

What Drives the Skill Premium: Technological Change or Demographic Variation?

by Hui He,^{*}
Department of Economics
University of Hawaii at Manoa

Working Paper No. 09-11
March 2009

Abstract

This paper quantitatively examines the effects of two exogenous driving forces, investment-specific technological change (ISTC) and the demographic change known as “the baby boom and the baby bust,” on the evolution of the skill premium in the postwar U.S. economy. I develop an overlapping generations general equilibrium model with endogenous discrete schooling choice. The production technology features capital-skill complementarity as in Krusell et al. (2000). ISTC, through capital-skill complementarity, raises the relative demand for skilled labor, while demographic variation affects the skill premium through changing the age structure and hence relative supply of skilled labor. I find that demographic change is more important in shaping the skill premium before 1980. Since then, ISTC takes over to drive the dramatic increase in the skill premium.

Key Words: Skill Premium; Schooling Choice; Demographic and Technological Change; Capital-Skill Complementarity; Overlapping Generations

JEL Classification: E25; I21; J11; J24; J31; O33

^{*} Mailing Address: Department of Economics, University of Hawaii, 2424 Maile Way, Saunders Hall 542, Honolulu, HI 96822. E-mail: huihe@hawaii.edu.

What Drives the Skill Premium: Technological Change or Demographic Variation?

Hui He*

University of Hawaii

March 2009

Abstract: This paper quantitatively examines the effects of two exogenous driving forces, investment-specific technological change (ISTC) and the demographic change known as “the baby boom and the baby bust,” on the evolution of the skill premium in the postwar U.S. economy. I develop an overlapping generations general equilibrium model with endogenous discrete schooling choice. The production technology features capital-skill complementarity as in Krusell et al. (2000). ISTC, through capital-skill complementarity, raises the relative demand for skilled labor, while demographic variation affects the skill premium through changing the age structure and hence relative supply of skilled labor. I find that demographic change is more important in shaping the skill premium before 1980. Since then, ISTC takes over to drive the dramatic increase in the skill premium.

JEL classification: E25 (Aggregate factor income distribution); I21 (Analysis of education); J24 (Human capital); J31 (Wage differentials by skill); O33 (Technological change).

Key Words: Skill Premium; Schooling Choice; Capital-Skill Complementarity; Investment-Specific Technological Change; Demographic Change.

1 Introduction

The *skill premium*, which is defined as the ratio of skilled labor (workers holding college degrees) wage to unskilled labor (workers holding high school diplomas) wage, has gone through dramatic changes in the postwar U.S. economy. As Figure 1 (taken from Acemoglu (2003)) shows, starting from 1949 the evolution of the skill premium exhibited an “N” shape: it increased in the 1950s and 1960s, then decreased throughout the 1970s, and has increased dramatically since then. Meanwhile, as is also shown in the figure, the relative supply of skilled labor (the ratio of weeks worked by skilled labor to unskilled labor) has been increasing over time.

A number of researchers have asked why the pattern of the skill premium looks like it does. Popular explanations include investment-specific technological change through capital-skill complementarity (see

*Mailing Address: Department of Economics, University of Hawaii, 2424 Maile Way, Saunders Hall 542, Honolulu, HI 96822.
E-mail: huihe@hawaii.edu.

Krusell, Ohanian, Rios-Rull, and Violante (2000), hereafter KORV), international trade induced skill-biased technological change (Acemoglu (2003)), and skill-biased technological change associated with the computer revolution (Autor, Katz, and Krueger (1998)). Probably the most popular story is the one proposed in Katz and Murphy (1992). They claim that a simple supply and demand framework can explain the dynamics of the skill premium: “A smooth secular increase in the relative demand for college graduates combined with the observed fluctuations in the rate of growth of relative supply could potentially explain the movements in the college wage premium from 1963 to 1987.”¹ (Katz and Murphy (1992), page 50.) Table 1 (taken from Autor, Katz, and Krueger (1998)) demonstrates the basic idea. Since 1950 there has been steady growth in the relative demand for skilled labor,² but the growth rate of the relative supply of skilled labor has fluctuated. This growth rate was quite stable from 1940 to 1970 at around 2.5% per year, then increased dramatically to 4.99% per year during the 1970s, and it dropped back to the original average after 1980. Therefore, if we put the supply and demand changes together, we will see that the relative price of skilled labor, that is, the skill premium, dropped during the 1970s, (since supply exceeded demand,) while it increased in other decades.

Why did the relative supply of skilled labor increase dramatically during the 1970s? Katz and Murphy attribute this pattern to the baby boom. High fertility rates in the U.S. from 1946 to around 1960 resulted in a huge increase in the number of college graduates in the labor force since the late 1960s. In turn, the passage of the baby boom cohorts into mid-career, together with the accelerating skill-biased technological change in the 1980s, contributed to the rising college wage premium since 1980. In other words, the demographic change, together with the trend in skill-biased technological change, explains the dynamics of the skill premium.

But what is the source of the skill-biased technological change? KORV (2000) suggest an explanation. Modeling capital-skill complementarity in a neoclassical aggregate production function, they claim that the growth in the stock of capital equipment will complementarily increase the marginal product of skilled labor and hence raise its relative demand. They quantitatively evaluate how much this capital-skill complementarity has affected the skill premium from 1963 to 1992 and find that changes in observed factor inputs can account for most of the variation in the skill premium over these 30 years.

However, which driving force is more important in shaping the skill premium in the different periods remains an unexplored topic. This paper fill the void by asking a quantitative question: *to what extent can skill-biased technological change and demographic change account for the dynamics of the skill premium, respectively?*

The demographic change affects the skill premium mainly through changes in relative supply of skilled labor, while skill-biased technological change drives the skill premium via changes in relative demand for skilled labor. In order to evaluate the impacts from both sources, I have to endogenize relative supply and demand of skilled labor. Therefore, I develop a general equilibrium overlapping generations model

¹Goldin and Katz (2007) echoes this idea in explaining the whole twentieth century history of wage inequality.

²Katz and Murphy (1992) find that this secular growth in the relative demand for college graduates can be proxied by a linear time trend of 3.3% per year, which can be viewed as a proxy for skill-biased technological change (SBTC). But what drives this secular trend remains a “black box” in their paper and in subsequent work along this line. For example, Bound and Johnson (1992) also attribute much of the variation in the skill premium to a residual trend component that is interpreted as SBTC.

Period	Relative Wage (%)	Relative Supply (%)	Relative Demand (%)
1950-60	0.83	2.91	4.08
1960-70	0.69	2.55	3.52
1970-80	-0.74	4.99	3.95
1980-90	1.51	2.53	4.65
1990-98	0.36	2.25	2.76

Table 1: Growth of College/High School Relative Wage, Supply and Demand: 1950-1998 (Annualized Percent Changes)

with endogenous discrete schooling choice. The model includes three key features. First, with *ex-ante heterogeneity in the disutility cost of schooling*, individuals in each birth cohort (high school graduates) choose to go to college or not based on their expected future wage differentials, their forgone wages during the college years, their tuition payments, and their idiosyncratic disutility cost. This microfoundation gives us the standard features found in the human capital investment literature. (See, for instance, Ben-Porath (1967).) Second, the production technology has the feature of *capital-skill complementarity* as in KORV (2000), that is, capital is more complementary to skilled than unskilled labor. Third, following Greenwood, Hercowitz, and Krusell (1997, hereafter GHK), I assume existence of *investment-specific technological change*, i.e., a technological change on investment goods.

Under this theoretical framework, I calibrate the model and then quantitatively examine the effects of two exogenous forces, *investment-specific technological change* (ISTC) as described above and *demographic change* as represented by the growth rate of the cohort size of high school graduates, on the skill premium over the U.S. post-war period. I find that demographic change dwarfs ISTC before the late 1960s and accounts for about one-third of the decline in the skill premium in the 1970s. However, after the late 1970s ISTC takes over to drive the dramatic increase in the skill premium.

In this model, ISTC affects the skill premium through a simple economic mechanism.³ When ISTC speeds up, investment becomes increasingly efficient over time, so the relative price of the capital stock falls. This encourages higher investment and hence the capital stock increases. Due to capital-skill complementarity, an increase in the capital stock raises the *relative demand* for skilled labor, which raises the skill premium. However, in the general equilibrium framework of this model, rising skill premium will also induce forward-looking individuals more likely go to college and hence increase the relative supply of skilled labor, which tends to offset the increase in the skill premium.

Demographic change affects the skill premium through a different channel. Growing birth cohorts (“baby boom”) change the age structure in the economy and make it skew towards younger (college-aged population) cohorts. On the one hand, more people stay in college. Meanwhile more unskilled workers join the labor force. Therefore, the *relative supply* of skilled labor decreases, which tends to raise the skill premium. On

³Fernandez-Villaverde (2001) proposes a similar mechanism, but targets on a totally different question. He tries to answer why growth in per capita income has coexisted with a decline in fertility during the past 150 years. The idea is that capital-specific technological change brings more productive capital, which, through capital-skill complementarity, raises the skill premium (i.e., return to human capital). The increment in the skill premium induces parents to choose more education instead of more children, hence causing the decline in fertility.

the other hand, young people hold fewer assets over their life cycle, and thus the change in age structure also slows down asset accumulation. The decrease in capital stock, through capital-skill complementarity, tends to lower the skill premium.

Since both driving forces have opposite impacts on the skill premium, total effect of each force on the skill premium thus is ambiguous and has to be investigated quantitatively.

In quantitative terms, before the late 1960s the U.S. had undergone dramatic population growth (see Figure 2 and 3), while the change in ISTC was only moderate (see Figure 12). Demographic change outweighed technological change, and hence it dominated the impact on the skill premium. After the late 1970s, the magnitude of the baby bust was much smaller compared to the baby boom, while ISTC speeded up dramatically to become the major driving force.

This paper extends the existing literature on the effects of skill-biased technological change on wage inequalities. In comparison to KORV (2000), I endogenize the supply of skilled labor and put their aggregate production function into a dynamic general equilibrium setup. Therefore, the model is able to capture the dynamic interaction between the skill premium and relative supply of skilled labor. And only under this framework, we are able to quantify the relative importance of the two widely discussed driving forces, ISTC that was highlighted by KORV, and the demographic change that has been ignored by their paper, in shaping the dynamics of the skill premium.

In spirit, this paper is also close to Heckman, Lochner, and Taber (1998). They develop and estimate an overlapping generations general equilibrium model of labor earnings and skill formation with heterogeneous human capital. They test their framework by building into the model a baby boom in entry cohorts and an estimated time trend of increase in the skill bias of aggregate technology. They find that the model can explain the pattern of wage inequality since the early 1960s. However, they do not provide a micro-foundation about the source of skill-biased technological change as in this paper. They also do not ask the research question on quantitative decomposition of impacts of ISTC and demographic change on the skill premium.⁴

This paper extends He and Liu (2008), who provide a unified framework in which the dynamics of the relative supply of skilled labor and the skill premium arise as an equilibrium outcome driven by measured investment-specific technological change. This paper provides a microeconomic foundation for He and Liu (2008) by going deeper into the college choices that determine the supply of skilled labor. An overlapping generations framework is used here so that I can quantify the effect of demographic change, which is missing in He and Liu (2008).

The remainder of this paper is organized as follows. Section 2 documents some stylized facts about the dynamics of the cohort size of high school graduates, the college enrollment rate, and college tuition in the postwar U.S. economy. It also emphasizes linkages among these facts. Section 3 presents the economic

⁴Several other papers follow Heckman, Lochner, and Taber (1998) to emphasize the impact of skill-biased technological change on the skill premium. For example, Guvenen and Kuruscu (2006) present a tractable general equilibrium overlapping generations model of human capital accumulation that is consistent with several features of the evolution of the U.S. wage inequality from 1970 to 2000. Their work shares a similar micro-foundation of schooling choice as in this paper. But they do not have capital stock in the production technology, and hence no capital-skill complementarity. The only driving force in their paper is skill-biased technological change, which is calibrated to match the total rise in wage inequality in the U.S. data between 1969 and 1995.

model of college-going decisions, describes the market environment, and defines the general equilibrium in the model economy, thus laying out the theoretical foundation for the later data analysis and calibration exercise. Section 4 shows how to parameterize the model economy. Section 5 provides calibration results for the pre-1951 steady state. Section 6 computes the transition path of the model economy from 1951 to 2000 and compares the results with the data. It also conducts some counterfactual experiments to isolate the effects of investment-specific technological change and demographic change on the skill premium. Finally, Section 7 concludes.

2 Stylized Facts

This section summarizes the data pattern regarding the college choice which determines the supply of skilled labor. Figure 2 shows the cohort size of high school graduates. It was very stable before the early 1950s, then increased until 1976, and has decreased since then. Since the common age of high school graduation is around 18, we can view this graph as a 18-year lag version of U.S. fertility growth, that is, it reflects the baby boom and baby bust.⁵

Figure 3 measures the college-age population. I report the age 17-21 and 18-21 population in the U.S. since 1955. These series follow a similar pattern as in Figure 2. The baby boom pushed the college-aged population up until the fertility rate reached its peak around 1960, corresponding to the peak of the college-aged population around 1980. The baby bust then dragged the population size down.

The two figures above show changes in the population base of potential college students, but does the *proportion* of people going to college change over time? Figure 4 shows the college enrollment rate of recent high school graduates. It began growing in the early 1950s until 1968, when it started to decline; the entire 1970s was a depressed decade for college enrollment, and it was not until 1985 that the enrollment rate exceeded the level in 1968. Starting from 1980, the enrollment rate kept increasing for nearly 20 years. This pattern is also confirmed by other studies. (See Macunovich (1996) Figures 1.a, 1.b, 2.a, and 2.b, and Card and Lemieux (2000) Figure 3.)

Combining Figure 3 and 4, we can see the dramatic increase of the relative supply of skilled labor in the 1970s is due to a combination of accumulative rising college-age population and college enrollment rate in the 1960s. The demographic change only affects the *cohort size* of college-aged population (in this paper I restrict the college age to be 18-21 years old). Higher enrollment rate shifts the proportion of the college-age population into the skilled labor. The college enrollment rate thus is an important determinant of the relative supply of skilled labor.

By comparing the skill premium in Figure 1 and the college enrollment rate in Figure 4, one can see that they share a very similar pattern. This similarity implies a tight link between the college-going decision and the expected skill premium. The expected skill premium represents the expected gain from higher education. As the expected benefits increase, the enrollment rate increases. This finding motivates me to explicitly endogenize the college choice in the model. As Goldin and Katz (2007) point out, understanding changes in the relative supply of skilled labor is crucial in economic analysis of changes in wage structure

⁵Cohort size has been increasing again since 1995 because the baby boomers' children reached college age around the mid-1990s.

and returns to skill. In order to understand the evolution of skill premium, we cannot ignore the impact from the supply side.

To fully understand the determinants of schooling choice, we should also look at the cost side of college-going. In Figure 5 I report the real tuition, fees, room, and board (TFRB) per student charged by an average four-year institution (average means the enrollment weighted average of four-year public and private higher education institutions; see Appendix A for details). Again, we see a pattern similar to that of the skill premium and the college enrollment rate. TFRB increased over time except in the 1970s. Starting in 1980, real TFRB has raised dramatically.

The similarity is not surprising since it reflects supply and demand in the higher education market. Higher demand for skilled labor in the 1980s and 1990s pushed up the skill premium as we see in Figure 1. More people wanted to go to college, hence the enrollment rate increased as shown in Figure 4. In turn, higher demand for college education raised the price of college education, as shown in Figure 5.

The stylized facts relevant to this paper can be briefly summarized as follows:

1. The skill premium rose during the 1950s and 1960s, then fell from 1971 to 1979, and has increased dramatically since 1980.
2. The relative supply of skilled labor has increased since the 1940s.
3. The college enrollment rate exhibits a similar pattern as the skill premium, as do tuition payments.

The stylized facts about the skill premium observed in Figure 1 is the target of this paper. To answer the quantitative question raised in Section 1, I will take the demographic change in Figure 2 and the measured investment-specific technological change in Figure 12 as exogenously given, feed them into a dynamic general equilibrium model, and see what percentage of change in the skill premium can be explained by each of these two exogenous forces.

3 Model

In this section, I present the economic model that will be used later for calibration. It is a discrete-time overlapping generations (OLG) model. Individuals make the schooling choice in the first period. There is only one good in the economy that can be used in either consumption or investment.

3.1 Demographics

The economy is populated by overlapping generations. People enter the economy when they are 18 years old and finish high school, which I call the birth cohort and model as age $j = 1$. I assume people work up to age J , which is the maximum life span.⁶ To distinguish between the age of a cohort and the calendar time, I use j for the age, and t for the calendar time. For example, $N_{j,t}$ is the population size of the age- j cohort at time t .

⁶In an unreported experiment, I extend the current model to include retirement, social security, and lifetime uncertainty. The quantitative results are very similar to the ones presented here, while the lifecycle profiles of consumption and asset holdings are more realistic (hump-shaped). Please also refer to the sensitivity analysis in Section 6.3.

In every period t a new birth cohort enters the economy with cohort size $N_{1,t}$. It grows at rate n_t . Therefore, I have

$$N_{1,t} = (1 + n_t)N_{1,t-1}. \quad (1)$$

The fraction of the age- j cohort in the total population at time t is

$$\mu_{j,t} = \frac{N_{j,t}}{N_t} = \frac{N_{j,t}}{\sum_{i=1}^J N_{it}}. \quad (2)$$

This fraction will be used to calculate the aggregate quantities in the economy as cohort weights throughout the transition path.

The birth cohort in the model corresponds to the high school graduates (HSG) in Figure 2, and the growth rate of HSG cohort size is the data counterpart of n_t . Therefore, the “baby boom” corresponds to the 1951-1976 period when n_t increased over time, while the “baby bust” period is from 1976 to 1990 when n_t decreased over time.

3.2 Preferences

Individuals born at time t want to maximize their discounted lifetime utility

$$\sum_{j=1}^J \beta^{j-1} u(c_{j,t+j-1}).$$

The period utility function is assumed to take the CRRA form

$$u(c_{j,t+j-1}) = \frac{c_{j,t+j-1}^{1-\sigma}}{1-\sigma}. \quad (3)$$

The parameter σ is the coefficient of relative risk aversion, therefore $\frac{1}{\sigma}$ is the intertemporal elasticity of substitution. Since leisure does not enter into the utility function, each individual will supply all her labor endowment, which is normalized to be one.

3.3 Budget Constraints

An individual born at time t chooses whether or not to go to college at the beginning of the first period. I use $s \in \{c, h\}$ to indicate this choice. If an individual chooses $s = h$, she ends up with a high school diploma and goes on the job market to work as an unskilled worker up to age J , and earns high school graduate wage sequence $\{w_{j,t+j-1}^h\}_{j=1}^J$. Alternatively, she can choose $s = c$, spend the first four periods in college as a full-time student, and pay the tuition. I assume that an individual enters college will successfully graduate from college (there is no college dropout in the model). After college, she goes on the job market to find a job as a skilled worker and earns a college graduate wage sequence $\{w_{j,t+j-1}^c\}_{j=1}^J$. After the schooling choice, within each period, an individual makes consumption and asset accumulation decisions according to her choice.

For $s = c$, the budget constraints of the cohort born at time t are

$$c_{j,t+j-1} + tuition_{t+j-1} + a_{j,t+j-1} \leq (1 + r_{t+j-1})a_{j-1,t+j-2} \quad \forall j = 1, 2, 3, 4 \quad (4)$$

$$c_{j,t+j-1} + a_{j,t+j-1} \leq (1 + r_{t+j-1})a_{j-1,t+j-2} + w_{t+j-1}^c \varepsilon_j^c \quad \forall j = 5, \dots, J \quad (5)$$

$$c_{j,t+j-1} \geq 0, a_{0,t-1} = 0, a_{J,t+J-1} \geq 0,$$

where $\{\varepsilon_j^c\}_{j=5}^J$ is the age efficiency profile of college graduates. It represents the age profile of the average labor productivity for college graduates. Notice that individuals have zero initial wealth and cannot die in debt.⁷

For $s = h$, the budget constraints of the cohort born at time t are

$$\begin{aligned} c_{j,t+j-1} + a_{j,t+j-1} &\leq (1 + r_{t+j-1})a_{j-1,t+j-2} + w_{t+j-1}^h \varepsilon_j^h \quad \forall j = 1, \dots, J \\ c_{j,t+j-1} &\geq 0, a_{0,t-1} = 0, a_{J,t+J-1} \geq 0. \end{aligned} \quad (6)$$

Similarly, $\{\varepsilon_j^h\}_{j=1}^J$ is the age efficiency profile of high school graduates.

3.4 Schooling Choice

Next, I would like to explicitly model an individual's schooling choice. In order to generate a positive enrollment rate in the model, I need to introduce some *ex-ante* heterogeneity within each birth cohort. Without this within-cohort heterogeneity, the enrollment rate would be either zero or one.

Following Heckman, Lochner, and Taber (1998), I assume that different individuals within each birth cohort are endowed with different levels of disutility cost of schooling. I index people by their disutility level $i \in [0, 1]$, and the associated disutility cost that individual i bears is $DIS(i)$. I assume $DIS'(i) < 0$.⁸ The Cumulative Distribution Function (CDF) of disutility cost is denoted by F , $F(i_0) = \Pr(i \leq i_0)$. Now an individual i born at time t has her own expected discounted lifetime utility

$$\sum_{j=1}^J \beta^{j-1} u(c_{j,t+j-1}) - \mathcal{I}_i DIS(i), \quad (7)$$

where

$$\mathcal{I}_i = \begin{cases} 1 & \text{if } s_i = c \\ 0 & \text{if } s_i = h \end{cases},$$

subject to the conditional budget constraints (4)-(5) or (6), depending on individual i 's schooling choice s_i . Notice that idiosyncratic disutility cost $DIS(i)$ does not enter into the budget constraints, so everyone within the same cohort and with the same education status will have the same lifetime utility derived from physical consumption, which simplifies the computation. I use $UTIL_t^c$ to denote the discounted lifetime utility derived from people who are born at time t and choose to go to college ($s = c$) and $UTIL_t^h$ to denote the discounted lifetime utility derived from people who choose not to go to college ($s = h$). Therefore, $UTIL_t^c - UTIL_t^h$ represents the utility gain from attending college. Obviously, individual i will choose go to college if $DIS(i) < [UTIL_t^c - UTIL_t^h]$, will not go if $DIS(i) > [UTIL_t^c - UTIL_t^h]$, and is indifferent if $DIS(i) = [UTIL_t^c - UTIL_t^h]$.

It is easy to show that this model implies

$$UTIL_t^c - UTIL_t^h \stackrel{\geq}{\leq} 0 \text{ iff } NPV_t \stackrel{\geq}{\leq} 0,$$

⁷Notice that the model does not have exogenous (ad hoc) borrowing constraints. However, the standard properties of the utility function and the restriction that the agent cannot die in debt impose an endogenous (natural) borrowing constraint at every period.

⁸Navarro (2007) finds ability is the main determinant of this "psychic" cost and it plays a key role in determining schooling decisions. High ability individuals face very low disutility cost while low ability individuals face large disutility cost of attending college. Therefore, we can also view i as the index of individuals' "learning ability".

where

$$NPV_t = \sum_{j=5}^J \frac{w_{t+j-1}^c \varepsilon_j^c - w_{t+j-1}^h \varepsilon_j^h}{\prod_{i=2}^j (1 + r_{t+i-1})} - \sum_{j=1}^4 \frac{w_{t+j-1}^h \varepsilon_j^h}{\prod_{i=2}^j (1 + r_{t+i-1})} - \sum_{j=1}^4 \frac{tuition_{t+j-1}}{\prod_{i=2}^j (1 + r_{t+i-1})}. \quad (8)$$

Here NPV stands for the net present value of higher education. It consists of three terms. The first term represents the benefit of schooling, as college graduates can earn more through the skill premium. The second term represents the opportunity cost of schooling. It is the four-year forgone wage income for the college students. The third term is the present value of tuition paid during college, which represents the direct cost of schooling. From this representation it is very clear how the skill premium is going to affect an individual's schooling decision. Keeping other things equal, an increase in the skill premium will raise the benefit of schooling, thus raising NPV . A higher NPV will induce a higher utility gain from schooling $UTIL_t^c - UTIL_t^h$. If we assume that the distribution of disutility cost is stationary, a higher utility gain from schooling means it is more likely that $DIS(i) < [UTIL_t^c - UTIL_t^h]$, which implies that more people would like to go to college. This mechanism will generate the co-movement between skill premium and enrollment rate as observed in the data.

3.5 Production

I close the model by describing the production side of the economy. The representative firm in the economy uses capital stock (K), skilled labor (S), and unskilled labor (U) to produce a single good. Here skilled labor consists of college graduates, and unskilled workers are high school graduates. Following KORV (2000), I adopt an aggregate production function with capital-skill complementarity as follows:⁹

$$\begin{aligned} Y_t &= A_t F(K_t, S_t, U_t) \\ &= A_t [\mu U_t^\theta + (1 - \mu)(\lambda(B_t K_t)^\rho + (1 - \lambda)S_t^\rho)^{\frac{\theta}{\rho}}]^{1/\theta}, \end{aligned} \quad (9)$$

where A_t is the level of total factor productivity (TFP), B_t is the level of capital productivity and represents capital-embodied technological change, $0 < \lambda, \mu < 1$, and $\rho, \theta < 1$. This production technology is of constant returns to scale. The elasticity of substitution between the capital-skilled labor combination and unskilled labor is $\frac{1}{1-\theta}$ and the one between capital and skilled labor is $\frac{1}{1-\rho}$. For the capital-skill complementarity, we require $\frac{1}{1-\rho} < \frac{1}{1-\theta}$, which means $\rho < \theta$.

The difference between my production function and the one in KORV (2000) is that I do not distinguish between structures and equipment, so the capital K in my model is just the total capital stock.

The representative firm rents capital, skilled labor, and unskilled labor from households at the rates r_t , w_t^c , and w_t^h . Its profit maximization implies the first-order conditions

$$r_t = \lambda(1 - \mu)A_t B_t^\rho H_t (\lambda(B_t K_t)^\rho + (1 - \lambda)S_t^\rho)^{\frac{\theta}{\rho}-1} K_t^{\rho-1} - \delta, \quad (10)$$

$$w_t^c = (1 - \mu)(1 - \lambda)A_t H_t (\lambda(B_t K_t)^\rho + (1 - \lambda)S_t^\rho)^{\frac{\theta}{\rho}-1} S_t^{\rho-1}, \quad (11)$$

$$w_t^h = \mu A_t H_t U_t^{\theta-1}, \quad (12)$$

$$\text{where } H_t = [\mu U_t^\theta + (1 - \mu)(\lambda(B_t K_t)^\rho + (1 - \lambda)S_t^\rho)^{\frac{\theta}{\rho}}]^{1/\theta-1}.$$

⁹Griliches (1969) provides evidence from U.S. manufacturing industries data that skill is more complementary with capital than unskilled labor. Duffy et al. (2004) show that there is some empirical support for the capital-skill complementarity hypothesis by using a macro panel set of 73 countries over the period 1965-1990.

δ is the capital depreciation rate. Dividing (11) by (12), I derive the expression for the skill premium:

$$\frac{w_t^c}{w_t^h} = \frac{(1-\mu)(1-\lambda)}{\mu} \left[\lambda \left(\frac{B_t K_t}{S_t} \right)^\rho + (1-\lambda) \right]^{\frac{\theta-\rho}{\rho}} \left[\frac{S_t}{U_t} \right]^{\theta-1}. \quad (13)$$

Log-linearizing (13), differentiating it with respect to time, and using “hat” to denote the rate of change ($\hat{X} = \frac{\dot{X}}{X}$), I obtain (ignoring time subscripts for convenience)

$$\left(\frac{w^c}{w^h} \right) \approx \lambda(\theta - \rho) \left(\frac{BK}{S} \right)^\rho [\hat{B} + \hat{K} - \hat{S}] + (\theta - 1) [\hat{S} - \hat{U}]. \quad (14)$$

This equation is exactly the same as in KORV (2000) except for the B term. It says that the growth rate of the skill premium is determined by two components. One is the growth rate of the relative supply of skilled labor $[\hat{S} - \hat{U}]$. Since $\theta < 1$, relatively faster growth of skilled labor will reduce the skill premium. This term is called “*relative quantity effect*” in KORV (2000). The other term $\lambda(\theta - \rho) \left(\frac{BK}{S} \right)^\rho [\hat{B} + \hat{K} - \hat{S}]$ is called the “*capital-skill complementarity effect*”. If capital grows faster than skilled labor, this term will raise the skill premium due to $\rho < \theta$. The dynamics of the skill premium depend on the magnitudes of these two effects.

The law of motion for the capital stock in this economy is expressed as

$$K_{t+1} = (1 - \delta)K_t + X_t q_t,$$

where X_t denotes capital investment. Following GHK (1997), I interpret q_t as the current state of the technology for producing capital, hence changes in q represent the notion of investment-specific technological change (ISTC). When q increases, investment becomes increasingly efficient over time.

To simplify computation, I follow Fernandez-Villaverde (2001) and map ISTC into the changes in the capital productivity level B_t .¹⁰ Therefore, increases in q_t will be transformed into increases in B_t . As shown in equation (14), when B_t increases, through the capital-skill complementarity effect, it will raise the skill premium. ISTC thus is also skill-biased.¹¹

Finally, the resource constraint in the economy is given by

$$C_t + TUITION_t + X_t = Y_t,$$

where C_t is total consumption and $TUITION_t$ is the total tuition payment.

3.6 The Recursive Competitive Equilibrium

The model above is a standard OLG setting with discrete schooling choices. And I assume that individuals have perfect foresight about the paths of exogenous changes $\{n_t\}$ and $\{q_t\}$.¹² Suppose an individual i born

¹⁰The transformation is $B_t = 1 + (1 - (1 - \delta)^{t-1}) \left(\frac{q_t}{q_1} - 1 \right)$. I normalize $q_1 = 1$ in initial steady state. Please refer to Fernandez-Villaverde (2001) for details.

¹¹More accurately, it can be easily shown that an economy with investment-specific technological change q , but without capital-embodied technological change B (which is my benchmark economy), can be equivalent to another economy with capital-embodied technological change B , but without ISTC in terms of allocation. (See GHK (1997) for details.) Since changes in B_t will increase the skill premium, due to the equivalence the ISTC in my benchmark economy has the same effect.

¹²Perfect foresight assumption is quite common in this type of research. McGrattan and Ohanian (2008) use this assumption and conduct deterministic simulations to study the macroeconomic impact of fiscal shocks during World War II. Chen et al. (2006) take the actual time path of TFP growth rate to investigate its impact on post-war Japanese saving rate. Their sensitivity analysis shows that alternative expectations hypotheses do not change the quantitative results significantly.

at time t has already made the schooling decision $s_{i,t}$. Conditional on this choice, I can present her utility maximization problem in terms of a dynamic programming representation.

For $s_{i,t} = c$, let $V_{t+j-1}^c(a_{j-1,t+j-2}, j)$ denote the value function of an age- j individual with asset holding $a_{j-1,t+j-2}$ at beginning of time $t+j-1$. It is given as the solution to the dynamic problem

$$V_{t+j-1}^c(a_{j-1,t+j-2}, j) = \max_{\{c_{j,t+j-1}, a_{j,t+j-1}\}} \{u(c_{j,t+j-1}) + \beta V_{t+j}^c(a_{j,t+j-1}, j+1)\} \quad (15)$$

subject to (4)-(5).

For $s_{i,t} = h$, the corresponding value function is

$$V_{t+j-1}^h(a_{j-1,t+j-2}, j) = \max_{\{c_{j,t+j-1}, a_{j,t+j-1}\}} \{u(c_{j,t+j-1}) + \beta V_{t+j}^h(a_{j,t+j-1}, j+1)\} \quad (16)$$

subject to (6).

Individuals solve their perfect foresight dynamic problem by using backward induction. Back to age 1, an individual with disutility index i will choose $s_{i,t}$ based on the criterion below

$$\begin{aligned} s_{i,t} &= c && \text{if } V_t^c(a_{0,t-1} = 0, 1) - DIS(i) > V_t^h(a_{0,t-1} = 0, 1), \\ s_{i,t} &= h && \text{if } V_t^c(a_{0,t-1} = 0, 1) - DIS(i) < V_t^h(a_{0,t-1} = 0, 1), \\ s_{i,t} &= \text{indifferent} && \text{if } V_t^c(a_{0,t-1} = 0, 1) - DIS(i) = V_t^h(a_{0,t-1} = 0, 1). \end{aligned} \quad (17)$$

Based on the individuals' dynamic program and the schooling choice criterion above, the definition of the competitive equilibrium in this model economy is standard.

Definition 1 Let $\mathcal{A} = \{a : -b \leq a \leq a_{\max}\}$, $\mathcal{S} = \{h, c\}$, $\mathcal{J} = \{1, 2, \dots, J\}$, $\mathcal{D} = [0, 1]$, and $\mathcal{T} = \{1, 2, \dots, T\}$. Given the age structure $\{\{\mu_{j,t}\}_{j=1}^J\}_{t=1}^T$, a Recursive Competitive Equilibrium is a sequence of individual value functions $V_t^s : \mathcal{A} \times \mathcal{J} \rightarrow R$; individual consumption decision rules $C_t^s : \mathcal{A} \times \mathcal{J} \rightarrow R_+$; individual saving decision rules $A_t^s : \mathcal{A} \times \mathcal{J} \rightarrow \mathcal{A}$ for $s \in \mathcal{S}$ and $t \in \mathcal{T}$; an individual i 's period 1 schooling choice $s_{i,t}^*$ for $s \in \mathcal{S}$, $i \in \mathcal{D}$, and $t \in \mathcal{T}$; an allocation of capital and labor (skilled and unskilled) inputs $\{K_t, S_t, U_t\}_{t=1}^T$ for the firm; a price system $\{w_t^c, w_t^h, r_t\}_{t=1}^T$; and a sequence of measures of individual distribution over age and assets $\lambda_t^s : \mathcal{A} \times \mathcal{J} \rightarrow R_+$ for $s \in \mathcal{S}$ and $t \in \mathcal{T}$ such that:

1. Given prices $\{w_t^c, w_t^h, r_t\}$, the individual decision rules C_t^s and A_t^s solve the individual dynamic problems (15) and (16).
2. Optimal schooling choice $s_{i,t}^*$ is the solution to the schooling choice criterion in (17) for each individual i .
3. Prices $\{w_t^c, w_t^h, r_t\}$ are the solutions to the firm's profit maximization problem (10)-(12).
4. The time-variant age-dependent distribution of individuals choosing s follows the law of motion

$$\lambda_{t+1}^s(a', j+1) = \sum_{a:a' \in A_t^s(a, j)} \lambda_t^s(a, j). \quad (18)$$

5. Individual and aggregate behaviors are consistent

$$K_t = \sum_j \sum_a \sum_s \mu_{j,t} \lambda_t^s(a, j) A_t^s(a, j - 1), \quad (19)$$

$$S_t = \sum_j \sum_a \mu_{j,t} \lambda_t^c(a, j) \varepsilon_j^c, \quad (20)$$

$$U_t = \sum_j \sum_a \mu_{j,t} \lambda_t^h(a, j) \varepsilon_j^h. \quad (21)$$

6. Goods market clears

$$\sum_{j=1}^J \sum_a \sum_s \mu_{j,t} \lambda_t^s(a, j) C_t^s(a, j) + \sum_{j=1}^4 \sum_a \mu_{j,t} \lambda_t^c(a, j) tuition_{j,t} + X_t = Y_t, \quad (22)$$

or

$$C_t + TUITION_t + X_t = Y_t.$$

When ISTC and demographic change both stabilize at some constant levels, that is, $q_t = q$ and $n_t = n, \forall t$, the economy reaches a steady state. In such a steady state, the age structure, the distribution of individuals over assets and age, and the individual decision rules are all age-dependant but time-invariant. Therefore, I can define the stationary competitive equilibrium accordingly.

4 Parameterization

In this section, I calibrate the model economy to replicate certain properties of the U.S. economy in the pre-1951 initial steady state. More specifically, my strategy is to choose parameter values to match *on average* features of the U.S. economy from 1947 to 1951.¹³

4.1 Data Work of Cohort-Specific Skill Premium

The skill premium data I report in Figure 1 is the average skill premium across all age groups in a specific year. However, since the model presented here is a cohort-based OLG model, each cohort's college-going decision is based on this cohort's specific lifetime skill premium profile. For example, for the cohort born at time t , the lifetime cohort-specific skill premium is $\left\{ \frac{w_{t+j}^c}{w_{t+j}^h} \right\}_{j=1}^J$. In order to understand the mechanism of schooling decision for each cohort, I need to find the data counterpart of this cohort-specific skill premium.

I use March CPS data from 1962 to 2003, plus 1950 and 1960 Census data to construct the cohort-specific skill premium profiles for the 1948-1991 cohorts. (I choose to end the sample in 1991 due to the data quality—the 1991 cohort only has 12 year HSG wage and 8 year CG wage data.) In order to make my results comparable to the literature, I follow Eckstein and Nagypál (2004) in restricting the data (please refer to their paper for the details). The sample includes all full-time full-year (FTFY) workers between ages 18 and 65. To be consistent with the model, I only look at high school graduates (HSG) and college graduates (CG). The wage is the annualized real wage (in terms of 2002 U.S. dollars). In Figure 6, I show the mean

¹³I choose the U.S. economy from 1947 to 1951 as the initial steady state based on the observations that both the ISTC and demographic changes were quite stable for this time period.

CG and HSG wages (top panel) and the skill premium (bottom panel) which is the ratio between these two means for the sample period 1949-2002. It is similar to the pattern of the skill premium shown in Figure 1, which includes post-college graduates in the skilled labor group. However, the decline during the 1970s is flatter and the magnitude of the increase since 1980 is smaller. This is because the wage of post-college graduates increases even faster than CG.¹⁴ Including them (as Autor and Katz (1999) do) further widens the wage gap between skilled and unskilled labor.

Since CPS is not a panel data set, theoretically speaking I cannot track specific cohorts from it. However, since it is a repeated cross-sectional data set, I can use a so-called “synthetic cohort construction method” to construct a proxy of a cohort’s specific skill premium. For example, the 1962 cohort’s (18-year-old HSG in 1962) lifetime (18-65 years old) HSG wage profile $\{w_{1961+j}^h \varepsilon_j^h\}_{j=1}^{48}$ is constructed as follows: I take the 18-year-old HSGs in 1962 and calculate their mean wage. Then I take the 19-year-old HSGs in 1963 and calculate their mean wage. Next, the 20-year-old HSGs in 1964, the 21-year-old HSGs in 1965, and so on, up to the 58-year-old HSGs in 2002. Later I will show how to predict the mean HSG wage after age 58 to complete the lifecycle wage profile for this cohort.

I perform a similar procedure to construct the 1962 cohort’s CG wage profile $\{w_{1961+j}^c \varepsilon_j^c\}_{j=1}^{48}$. But I start from 1966 because if someone from the 1962 cohort chose to go to college, she would spend four years in college, graduate in 1966, and start to earn CG wages thereafter. Therefore, I take the 22-year-old CGs in 1966, calculate their mean wage, and then follow the above procedure again.

Using this method repeatedly for each birth cohort, I have the original data sequences of cohort-specific HSG and CG lifetime wage profiles for the 1948-1991 cohorts. However, due to the time range of the CPS data, some data points are missing for a complete lifetime profile for every cohort. For example, some cohorts are missing at the late age data points (cohorts after 1962) and some are missing at the early age data points (e.g., cohorts 1948-1961). I use econometric method to predict the mean wage at those specific age points and interpolate the missing data. I predict them by either second- or third- order polynomials, or a conditional Mincer equation as follows

$$\begin{aligned} \log[\text{HSGwage}(\text{age})] &= \beta_0^h + \beta_1^h \text{experience}_h + \beta_2^h \text{experience}_h^2 + \varepsilon^h, \text{ experience}_h = \text{age}-18 \\ \log[\text{CGwage}(\text{age})] &= \beta_0^c + \beta_1^c \text{experience}_c + \beta_2^c \text{experience}_c^2 + \varepsilon^c, \text{ experience}_c = \text{age}-22. \end{aligned}$$

The criterion is basically the goodness of fit. I check with the neighborhood cohorts to make sure the predicted value is reasonable. The “rule of thumb” of a hump-shaped profile also applies here to help make choices. As an example, in Figures 7 and 8 I show the prediction for the 1955 cohort by using the trendlines of second- and third- order polynomials. Obviously, the third-order one fits the data better, therefore, I use it to predict the missing data points for this cohort.

4.2 Distribution of Disutility Cost

The distribution of disutility cost $DIS(i)$ becomes very crucial in the computation because it is this distribution that determines the enrollment rate and hence the relative supply of skilled labor in the model. The problem is how to obtain it.

¹⁴Eckstein and Nagypál (2004) document this fact. Please refer to their paper for more details.

The schooling choice criterion embodied in (17) actually sheds some light on how to compute the distribution of disutility cost. Note that the person i^* who is indifferent between going to college or not will have

$$V_t^c(a_{0,t-1} = 0, 1) - DIS(i^*) = V_t^h(a_{0,t-1} = 0, 1),$$

that is, her disutility cost is exactly the difference between two conditional value functions. Since disutility cost is a decreasing function of index i , people with disutility index $i > i^*$ will go to college. Therefore, for a specific cohort t , if we calculate the difference between two conditional value functions $V_t^c(a_{0,t-1} = 0, 1) - V_t^h(a_{0,t-1} = 0, 1)$, we obtain the cut-off disutility cost for this cohort. If we also know the enrollment rate of this cohort, it tells us the proportion of people in this cohort who have less disutility than i^* at that specific cut-off point of disutility cost. In this way, I can pin down one point on the CDF of disutility cost. Applying this procedure to different cohorts will give me a picture of how disutility cost is distributed.¹⁵

Fortunately, I have cohort-specific lifetime wage profile data from 1948 to 1991 as described in Section 4.1. I set the interest rate equal 3%, the discount parameter β to 1.03, and the preference parameter σ to 1.5.¹⁶ For each cohort born at time t , I normalize the 18-year-old HSG wage (which is $w_t^h \varepsilon_1^h$ in the model) to one and feed in the cohort-specific lifetime wage profiles from the data. I go through the backward induction of the Bellman equation as described in Section 3.6 to obtain the value function difference $V_t^c(a_{0,t-1} = 0, 1) - V_t^h(a_{0,t-1} = 0, 1)$ for every cohort t . In Figure 9, by plotting them against enrollment rate data in the same time range, I have 44 points on the possible CDF of the disutility cost. By assuming disutility costs follow a normal distribution, I then estimate the CDF function.¹⁷ Later in the computation of the stationary equilibrium and transition path, during each iteration when I obtain factor prices $\{w_t^c, w_t^h, r_t\}$, I can conduct backward induction of Bellman equations loop to obtain the conditional value functions. Feeding the difference between these two functions into the estimated CDF, I get the corresponding enrollment rate.

4.3 Demographic

The model period is one year. Agents enter the model at age 18 ($j = 1$), work up to age 65 ($J = 48$), and die thereafter.

The growth rate of cohort size n is calculated as the average growth rate of the HSG cohort size from 1948 to 1951, which is 0%.

4.4 Preferences and Endowments

I pick CRRA coefficient $\sigma = 1.5$, which is in the reasonable range between 1 and 5 and is widely used in the literature (e.g., Gourinchas and Parker (2002)).

¹⁵Here I assume the distribution of disutility cost is stationary.

¹⁶Since I obtain the disutility cost from the partial equilibrium computation given the preference parameter σ , the discount rate β , and the interest rate r , later when I calibrate the model to match pre-1951 steady state, I need to make sure these three values are consistent with those used in the general equilibrium. The values I give here are consistent with those I obtain later in calibrating the initial steady state.

¹⁷Heckman, Lochner, and Taber (1998) also assume the “nonpecuniary benefit of attending college” is normally distributed. A more flexible Beta distribution yields a very similar estimated CDF as normal distribution within the reasonable range of disutility cost.

Parameter	Description	Value and Source
J	maximum life span	48, corresponding to age 65 in real life
$\{\varepsilon_j^s\}_{j=1}^J, s = c, h$	age efficiency profiles	1962-2003 CPS, and 1950, 1960 Census
σ	CRRRA coefficient	1.5, Gourinchas and Parker (2002)
θ	elasticity b/w U and K	0.401, KORV (2000)
ρ	elasticity b/w S and K	-0.495, KORV (2000)
δ	depreciation rate	0.069, İmrohorođlu, İmrohorođlu, and Joines (1999)

Table 2: Parameter Values from Outside Sources

The age efficiency profile of high school graduates $\{\varepsilon_j^h\}_{j=1}^J$ and college graduates $\{\varepsilon_j^c\}_{j=1}^J$ are calculated as follows: from the 1962-2003 CPS and the 1950 and 1960 Census data I calculate the mean HSG and CG wages across all ages for the time period 1949-2002, then I obtain the mean HSG and CG wages in the same time period for each age group. Thus, the age efficiency profiles are expressed as

$$\varepsilon_j^h = \frac{HSGwage_j}{HSGwage}, \varepsilon_j^c = \frac{CGwage_j}{CGwage}, \forall j = 1, \dots, 48.$$

The result is shown in Figure 10. Both profiles exhibit a clear hump shape and reach a peak around age 55. Also notice that $\varepsilon_j^c = 0, \forall j = 1, \dots, 4$, since I assume CGs never work during study.

4.5 Technology

Two key elasticity parameters in the production function, the coefficient for elasticity of substitution between capital and skilled labor $\rho = -0.495$ and the coefficient for elasticity of substitution between unskilled labor and the capital-skilled labor combination $\theta = 0.401$, are taken directly from KORV (2000). This implies the elasticity of substitution between capital and skilled labor is 0.67 and the one between unskilled and skilled labor is 1.67. Capital-skill complementarity is satisfied.

In the initial steady state, both TFP level A and capital productivity B are normalized to unity. I set the depreciation rate of capital δ to 0.069 by following İmrohorođlu, İmrohorođlu, and Joines (1999), who calculate this parameter from the annual U.S. data since 1954.

Table 2 summarizes the choices of parameter values from outside sources.

This leaves four parameter values to be calibrated. The subjective discount rate β is set equal to 1.03 to replicate the target capital-output ratio 2.67, which is the average value from 1947 to 1951.¹⁸ The income share of capital in the capital-skilled labor combination $\lambda = 0.645$ is chosen to match the income share of capital in NIPA for the 1947-1951 period. The income share of unskilled labor $\mu = 0.418$ is chosen to match the average skill premium in the 1949 Census data (1.4556). The scale factor of the disutility cost sd (see Appendix B for detail) is chosen to match the average enrollment rate between 1947 and 1951. Table 3 summarizes the discussion above.

The computation method of the steady state is described in detail in Appendix B.

¹⁸It is plausible for a subjective discount factor greater than one in an overlapping generations setting. See İmrohorođlu et al. (1995) for a detailed discussion.

Parameter	Value	Match the moment
β	1.03	$\frac{K}{Y} = 2.67$
λ	0.645	Capital income share= 27.57%
μ	0.418	Skill premium= 1.4556
sd	3.10	enrollment rate= 41.54%

Table 3: Parameter Values from Calibration

Variable	Model	Data
$\frac{w^c}{w^h}$	1.4541 (construction)	1.4556 (1949 Census data)
e	41.51% (construction)	41.54% (1947-51 average)
K/Y	2.77 (construction)	2.67 (1947-51 average)
$(w^hU + w^cS)/Y$	72.47% (construction)	72.43% (1947-51 average)
S/U	67.67%	66.28% (1949 Census data)
C/Y	80.33%	79.57% (1947-51 average)
X/Y	19.09%	20.17% (1947-51 average)
r	3%	
$UTIL^c$	-100.2247	
$UTIL^h$	-157.1638	

Table 4: Macro Aggregates in the Benchmark Economy: Initial Steady State

5 Steady State Results

In this section, I report the numerical simulations for the stationary equilibrium of the benchmark economy and compare the results with the pre-1951 U.S. data. The macro aggregates that the model generates are shown in Table 4.

The simulations show that the model does well in matching the data. It matches our targets—skill premium, enrollment rate, capital-output ratio (K/Y), and labor income-output ratio ($\frac{w^cS+w^hU}{Y}$) by construction. Additionally, several key macro aggregate ratios such as consumption-output ratio (C/Y) and investment-output ratio (X/Y) are also in line with the U.S. average data. The model also matches the relative supply of skilled labor very well. The risk-free real interest rate is 3%. Average CG enjoys higher lifetime utility than average HSG because a CG has higher consumption over the lifecycle. We can see this clearly in the lifecycle profiles below.

5.1 Lifecycle Profiles

This model also generates the lifecycle profiles for CG and HSG, respectively. Figure 11 shows the lifecycle profiles of wealth accumulation, consumption, and income for CG and HSG. Panel A shows that since CGs have no income in the first four periods, they have to borrow to pay for tuition and consumption. Therefore, they accumulate negative wealth over the first four periods. After graduating from college, they start to earn the CG wage and are able to pay the loans. By age 34, CGs pay back all the loans borrowed from

n (%)	w^c/w^h	S/U (%)	BK/S	$UTIL^c$	$UTIL^h$
0 (benchmark)	1.4541	67.67	5.72	-100.2	-157.2
4.06	1.5085	61.54	5.43	-100.9	-156.8
-1.57	1.4400	69.55	5.83	-100.0	-157.1

Table 5: Effect of Population Growth on Steady State

previous years and begin to accumulate positive wealth, reaching a peak around their mid-50s. At that time they begin to dissave. There is no bequest motive in this model, therefore people die with zero assets remaining. The same hump shape is also observed for HSG, except that they accumulate positive assets from the beginning.

The lifecycle profile of consumption in Panel B is worthy of explanation. It keeps increasing until the deaths of both CG and HSG. The reason is clear when we look at the intertemporal Euler equation derived from the model

$$\left(\frac{c_{j+1}}{c_j}\right)^\sigma = \beta(1+r). \quad (23)$$

Given $\beta = 1.03$ and $r = 0.03$ as in the results, $\beta(1+r) = 1.061$. Therefore, the right-hand side of equation (23) is larger than one, inducing a positive growth rate of consumption over the life cycle.

Although CGs do not have any income during the first four periods, they have higher consumption than HSGs at any age. This is because in this deterministic model, the consumption path is determined by permanent income, and CGs have higher discounted lifetime income.

Finally, panel C shows the hump-shaped lifetime labor income profiles for CG and HSG, which are affected by the hump-shaped CG and HSG age efficiency profiles $\{\varepsilon_j^c\}_{j=1}^J$ and $\{\varepsilon_j^h\}_{j=1}^J$.

5.2 Comparative Static Experiments

Based on the steady state results, I carry out some comparative static exercises to study the effects of the growth rate of cohort size by changing n and the effects of investment-specific technological change by changing q . I summarize the corresponding results in Tables 5 and 6, respectively. In Table 5, 0% is the average growth rate of the HSG cohort size from 1947 to 1951, which is our benchmark case, 4.06% is the average growth rate of the HSG cohort size from 1952 to 1976, the “baby boom” period, and -1.57% is the average growth rate from 1977 to 1991, the period when n_t continuously decreased. The results show that as the growth rate of the HSG cohort size increases, the skill premium also increases, and vice versa.

Why does the increase in the HSG cohort size cause an increase in the skill premium? The intuition is as follows: an increase in n will change the age structure $\{\mu_j\}_{j=1}^J$ in the economy, making it skew towards younger cohorts. Keeping the enrollment rate unchanged, more people from the college-aged cohort stay in college. Meanwhile, more people from the college-aged cohort also join the labor force as unskilled labor. This results in relatively less *out-of-school* skilled labor in the current labor market, as is shown in Table 5. When n increases to around 4%, the relative supply of skilled labor S/U decreases by 9.1%. This change tends to raise the relative price of skilled labor, which is the skill premium, through the relative quantity effect. However, a change in age structure also has an impact on asset accumulation. Recall that the lifecycle profile of asset holdings for CG and HSG in Figure 11. People accumulate fewer assets during early

B	w^c/w^h	S/U (%)	BK/S	$UTIL^c$	$UTIL^h$
1 (benchmark)	1.4541	67.67	5.72	-100.2	-157.2
3.28 (2000 level)	1.8516	86.00	18.61	-88.8	-125.3

Table 6: Effect of Investment-Specific Technological Change on Steady State

working years. A shift towards younger cohorts in the demographic structure thus decreases the incentive of asset accumulation in the economy. As a result, the capital-output ratio (K/Y) decreases from 2.77 in the benchmark case to 2.57 in the $n = 4.06\%$ case. It also leads to a decrease in effective capital-skilled labor ratio (BK/S). Then, through the capital-skill complementarity effect, it tends to decrease the skill premium. Quantitatively, the impact of the demographic change on the relative supply of skilled labor dominates that on the relative demand of skilled labor through capital-skill complementarity. Thus, the skill premium increases. On the other hand, a decrease in n will make the age structure favor the older cohort, and hence will increase the relative supply of skilled labor and raise the incentive for asset accumulation. These two impacts again tend to offset each other. Quantitatively, a change in n from 0% to -1.57% only slightly decreases the skill premium.

Next, I show the effect of a permanent change in q on the steady state. Consistent with GHK(1997) and KORV(2000), in this model, due to the existence of ISTC the relative price of capital goods is equal to the inverse of the investment-specific technological change q . Therefore, I can use the relative price of capital to identify ISTC q . I take the price index of personal consumption expenditures from NIPA, and the quality-adjusted price index of total investment (equipment and structures) from Cummins and Violante (2002) for the time period 1951-2000. I then divide these two sequences to obtain the data counterpart of q . Figure 12 shows the log of the time series of q_t . It was fairly stable before 1957, then started to grow. The average growth rate of q in the 1960s and 1970s was 1.8% and 1.7%, respectively. It has speeded up since the early 1980s. The average growth rate in the 1980s was 3.2% and it was even higher in the 1990s (4.4%). Finally, I use the mapping mentioned before to transform the sequence of q_t to the changes in the capital productivity level B_t . Normalizing the initial steady state value B_{1951} to one, I have $B_{2000} = 3.28$. During these 50 years, the decline in the relative price of the capital goods is equivalent to an increase in capital productivity by approximately 3.3 times.

Suppose that the U.S. economy reaches the steady state again after 2000. Keeping other things equal, Table 6 shows the effects of this permanent change from B_{1951} to B_{2000} .

Investment-specific technological change, through capital-skill complementarity, increases the skill premium significantly. The mechanism is as follows: investment-specific technological change raises capital productivity B_t , and hence raises the effective capital stock, $B_t K_t$. Since capital is complementary with skilled labor, increases in effective capital also raise the demand for skilled labor. Therefore, it tends to increase the skill premium. However, rising skill premium gives individuals stronger incentive to go to college. The relative supply of skilled labor also increases. The relative quantity effect thus dampens the increase in the skill premium. The quantitative results in Table 6 confirm that the first-order impact of ISTC is on the demand side of skilled labor through capital-skill complementarity. This impact dominates the repercussion effect from the relative supply side. Hence, the skill premium increases from 1.4541 to 1.8516 in the model, which is quite close to 1.8357 in the data. The results also show that both CG and HSG significantly benefit

from this technological change.

6 Transition Path

The comparative static exercises above (especially the one with ISTC) show that we are on the right track in explaining the increases in the skill premium over time. However, to see how far the model can go to match the time series data of the skill premium in the postwar U.S. economy, comparative static analysis is not enough. One has to solve the model along a time path.

Following the spirit of the computation method in Chen, İmrohorođlu, and İmrohorođlu (2006) and Conesa and Kreuger (1999), I compute the model along a transition path from initial pre-1951 steady state towards a final steady state in the far future. The computation algorithm is described in detail in Appendix C.

6.1 Benchmark Case

In the benchmark case, I feed into the model the exogenous path of capital-specific technological change $\{B_t\}_{t=1951}^{2000}$ and demographic change embodied in the change in the growth rate of the HSG cohort size $\{n_t\}_{t=1951}^{2000}$. I also feed in the normalized tuition payments $\{tuition_t\}_{t=1951}^{2000}$. I then assume that capital-specific technological change gradually decelerates until it becomes stable in 2030, and it continues at this constant level until 2050. For simplicity, I also assume that after 2000 there is no demographic change and the tuition payment is constant at the 2000 level. Since I want to focus on the effect of ISTC, the neutral TFP change has been normalized to unity for all time periods through the transition.¹⁹ I compute the transition path of the benchmark economy between 1951 and 2050 and truncate it to the 1951-2000 period. The results are shown in Figure 13.

In Figure 13, the simulated skill premium from the benchmark economy overshoots the actual data since 1965, but it captures the increase after 1980 very well. From 1951 to 2000, the data show that the skill premium increases from 1.4546 to 1.8357 and the average growth rate (per year) during these 50 years is 0.49%. In the model the skill premium increases from 1.4543 to 1.8793 and the average growth rate is 0.54%. Focusing on 1951 to 1959, the average growth rate of the skill premium in the data is 0.73%, while in the model it is -0.32%. The model does not capture the increase through that decade. From 1963, I have annual data for the skill premium so the comparison between the data and the model's performance is more accurate. From 1963 to 1969, the average growth rate is 1.30% in the data and 1.52% in the model. From 1969 to 1981, the skill premium decreases at the average rate of 0.45%, but the model misses this decline by predicting an almost flat skill premium over this period, The average growth rate is 0.02%. The skill premium starts to increase dramatically beginning in 1981. From 1981 to 1990, the average growth rate in the data is 1.45%, while the model predicts 0.82%. That is, the model captures 56% of the change in the skill premium during this decade. From 1990 to 2000, the average growth rate in the data slows down to 0.96%, whereas the model predicts 1.01%, which overshoots the actual growth rate. Overall, for the three episodes in the “N” shape of the skill premium from 1963 to 2000, the model captures 117% of the changes

¹⁹See Section 6.3 for the sensitivity analysis when the variable neutral TFP change is allowed.

in the skill premium for the period 1963-1969 and 77% for the period 1981-2000. But the model fails to replicate the declining part of the skill premium from 1969 to 1981.²⁰²¹

6.2 Counterfactual Decomposition

To answer the quantitative question raised in the introduction, I conduct the following counterfactual experiments to isolate each exogenous change and investigate its impact on the skill premium.

I first shut down the investment-specific technological change, so the only exogenous force remaining is the demographic change. The results are shown in Figure 14.

Figure 14 shows that the model fits the skill premium data fairly well until 1980, and it generates the declining part of the skill premium during the 1970s. However, it cannot capture the dramatic increase since 1980 as shown in the data. More specifically, from 1963 to 1969, the model generates around 1% average growth rate in the skill premium, which can explain 77% of the change in this period. From 1969 to 1981, the data shows that the skill premium decreases on average at a rate of 0.45%, while the model generates an average growth rate of -0.14%. The model captures 31% of the decline in the skill premium for that period. However, from 1981, the model predicts a slight decrease in the skill premium (-0.007% per year), in contrast to the dramatic increase shown in the data.²²

Next, I shut down the demographic change. What remains is only the investment-specific technological change. Figure 15 shows that the results are similar to those in the benchmark case. The skill premium overshoots the data after the early 1970s and keeps increasing afterwards. Therefore, it captures the dramatic increase since the late 1970s, but misses the declining part of the skill premium during the 1970s. From 1963 to 1969, the average growth rate of the skill premium in the model is 0.55%. The model can thus explain about 42% of the change in the skill premium over this period. From 1969 to 1981, the data show that the skill premium decreases at an average rate of 0.45%, while the model goes the wrong direction to predict a growth at a rate of 0.26% on average. However, after 1981, the “ISTC only” model generates a 0.89%

²⁰The model also raises the enrollment rate from 41.52% to 48.07% for the period 1951-2000; while the enrollment rate in the data is from 41.54% to 63.3%. In other words, the model can explain about 30% of the increase in the enrollment rate during this period. Matching the college enrollment rate is not the target of this paper. The model here is a highly abstract one that excludes many of the important determinants of an individual’s schooling choice such as policy change (e.g., GI Bill and Vietnam War Draft) and the shifts in social norm that have affected especially the female college-going behavior and increased female enrollment

(see Goldin (2006)). As emphasized in the introduction, endogenous schooling choice is needed here to capture the dynamic interaction between the skill premium and relative supply of skilled labor. We thus can investigate better the general equilibrium effects of two exogenous changes on the skill premium. To understand why enrollment rate has changed over time is a deeper goal and it is beyond the scope of this paper.

²¹Speaking about the model performance on other dimensions of macro variables, the benchmark case predicts a rising rate of return to capital from the 1950s to mid-1980s and a declining rate since then. This trend is quite similar to the historical trend of nominal interest rate in the U.S.. The average rate of return in the model is around 4.2% over the period 1951-2000. The model also does a fairly ok job in tracking the U.S. investment-output share along the transition path. Model generates an average investment share of 20.74% from 1951 to 2000, while the U.S. data are 20.09%. For the sake of brevity, I do not present the results here. The figures are available upon the request.

²²When ISTC is shut down, the model generates little variation in the enrollment rate. From 1951 to 2000, the model predicts that the enrollment rate slightly decreases from 41.52% to 40.88%, while the data reflect an increase from 41.54% to 63.33%. The average growth rate of the enrollment rate is 0.95% in the data, while in the model it is -0.03%. Hence, demographic change does not have a significant effect on the enrollment rate over this period.

Period	Data (%)	Benchmark (%)	Demographic (%)	ISTC (%)
1963-2000	0.68	0.73	0.11	0.63
1963-1969	1.30	1.52	1.00	0.55
1969-1981	-0.45	0.02	-0.14	0.26
1981-1990	1.45	0.82	0.06	0.78
1990-2000	0.96	1.01	-0.07	0.98

Table 7: Average Annual Growth Rate of the Skill Premium: Model vs. Data

average growth rate in the skill premium over the period 1981-2000, which can explain about 74% of the increase in the skill premium in the data.²³

The decomposition results become more clear when we combine three cases discussed above in one graph. In Figure 16, I show the simulations of the skill premium in the benchmark case, in the “Demographic only” case, and in the “ISTC only” case. We can see that the “Demographic only” outcome is very close to the benchmark outcome until 1965; then these cases diverge. On the other hand, the “ISTC only” outcome closely follows the benchmark outcome since the late 1970s. From this observation, we can draw the conclusion that demographic change dwarfs ISTC before 1966, and it contributes to the decline in the skill premium in the 1970s. However, things reverse after the late 1970s, when ISTC takes over to drive the increase in the skill premium.

Table 7 compares the average annual growth rate of the skill premium in the data and in the three model cases for different periods.

The contribution of each force to the dynamics of the skill premium is summarized in Table 8. Here the contribution is measured by the ratio of the average annual growth rate of the skill premium in the model and in the data.²⁴ Overall, ISTC is much important than demographic change in explaining the pattern of the skill premium from 1963 to 2000 (93% vs. 17%). But demographic change dwarfs ISTC in the 1960s (77% vs. 42%). It can also explain about one-third of the declining for the period 1969-1981 while ISTC goes in the wrong direction. The relative importance of demographic change decreases dramatically since 1981, while ISTC becomes the major driving force behind the skill premium.

Since the effective capital-skilled labor ratio (BK/S) and relative supply of skilled labor (S/U) are the two major determinants of the skill premium in the model, I also present simulations of both ratios in Figures 17 and 18, respectively. Figure 17 shows the benchmark BK/S ratio closely tracks the “Demographic only” BK/S from 1951 to 1965, which implies that demographic change dominates ISTC over this period. The ups and downs from 1959 to 1965 are due to the dramatic demographic change during that time. From

²³In this case, the model predicts that the enrollment rate rises from 41.52% in 1951 to 48.07% in 2000, the same as in the benchmark case.

²⁴We should also be aware of that these two exogenous forces are *not* mutually exclusive. This explains why the summation of each contribution that takes separately is not equal to the contribution in the benchmark case. For example, in 1981-1990 period, the contribution of demographic change is 4% and that of ISTC is 54%. We should expect the total effect of combining the two forces together, as in the benchmark case, to be 58% if the effects on the skill premium from these two forces are orthogonal, but the actual number is 56%. The reason lies in the interaction between these two forces (refer to Section 5.2 for a discussion). Both forces contribute to the dynamics of the capital-skilled labor ratio (BK/S) and relative supply of skilled labor (S/U). And the skill premium is nonlinearly determined by these two ratios.

Period	Data (%)	Benchmark (%)	Demographic (%)	ISTC (%)
1963-2000	100	107	17	93
1963-1969	100	117	77	42
1969-1981	100	-5	31	-57
1981-1990	100	56	4	54
1990-2000	100	106	-7	103

Table 8: Decomposition of the Contribution to the Dynamics of the Skill Premium

1959 to 1960, n_t drops from 14.2% to 5.7%. This decrease in the growth rate of HSGs' cohort size lead to an increase in BK/S (recall the mechanism for the “baby bust” in Section 5.2). On the other hand, n_t increases from 1.3% in 1962 to 17.5% in 1963. This huge increase changes the age structure significantly and decreases the asset accumulation in the economy. Therefore, BK/S decreases drastically. Since 1965, the benchmark BK/S closely follows the one in the “ISTC only” case. This suggests that ISTC becomes the major driving force to affect this ratio. And it is the drastic rise in the capital-skilled labor ratio since the late 1970s, through capital-skill complementarity, that drives the rising skill premium.

Figure 18 shows the relative supply of skilled labor in the benchmark and two decomposition cases. The “Demographic only” case predicts decreasing S/U from 1951 to 1976 and increasing S/U since then. This is consistent with the mechanism mentioned in Section 5.2: the “baby boom” decreases S/U , while the “baby bust” does the opposite. In contrast, the “ISTC only” model shows a continuously rising S/U ratio by generating an increase in the enrollment rate over time. S/U in the benchmark case is in between these two decomposition cases, as it decreases until 1966, then begins to increase and converges to the ratio in “ISTC only” model after 1990. This “ J ” pattern confirms that demographic change is a dominating force in driving the evolution of the relative supply of skilled labor before 1966, but the effect phases out after that. ISTC catches up to become the major driving force.

6.3 Sensitivity Analysis

In this subsection, I show that the paper's results are robust to the choice of the model setting, alternative values of key parameters, including neutral technological change, increase in life expectancy, and the time of the final steady state.

6.3.1 Retirement and Social Security

First, the current model does not include retirement. However, a specification including retirement and social security yields very similar results. In benchmark economy, the skill premium increases from 1.4543 in 1951 to 1.8793 in 2000. While in an extension with life span of 82 periods (people retire at age 66 and live up to age 100) and a pay-as-your-go social security system (replacement ratio is 50%), the skill premium increases from 1.4546 to 1.8726 for the same period.

6.3.2 Elasticity of substitution

Next, since the two elasticity parameters in the production function, θ and ρ , are the key parameters in the model, I verify the sensitivity of the analysis to different values of these two parameters. Specifically, I use the values $\theta = 0.33$ and $\rho = -0.67$, which are taken from Fernandez-Villaverde (2001) since we share the same specification of the production function. This implies that the elasticity of substitution between capital and skilled labor is 0.60, and the one between unskilled and skilled labor is 1.49. I then recalibrate the model according to these new values. All the comparative static and transition path results are similar to the benchmark results shown here. For example, the skill premium increases from 1.4558 to 1.8233 for 1951-2000 period.

6.3.3 Neutral TFP Change

So far the model keeps the total factor productivity at a constant level and hence excludes the neutral technological change. Since the TFP change does not enter into equation (14) that determines the dynamics of the skill premium, we should not expect that including TFP growth would change the results significantly. As a robustness check, I carry on an experiment which allows TFP grows at a rate of 0.15% along the transition path. Including neutral TFP change slightly increases the skill premium. From 1951 to 2000, the skill premium increases from 1.4550 to 1.9124. While in the benchmark model, it increases from 1.4553 to 1.8793.

6.3.4 Increasing Life Expectancy

Demographic changes were not only happening in terms of fluctuations in the fertility rates. In 1950 life expectancy at birth in the US was only 68.2 years. It was at 77 years in 2000. To take into account the effects of this demographic dimension, I do the following experiment. In an extended model with retirement and social security (people retire at age 66 and live up to age 100), I first use the survival probability taken from US Life Table for 1949-1951 period to calibrate the model. It predicts that the skill premium in the initial steady state is 1.4562. I then keep all the other parameters unchanged except that now the survival probability is replaced by the data taken from US Life Table for 1999-2001 period. Although the survival probability has been increased substantially for these 50 years, which leads to the increase in life expectancy, the model predicts that the skill premium only changes slightly to 1.4218. This experiment implies that increasing life expectancy would not change our results significantly.

6.3.5 Timing of the Final Steady State

Finally, I test the robustness of the model to the timing of the final steady state. In the current model, I set it in an arbitrary way so that the economy reaches the final steady state after 2050. However, the choice of this timing does not affect the results significantly. For example, if the model reaches the final steady state right after 2000, I still obtain almost identical results as in the benchmark model. The skill premium increases from 1.4544 to 1.8854 for 1951-2000 period.

Figure 19 summarizes the simulated transition path for the skill premium in the different experiments mentioned above.

7 Conclusion

The skill premium (college wage premium) in the U.S. increased in the 1950s and 1960s, decreased in the 1970s, and has increased dramatically since 1980. What are the driving forces behind this “N” shape? The previous literature proposes several explanations including skill-biased technological change (SBTC) and demographic change. However, little attention has been paid to investigate the relative importance of each driving force on the evolution of the skill premium. In this paper, I establish and compute an overlapping generations general equilibrium model with endogenous schooling choice to answer an important quantitative question: what percentage of the change in the skill premium for the post-war period in the U.S. can be explained by demographic change and investment-specific technological change (ISTC), respectively?

In this model, ISTC and demographic change drive the equilibrium outcomes of the skill premium by dynamically affecting the relative demand and supply of skilled labor. ISTC, through the key feature of capital-skill complementarity in the production technology, increases the relative demand of skilled labor, and thus raises the skill premium. In turn, the rising skill premium encourages skill formation and increases the relative supply of skilled labor. In contrast, demographic change affects the age structure in the economy. A change in the age structure has a direct impact on the relative supply of skilled labor. In addition, since people have different saving tendencies along the life cycle, a change in the age structure also influences the relative demand of skilled labor through changing asset accumulation in the economy. The ultimate effects of these two forces on the skill premium depend on the quantitative magnitude of both demand and supply effects.

I calibrate the model to match U.S. data for the period 1947-1951 as the initial steady state. Then, by feeding in the ISTC data from Cummins and Violante (2002) and the growth rate of the HSG cohort size from 1951 to 2000, I conduct perfect foresight deterministic simulations to compare with the data of the 1951-2000 period and counterfactual decomposition experiments to identify the effects of each force.

The results show that demographic change dwarfs ISTC before the late 1960s and accounts for about one-third of the decline in the skill premium in the 1970s. However, ISTC takes over to drive the dramatic increase in the skill premium since the early 1980s, explaining about three-fourth of the increase in the skill premium since 1981.

By carefully conducting this calibration exercise, the current paper helps to narrow down the explanations of rising skill premium. In a more comprehensive dynamic general equilibrium framework, this paper confirms KORV’s (2000) main finding on quantitative importance of capital-skill complementarity in affecting the skill premium in the post-war period. Goldin and Katz (2007) claim that the slowdown in the growth of the relative supply of college workers starting around 1980 was a major reason for the rising skill premium from 1980 to 2005. Education has been lost to technology in the race. This paper shows that the race has been lost is mostly due to the speedup of the technology rather than the slowdown of the supply.

A possible direction to extend the current study is to relax the assumption of perfect substitution across age groups in skilled and unskilled labor. Card and Lemieux (2001) show that virtually the entire rise in the skill premium is attributable to the changes in the relative earnings of *younger* college-educated workers. And they claim that shifts in cohort-specific supplies of college-educated workers play a very important role. It would be interesting to see once we allow imperfect substitution between similarly educated workers in different age groups, can the model replicate the above fact and improve its performance.

References

- [1] Acemoglu, D. (2003): “Patterns of Skill Premia,” *Review of Economic Studies*, Vol. 70, 199-230.
- [2] Autor, D. H., L. F. Katz and A. B. Krueger (1998): “Computing Inequality: Have Computers Changed the Labor Market?” *Quarterly Journal of Economics*, Vol. 113, No. 4, 1169-1213.
- [3] Ben-Porath, Y. (1967): “The Production of Human Capital and the Life Cycle of Earnings,” *Journal of Political Economy*, Vol. 75, No. 4, 352-365.
- [4] Bound, J. and G. Johnson (1992): “Changes in the Structure of Wages in the 1980’s: An Evaluation of Alternative Explanations,” *American Economic Review*, Vol. 82, 371-392.
- [5] Card, D. and T. Lemieux (2000): “Dropout and Enrollment trends in the Post-War Period: What Went Wrong in the 1970s?” NBER Working Paper 7658.
- [6] Card, D. and T. Lemieux (2001): “Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis,” *Quarterly Journal of Economics*, Vol. 116, No. 2, 705-746.
- [7] Chen, K., A. İmrohorođlu and S. İmrohorođlu (2006): “Japanese Saving Rate,” *American Economic Review*, Vol. 96, No. 5, 1850-1858.
- [8] The College Board: Trends in College Pricing 2003.
- [9] Conesa, J. C. and D. Krueger (1999): “Social Security Reform with Heterogeneous Agents” *Review of Economic Dynamics*, Vol. 2, 757-795.
- [10] Cummins, J. G. and G. L. Violante (2002): “Investment-Specific Technical Change in the United States (1947-2000): Measurement and Macroeconomic Consequences,” *Review of Economic Dynamics*, Vol. 5, 243-284.
- [11] Duffy, J., C. Papageorgiou, and F. Perez-Sebastian (2004): “Capital-Skill Complementarity? Evidence from a Panel of Countries,” *Review of Economics and Statistics*, Vol. 86, No. 1, 327-344.
- [12] Eckstein, Z. and E. Nagypál (2004): “The Evolution of U.S. Earnings Inequality: 1961-2002,” *Federal Reserve Bank of Minneapolis Quarterly Review*, Vol. 28, No. 2, 10-29.
- [13] Fernandez-Villaverde, J. (2001): “Was Malthus Right? Economic Growth and Population Dynamics,” University of Pennsylvania, working paper.
- [14] Goldin, C. (2006): “The Quiet Revolution That Transformed Women’s Employment, Education, and Family,” *American Economic Review*, Vol. 96, No. 2, 1-21.
- [15] Goldin, C. and L. F. Katz (2007): “The Race between Education and Technology: The Evolution of U.S. Educational Wage Differentials, 1890 to 2005,” NBER Working Paper #12894.
- [16] Gourinchas, P. O. and J. A. Parker (2002): “Consumption over the Life Cycle,” *Econometrica*, Vol. 70, No. 1, 47-89.

- [17] Greenwood, J., Z. Hercowitz and P. Krusell (1997): “Long-Run Implications of Investment-Specific Technological Change,” *American Economic Review*, Vol. 87, No. 3, 342-362.
- [18] Griliches, Z. (1969): “Capital-Skill Complementarity,” *Review of Economic and Statistics*, Vol. 51, No. 4, 465-468.
- [19] Guvenen, F. and B. Kuruscu (2006): “Understanding Wage Inequality: Ben-Porath Meets Skill-Biased Technical Change,” University of Texas at Austin, working Paper.
- [20] He, H. and Z. Liu (2008): “Investment-Specific Technological Change, Skill Accumulation, and Wage Inequality,” *Review of Economic Dynamics*, Vol. 11, 314-334.
- [21] Heckman, J. J., L. Lochner and C. Taber (1998): “Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents,” *Review of Economic Dynamics*, Vol. 1, 1-58.
- [22] İmrohorođlu, A., S. İmrohorođlu and D. H. Joines (1995): “A Life Cycle Analysis of Social Security,” *Economic Theory*, Vol. 6, 83-114.
- [23] İmrohorođlu, A., S. İmrohorođlu and D. H. Joines (1999): “Social Security in an Overlapping Generations Model with Land,” *Review of Economic Dynamics*, Vol. 2, 638-665.
- [24] Katz, L. F. and K. M. Murphy (1992): “Changes in Relative Wages, 1963-1987: Supply and Demand Factors,” *Quarterly Journal of Economics*, Vol. 107, No. 1, 35-78.
- [25] Katz, L. F. and D. H. Autor (1999): “Changes in the Wage Structure and Earnings Inequality,” in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, vol. 3A, North-Holland, 1999, 1463-1555.
- [26] Krusell, P., L. E. Ohanian, J. V. Rios-Rull and G. L. Violante (2000): “Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis,” *Econometrica*, Vol. 68, No. 5, 1029-1053.
- [27] Macunovich, D. J. (1996): “Cohort Size Effects on US Enrollment Decisions,” Discussion Paper No. 36, Williams Project on the Economics of Higher Education, Williams College.
- [28] McGrattan, E. R. and L. E. Ohanian (2008): “Does Neoclassical Theory Account for the Effects of Big Fiscal Shocks? Evidence from World War II,” *Federal Reserve Bank of Minneapolis Research Department Staff Report* 315.
- [29] National Center for Education Statistics (NCES): *Digest of Education Statistics*, various issues.
- [30] Navarro, S. (2007): “Using Observed Choices to Infer Agent’s Information: Reconsidering the Importance of Borrowing Constraints, Uncertainty and Preferences in College Attendance,” manuscript, University of Wisconsin at Madison.
- [31] Standard Education Almanac, 1968, 1969, 1972, 1979-1980, 1981-82, 1984-85.

8 Appendix A. Data

The skill premium and relative supply of skilled labor data in Figure 1 are taken from Katz and Autor (1999). The data come from the 1940, 1950, 1960 Censuses and the 1964-1997 March CPS. The skill premium in their paper is the coefficient on workers with a college degree or above relative to high school graduates in a log weekly wage regression. The sample includes full-time full-year workers aged between 18 and 65. The relative supply of skilled labor is the ratio between college equivalents and non-college equivalents, using weeks worked as weights. Here, college equivalents= $CG+0.5\times$ workers with some college, and non-college equivalents= $High\ School\ Dropout+HSG+0.5\times$ workers with some college. Figure 1 is also the same as Figure 1 in Acemoglu (2003). Please refer to Katz and Autor (1999) or Data Appendix in Acemoglu (2003) for detailed data construction.

HSG cohort size data in Figure 2 are from the National Center for Education Statistics (NCES): *Digest of Education Statistics (DES)* 2002, Table 103. Data for 1941, 1943, and 1945 come from *DES* 1970 Table 66. In this figure “year” refers to school year, for example, 1939 refers to the school year 1939-1940.

The 18-21 year-old population data in Figure 3 come from different sources. 1970-2000 data are from NCES: *DES* 2002, Table 15. 1960-1969 data are from NCES: *DES* 1995. 1955-1959 data are from *Standard Education Almanac* 1968, Table 1. The 17-year-old population data are from NCES: *DES* 2002, Table 103.

College enrollment rates of HSG for 1960-2001 in Figure 4 are from NCES: *DES* 2002, Table 183. 1948-1959 data are calculated by the author. To construct them, first I take the 1948-1965 data of first-time freshmen enrolled in institutions of higher education (from NCES: *DES* 1967, Table 86), divided by the HSG cohort size as in Figure 2. Since first-time freshmen are not necessarily from recent HSG, I use the overlapped years 1960-1965 to calculate the average difference between my calculation and the true data, then adjust my calculation for the 1948-1959 period according to this difference.

Average TFRB charges data in Figure 5 are constructed as follows. First, I obtain data about estimated Average Charges to Full-Time Resident Degree-Credit Undergraduate Students between 1956-57 and 1966-67 from *Standard Education Almanac* 1969, Table 120; 1967-68 to 1973-74 from *Standard Education Almanac* 1981-82, P. 231-232; 1974-75 to 1983-84 data from *Standard Education Almanac* 1984-85, P. 328-329; 1984-85 to 2003-04 data from “The Trends in College Pricing 2003”, the College Board, Tables 5a, 5b. 1948-1955 data are from *Standard Education Almanac* 1968, Table 102: “Estimated Costs of Attending College, Per Student: 1931-1981”. To make it consistent with the data after 1955, I use the overlapped 1956 data to adjust. Second, I focus only on public or private four-year institutions. I obtain the TFRB charges for those institutions. Third, I calculate the enrollment share of public and private four-year institutions. For the 1948-1964 data, I obtain the total fall enrollment in degree-granting institutions by control of institution (private vs. public) from NCES: *DES* 2002 Table 172, noticing that applies to all higher education institutions. Then, from Table 173, I have total fall enrollment in degree-granting institutions by control and type of institution from 1965 to 2000. Fourth, I weight the average TFRB charges of public and private four-year institutions by enrollment share, then use the Personal Consumption Expenditure deflator from NIPA to convert them into constant 2002 dollars. Finally, the third-order moving average method is used to smooth the data.

The construction of the cohort-specific skill premium data is in the text. (See Section 4.1.)

The skill premium used in this paper is the ratio of the real mean annualized wage of CG and HSG as in Figure 6. The data counterpart of the relative supply of skilled labor (S/U) is the ratio of weeks worked of

CG plus some college and HSG. I obtain these data from Katz and Autor's (1999) data set.

9 Appendix B. Algorithm to Compute the Stationary Equilibrium

Given the parameter values as shown in the text, I compute the stationary equilibrium as follows:

1. Guess the initial values for capital stock K_0 and initial enrollment rate e_0 .
2. Given the initial guesses, calculate the skilled labor S_0 and the unskilled labor U_0 . Notice that for every j , $\sum_a \lambda^c(a, j) = \text{enrollment rate}$ and $\sum_a \lambda^h(a, j) = 1 - \text{enrollment rate}$, so by equation (20) and (21) I have

$$\begin{aligned} S_0 &= (\text{enrollment rate}) \cdot \sum_j \mu_j \varepsilon_j^c, \\ U_0 &= (1 - \text{enrollment rate}) \cdot \sum_j \mu_j \varepsilon_j^h. \end{aligned}$$

Given all the inputs, from the firm's FOCs (10)-(12), I can compute the interest rate r and wage rates w^c and w^h .

3. Discretize the asset level $-b \leq a \leq a_{\max}$ (make sure that the borrowing limit b and the maximum asset a_{\max} will never be reached). Given prices $\{w^c, w^h, r\}$, feed in the normalized tuition data. By using backward induction (remember that $a_J = 0$), I can solve the conditional value function $V^s(a, 1)$ for $s = h, c$, and therefore obtain the cut-off disutility cost $DIS(i^*) = (V^c(a = 0, 1) - V^h(a = 0, 1))/sd$, where sd is the scale factor of disutility cost, which is calibrated to replicate the enrollment rate data.
4. From the estimated CDF function of disutility cost in Figure 9, corresponding to $DIS(i^*)$, obtain the new enrollment rate e_1 . Check the convergence criterion ($\frac{|e_0 - e_1|}{e_0} \leq tol_e$). If it is not satisfied, update it by the relaxation method

$$e_2 = \kappa_e e_0 + (1 - \kappa_e) e_1,$$

where $0 < \kappa_e < 1$ is the relaxation coefficient for the enrollment rate.

5. Using the decision rules obtained in step 3, $A^s(a, j) \forall j, \forall a$, and equation (18), compute the age-dependent distributions by forward recursion. Then use these distributions $\lambda^s(a, j)$ and the age shares μ_j to compute the per capita (next period) capital stock K_1 as follows

$$K_1 = \left(\sum_j \sum_a \sum_s \mu_j \lambda^s(a, j) A^s(a, j) \right) / (1 + n),$$

where n is the growth rate of cohort size. Check the convergence criterion ($\frac{|K - K_1|}{K} \leq tol_K$) to see if it needs to stop. If not, use the relaxation method to update K

$$K_2 = \kappa_K K_0 + (1 - \kappa_K) K_1,$$

where $0 < \kappa_K < 1$ is the relaxation coefficient for capital. Then set $K_0 = K_2$, $e_0 = e_2$, and go back to step 1. The iteration will stop once all errors fall into the tolerance ranges.

6. Compute aggregate consumption, investment, tuition expense, and output by using the decision rules, age-dependent distributions, and age shares

$$\begin{aligned}
C &= \sum_{j=1}^J \sum_a \sum_s \mu_j \lambda^s(a, j) C^s(a, j), \\
TUITION &= \sum_{j=1}^4 \sum_a \mu_j \lambda^c(a, j) tuition_j, \\
X &= (n + \delta)K.
\end{aligned}$$

7. Finally, check if the market clearing condition given by equation (22) holds. If it does, stop.

10 Appendix C. Algorithm to Compute the Transition Path

In this paper I follow Chen, İmrohorođlu and İmrohorođlu (2006) (also see Conesa and Kreuger (1999)) in computing a transition path from the initial pre-1951 steady state towards a final steady state. In this way, I view 1952-2000 as a part of the transition path. Defining notation, I use $t = 1$ for 1951, $t = T$ for the final steady state, and $t = 2, \dots, T - 1$ for the transitional period. I take the following steps in the computation.

1. Compute the pre-1951 initial steady state by following the method described in Appendix B. Save the distribution $\lambda^s(a, j), \forall j, \forall s$ for later calculation.
2. Feed in the exogenous change in the growth rate of HSGs' cohort size $\{n_t\}_{t=1}^T$, transformed capital-specific technological change $\{B_t\}_{t=1}^T$, and the tuition payments $\{tuition_t\}_{t=1}^T$.
3. Compute the final steady state at $t = T$. Save the value function $V^s(a, j), \forall j, \forall s$ for later calculation.
4. Take the initial and final steady state values of capital stock K and enrollment rate e , and use linear interpolation to guess the sequences of $\{K_t\}_{t=1}^T$ and $\{e_t\}_{t=1}^T$. This is the initial guess for the transition path computation.
5. Start from $T - 1$, take the value function of the final steady state as the terminal values $V^s(a, j, T)$, solve the individual optimization problem by backward induction, and obtain the decision rules for all cohorts through the transition path.
6. Use the distribution of the pre-1951 steady state as the initial asset distribution, together with the decision rules collected from step 5, and calculate $\{\lambda^s(a, j, t)\}_{t=2}^T$ by forward recursion. Then use them to calculate new $\{K_t\}_{t=1}^T$ and $\{e_t\}_{t=1}^T$.
7. Compare the new sequences of endogenous variables $\{K_t\}_{t=1}^T$ and $\{e_t\}_{t=1}^T$ with the initial guess and iterate until convergence.

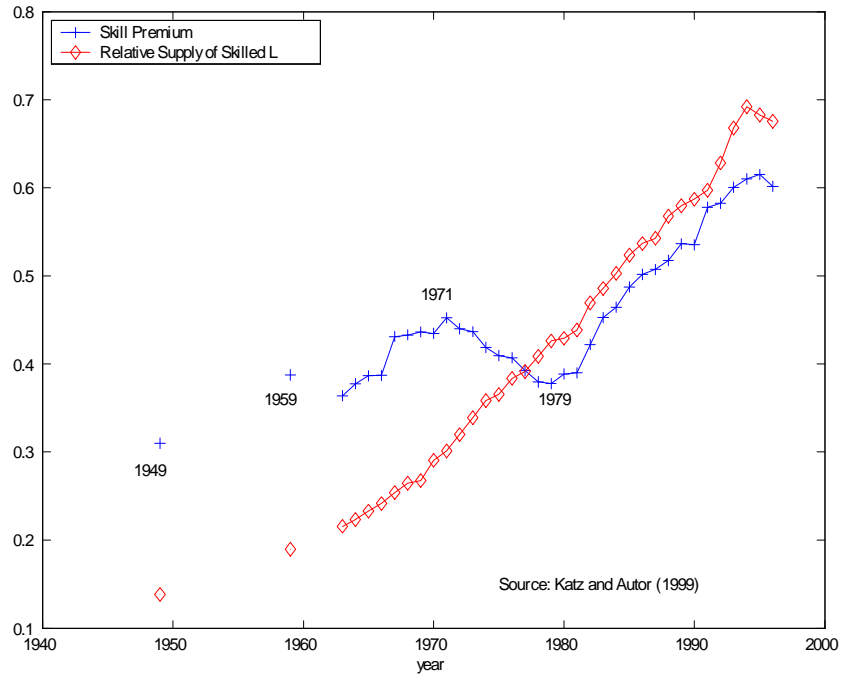


Figure 1: The Skill Premium (Log Units) and Relative Supply of Skilled Labor: 1949-1996

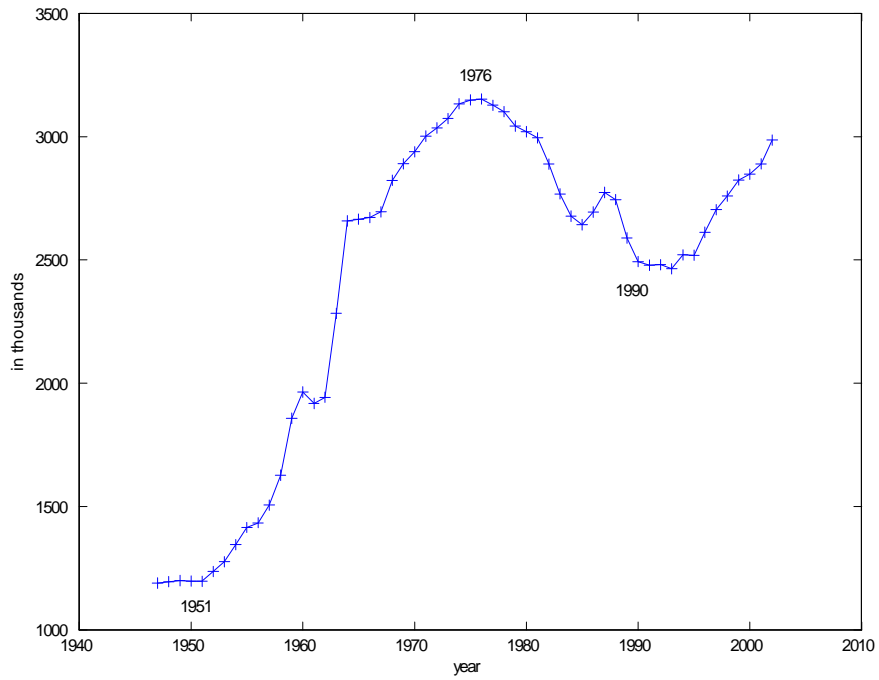


Figure 2: High School Graduates Cohort Size

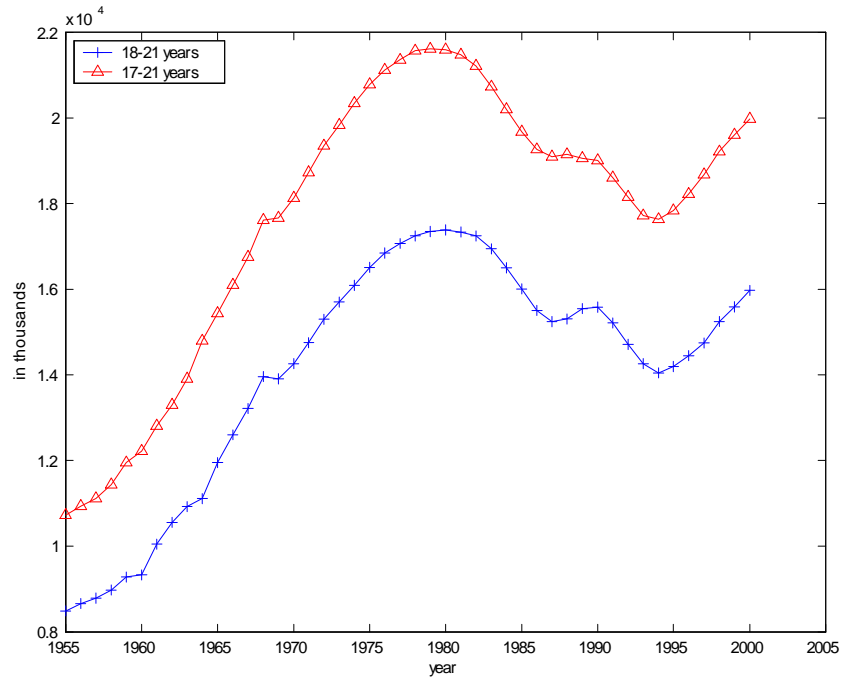


Figure 3: College Age Population: 1955-2000

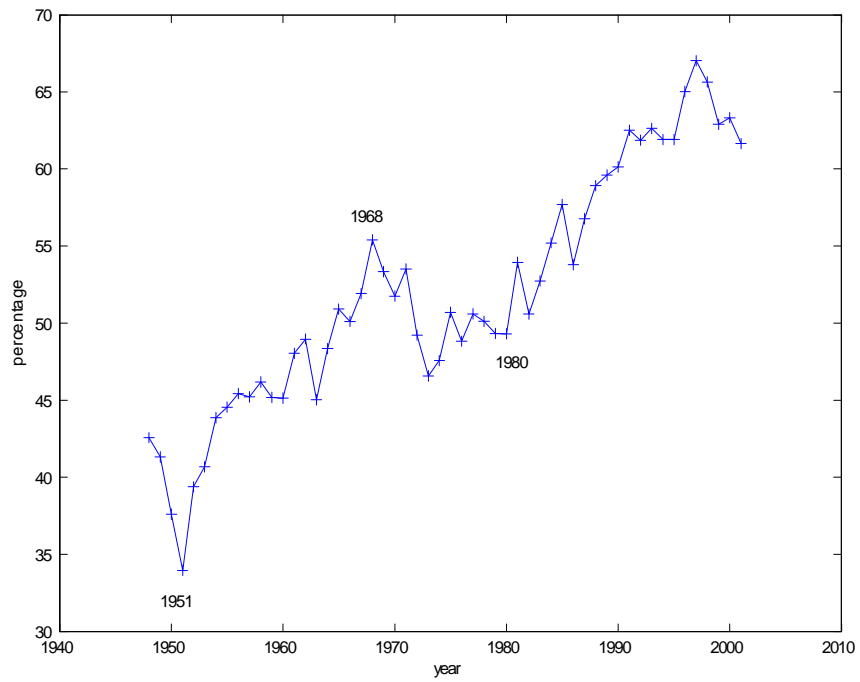


Figure 4: College Enrollment Rate of High School Graduates: 1948-2001

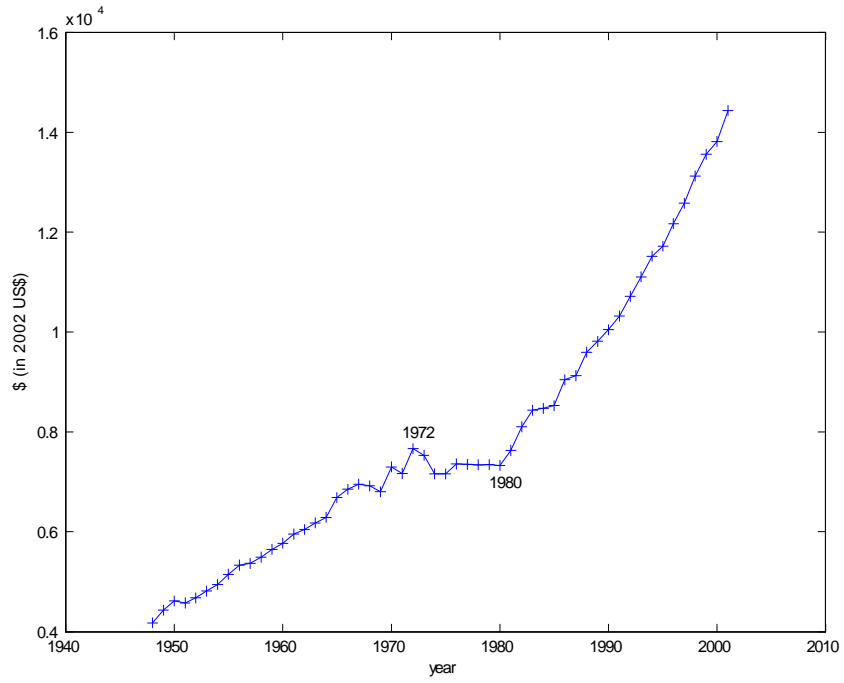


Figure 5: Average Real TFRB charges: 1948-2001

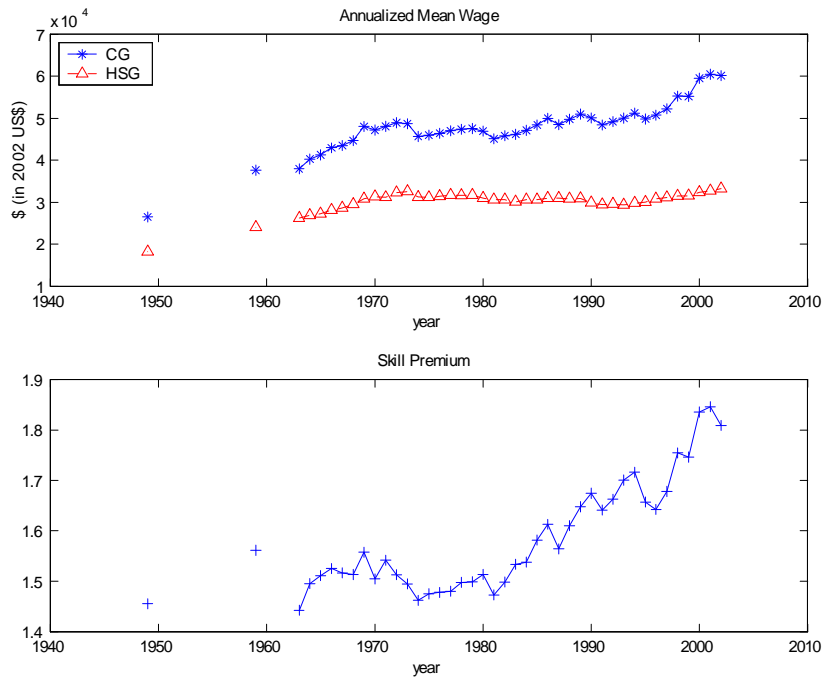


Figure 6: Real Annualized Mean CG and HSG Wage: 1949-2002

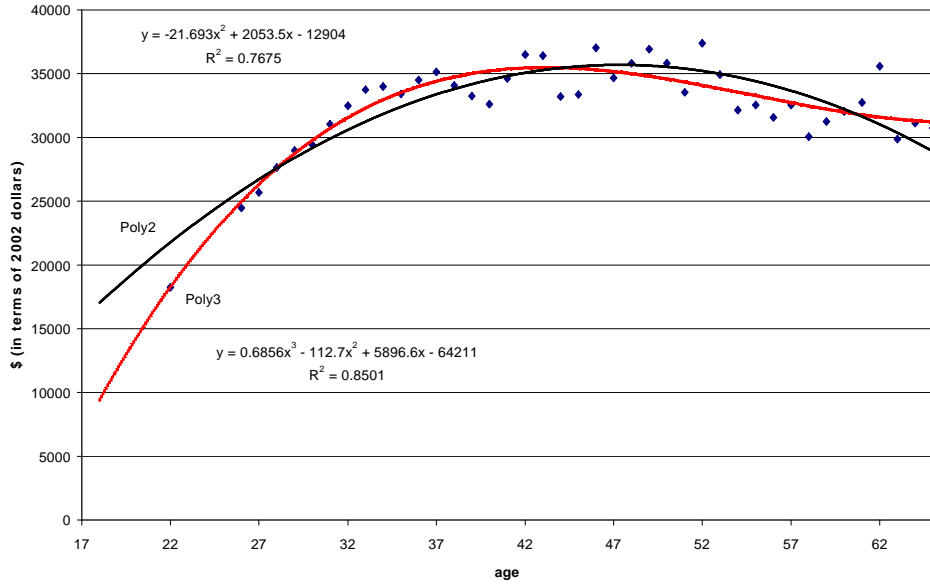


Figure 7: Lifecycle HSG Wage Profile: 1955 Cohort

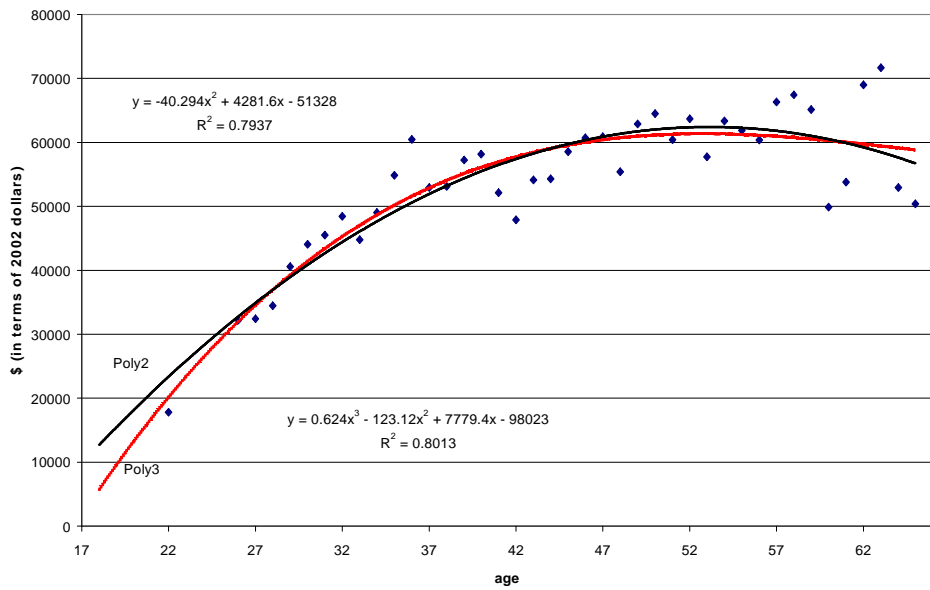


Figure 8: Lifecycle CG Wage Profile: 1955 Cohort

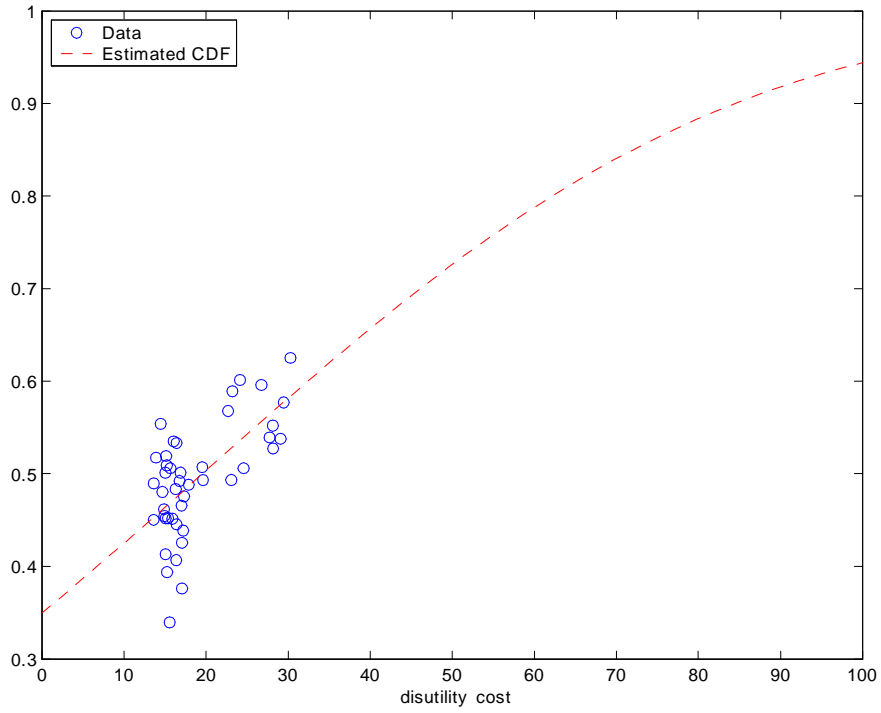


Figure 9: CDF of Disutility Cost

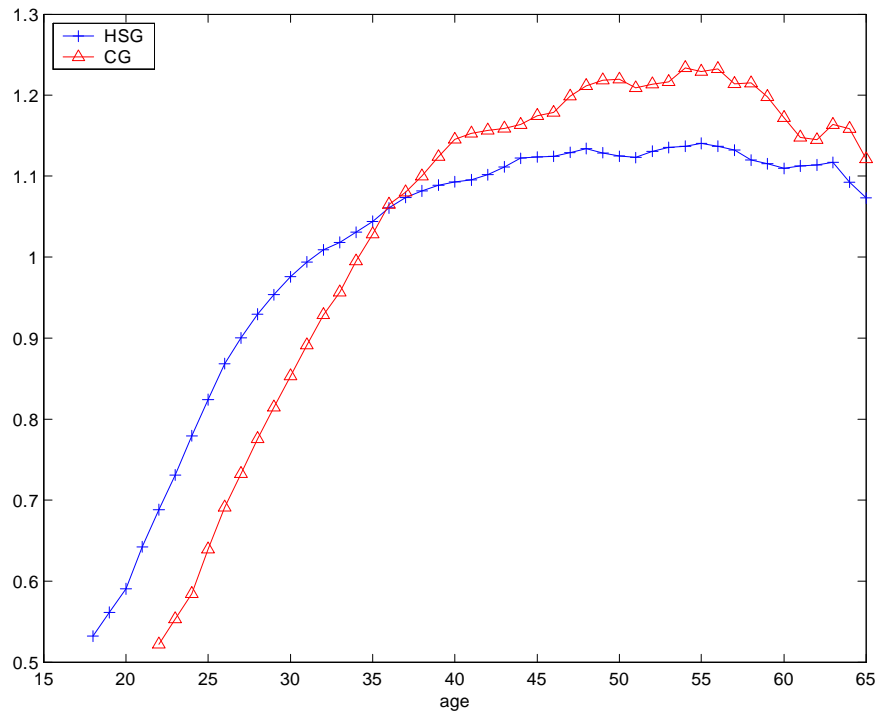


Figure 10: Age Efficiency Unit Profile: 1949-2002 Average

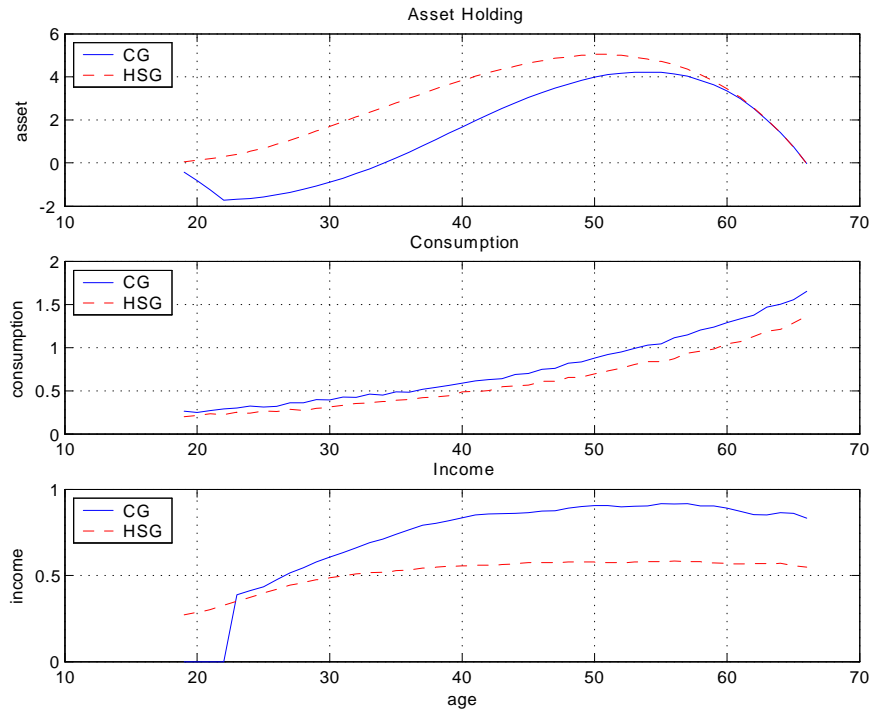


Figure 11: Lifecycle Profiles of HSG and CG

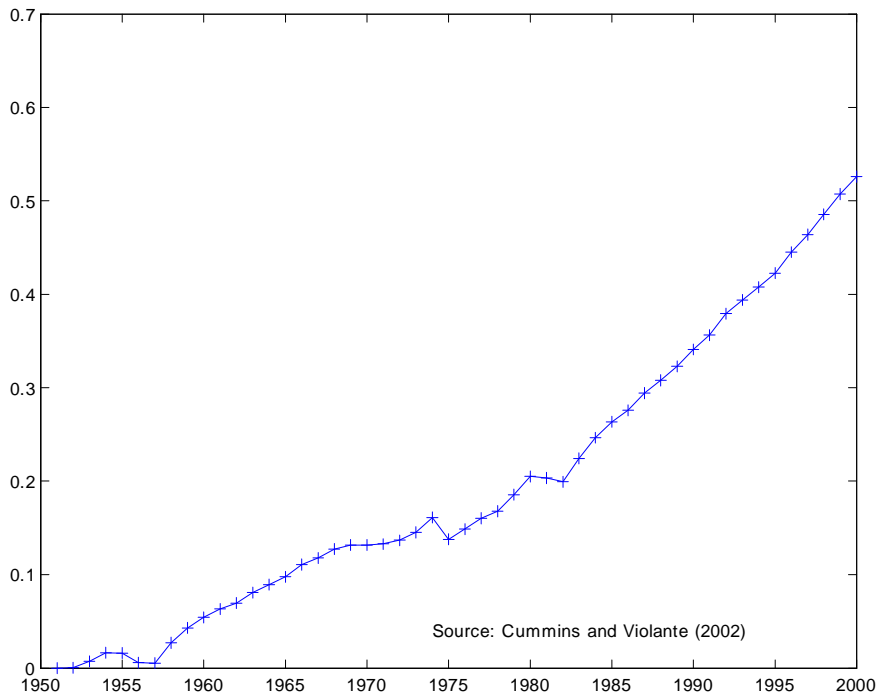


Figure 12: Investment-Specific Technological Change (Log Units)

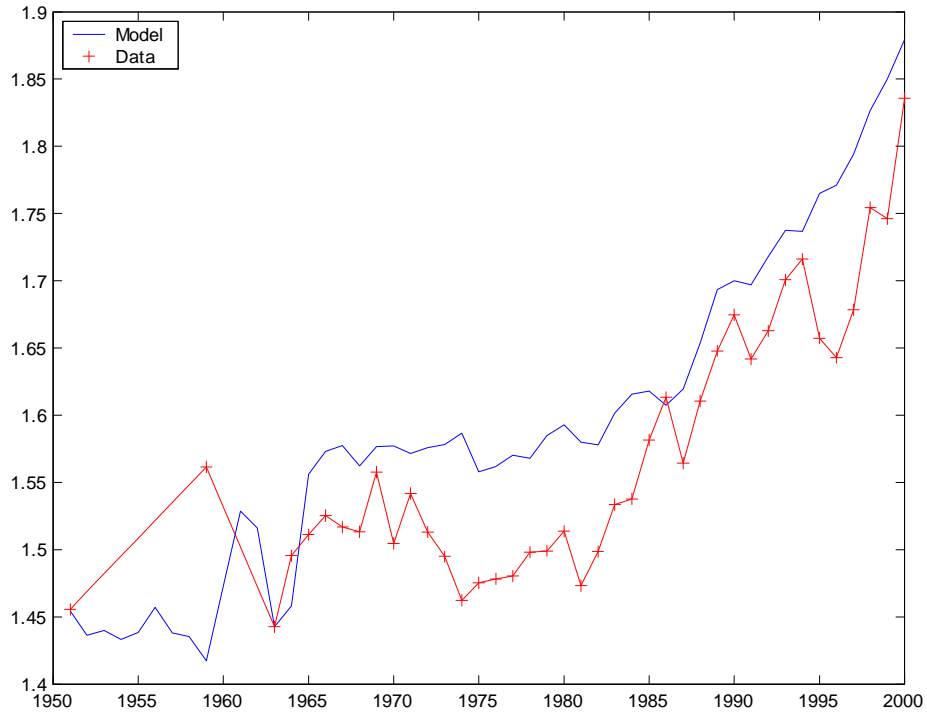


Figure 13: Model vs. Data: Benchmark Case

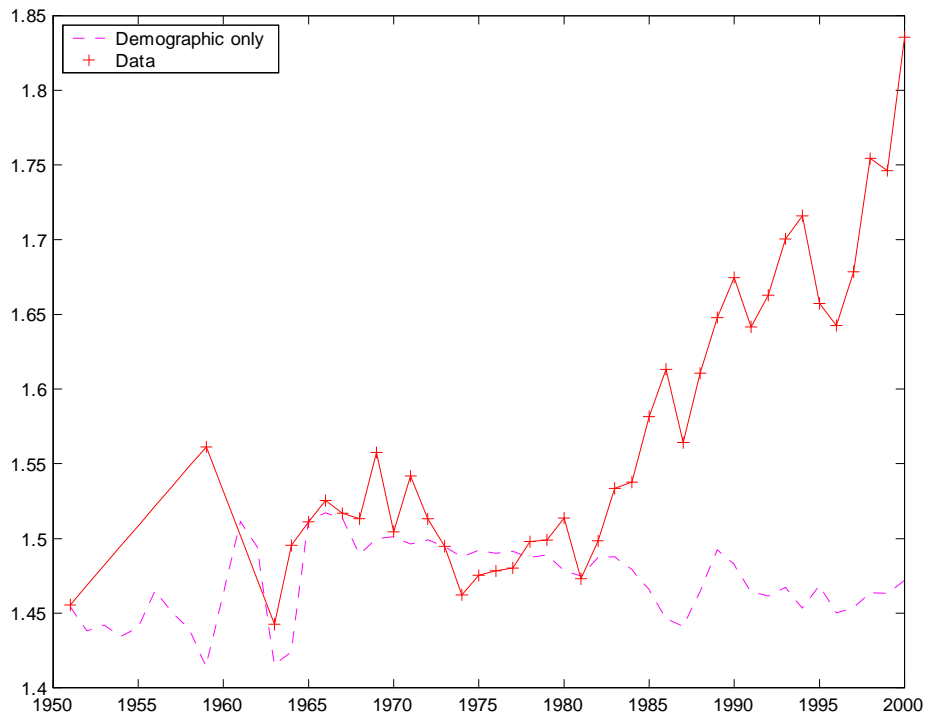


Figure 14: Model vs. Data: Demographic Change Only

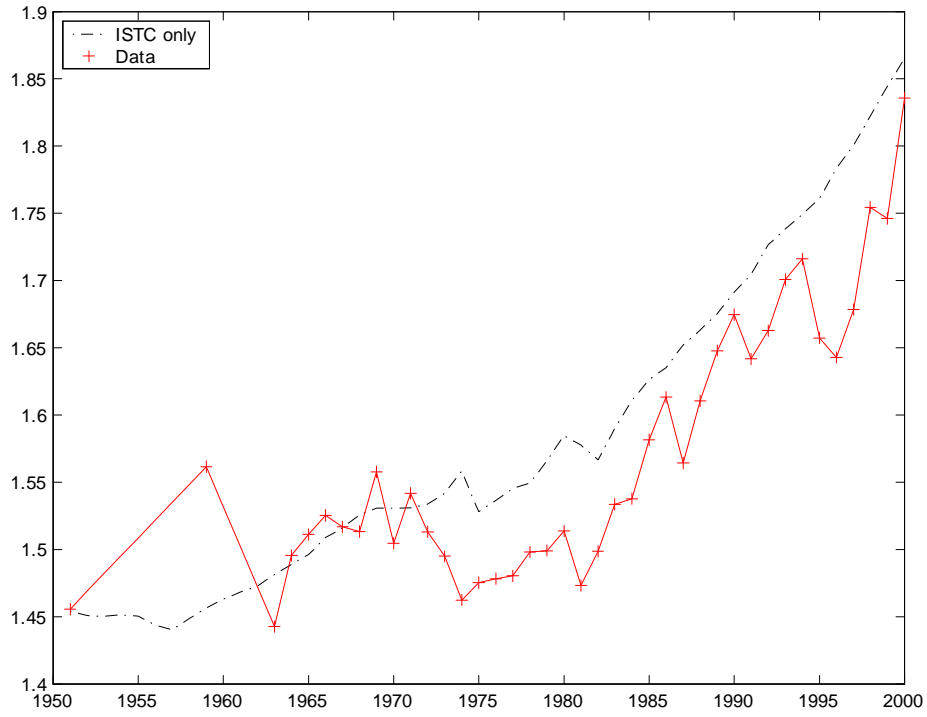


Figure 15: Model vs. Data: ISTC Only

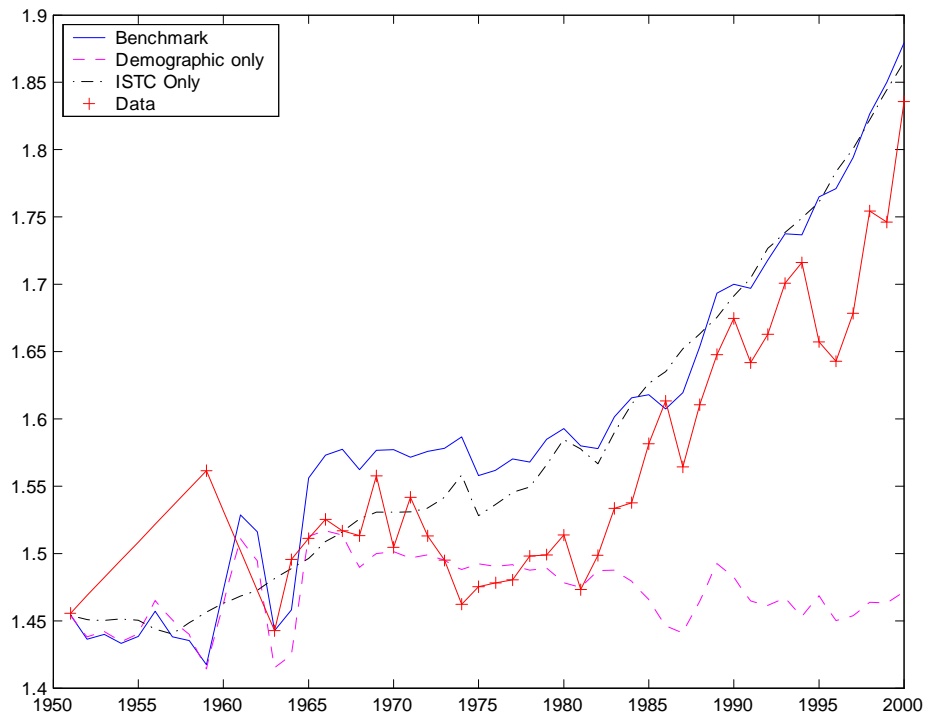


Figure 16: Skill Premium: Model vs. Data

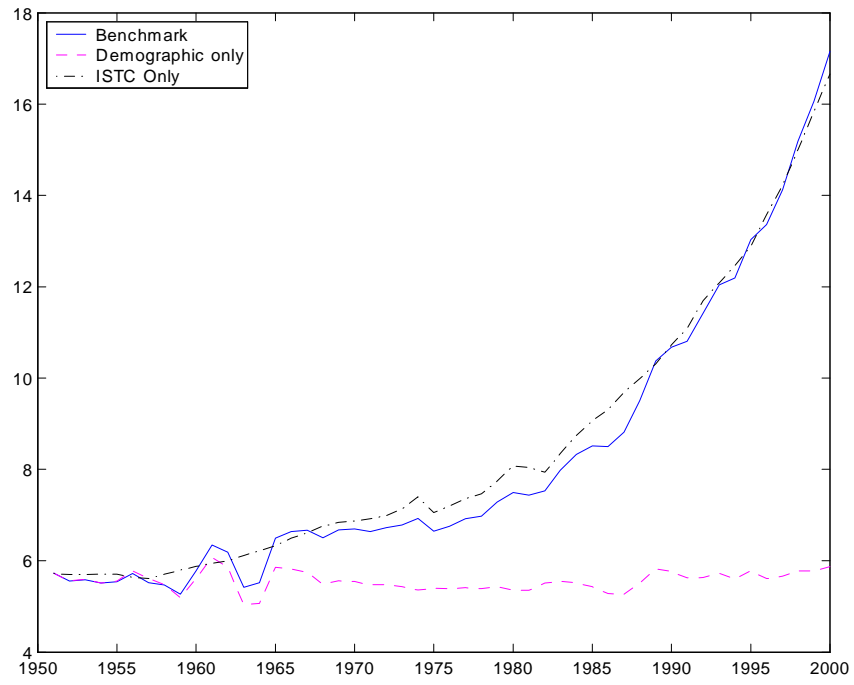


Figure 17: Capital-Skill Labor Ratio: Decomposition

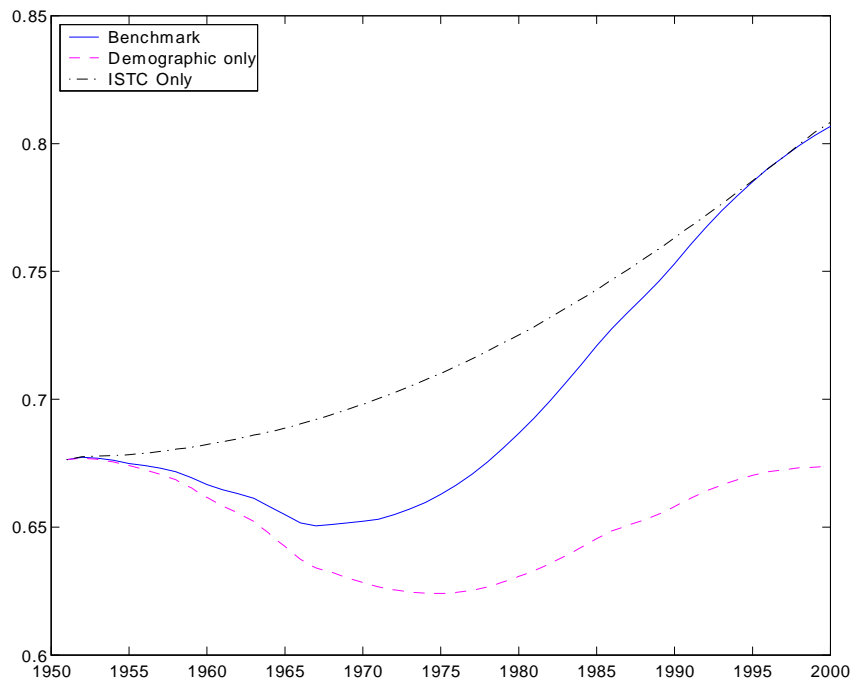


Figure 18: Relative Supply of Skilled Labor: Decomposition

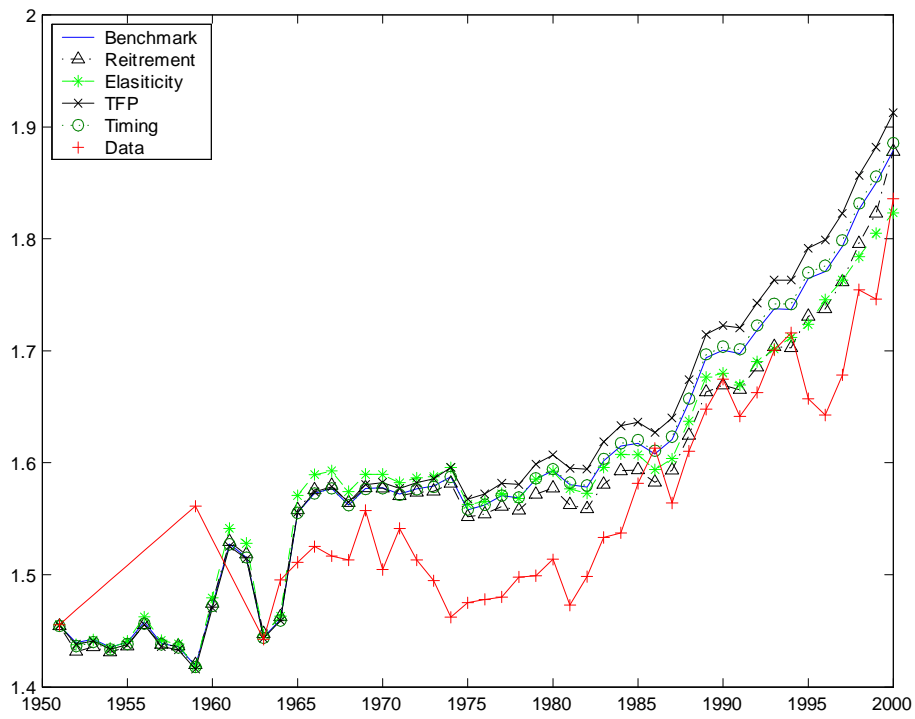


Figure 19: Skill Premium: Sensitivity Analysis