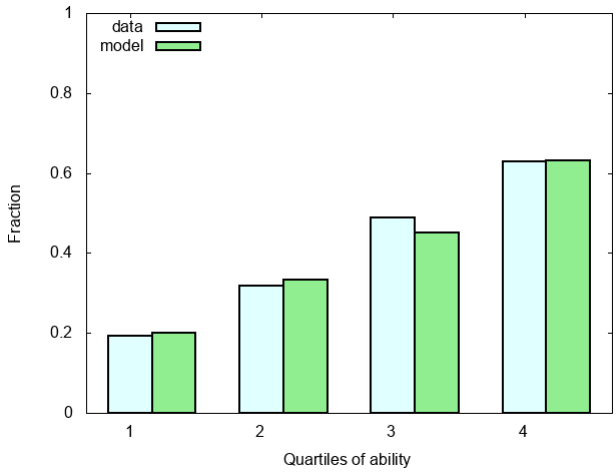
Table 3.13: Moments matched.

	status-quo	optimal mix
d	12.5%	8.91%
s ($\times$ benchmark)	1	12.9

Table 3.14: Optimal policy.

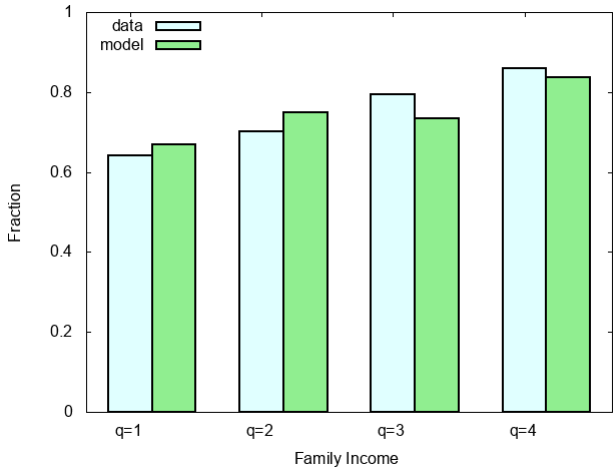


(a) Enrollment rate

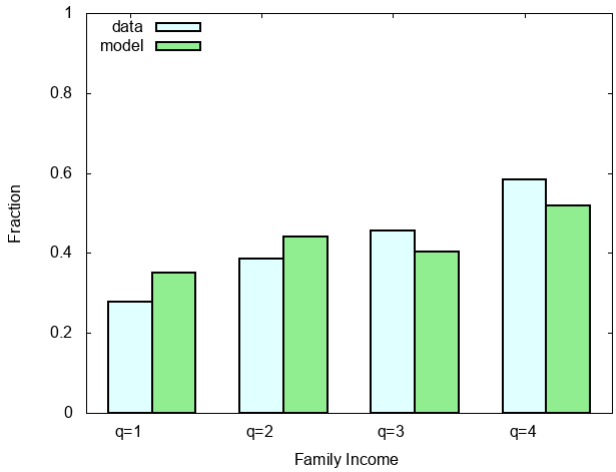


(b) Graduation rate

Figure 3.2: Model fit: enrollment and graduation rate for each ability quartile.



(a) Enrollment rate



(b) Graduation rate

Figure 3.3: Model fit: enrollment and graduation rates for each family income quartile.

Bibliography

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

Acemoglu, Daron, Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William Kerr. 2018. Innovation, reallocation, and growth. *American Economic Review* 108, no. 11:3450–91.

Acemoglu, Daron and David Autor. 2011. Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics*, vol. 4. Elsevier, 1043–1171.

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

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

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

Arcidiacono, Peter, Esteban Aucejo, Arnaud Maurel, and Tyler Ransom. 2016. Col-

lege attrition and the dynamics of information revelation. Tech. rep., National Bureau of Economic Research.

Athreya, Kartik and Janice Eberly. 2016. Risk, the college premium, and aggregate human capital investment. *Federal Reserve Bank of Richmond Working Paper* , no. 13-02R.

Autor, David H, Lawrence F Katz, and Melissa S Kearney. 2008. Trends in us wage inequality: Revising the revisionists. *The Review of economics and statistics* 90, no. 2:300–323.

Autor, David H, Frank Levy, and Richard J Murnane. 2003. The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics* 118, no. 4:1279–1333.

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

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

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

Bound, John and Sarah Turner. 2007. Cohort crowding: How resources affect collegiate attainment. *Journal of public Economics* 91, no. 5-6:877–899.

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

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

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

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

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

Castro, Rui and Daniele Coen-Pirani. 2016. Explaining the evolution of educational attainment in the united states. *American Economic Journal: Macroeconomics* 8, no. 3:77–112.

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

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

Cunha, Flavio. 2013. Investments in children when markets are incomplete. *Rev. Econ. Stud* .

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

Cutler, David M and Lawrence F Katz. 1992. Rising inequality? changes in the distribution of income and consumption in the 1980's. *The American Economic Review* 82, no. 2:546.

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

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2014a. The role of entrepreneurship in us job creation and economic dynamism. *Journal of Economic Perspectives* 28, no. 3:3–24.

———. 2014b. The secular decline in business dynamism in the us. *Unpublished draft, University of Maryland* .

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

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

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

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

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

Foster, Lucia, John Haltiwanger, and Cornell J Krizan. 2006. Market selection, reallocation, and restructuring in the us retail trade sector in the 1990s. *The Review of Economics and Statistics* 88, no. 4:748–758.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review* 98, no. 1:394–425.

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

Goldin, Claudia Dale and Lawrence F Katz. 2009. *The race between education and technology*. harvard university press.

Guvenen, Fatih, Burhanettin Kuruscu, and Serdar Ozkan. 2013. Taxation of human capital and wage inequality: A cross-country analysis. *Review of Economic Studies* 81, no. 2:818–850.

Haltiwanger, John, Ron S Jarmin, and Javier Miranda. 2013. Who creates jobs? small versus large versus young. *Review of Economics and Statistics* 95, no. 2:347–361.

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

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

Hathaway, Ian and Robert E Litan. 2014. Declining business dynamism in the united states: A look at states and metros. *Brookings Institution* :2.

Heckman, James J, Lance J Lochner, and Petra E Todd. 2006. Earnings functions,

rates of return and treatment effects: The mincer equation and beyond. *Handbook of the Economics of Education* 1:307–458.

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

Hopenhayn, Hugo and Richard Rogerson. 1993. Job turnover and policy evaluation: A general equilibrium analysis. *Journal of political Economy* 101, no. 5:915–938.

Hopenhayn, Hugo A. 1992. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society* :1127–1150.

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

Huergo, Elena and Jordi Jaumandreu. 2004. How does probability of innovation change with firm age? *Small Business Economics* 22, no. 3-4:193–207.

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

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

Jovanovic, Boyan. 1982. Selection and the evolution of industry. *Econometrica: Journal of the Econometric Society* :649–670.

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

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

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

Karahan, Fatih, Benjamin Pugsley, and Aysegül Sahin. 2015. Understanding the 30-year decline in the startup rate: A general equilibrium approach. *Unpublished manuscript, May* .

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

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

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

Krusell, Per, Lee E Ohanian, José-Víctor Ríos-Rull, and Giovanni L Violante. 2000. Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica* 68, no. 5:1029–1053.

Le, Anh T. 1999. Empirical studies of self-employment. *Journal of Economic surveys* 13, no. 4:381–416.

Lee, Sang Yoon, Yongseok Shin, and Donghoon Lee. 2015. The option value of human capital: Higher education and wage inequality. Tech. rep., National Bureau of Economic Research.

Lee, Yoonsoo and Toshihiko Mukoyama. 2008. Entry, exit and plant-level dynamics over the business cycle .

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

Matsuda, Kazushige. 2019. Optimal timing of college subsidies: Enrollment, graduation, and the skill premium .

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

Pugsley, Benjamin Wild and Ayegül ahin. 2018. Grown-up business cycles. *The Review of Financial Studies* 32, no. 3:1102–1147.

Schoar, Antoinette. 2010. The divide between subsistence and transformational entrepreneurship. *Innovation policy and the economy* 10, no. 1:57–81.

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

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

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

Stantcheva, Stefanie. 2017. Optimal taxation and human capital policies over the life cycle. *Journal of Political Economy* 125, no. 6:1931–1990.

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

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

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

Tauchen, George. 1986. Finite state markov-chain approximations to univariate and vector autoregressions. *Economics letters* 20, no. 2:177–181.

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

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

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