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

Hui He
University of Hawaii
December 28, 2011

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. We develop a general equilibrium overlapping generations 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 by changing the age structure and hence the relative supply of skilled labor. We find that demographic change is more important in shaping the skill premium before 1980. Since then, ISTC has taken over and has been driving the dramatic increase in the skill premium.

JEL classification: E24; E25; J24; J31; O33; I21.

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 shows, starting in 1949 the evolution of the skill premium (expressed in natural logarithm) exhibited an “N” shape: it increased in the 1950s and most of the 1960s, then decreased throughout the 1970s, and has increased dramatically since then. Meanwhile, the figure also shows the relative supply of skilled labor (the ratio of annual hours worked by skilled labor to annual hours worked by unskilled labor) has been increasing steadily over time.¹

¹Mailing Address: Department of Economics, University of Hawaii, 2424 Maile Way, Saunders Hall 542, Honolulu, HI 96822. E-mail: huihe@hawaii.edu. Tel: 1-808-956-7653; Fax: 1-808-956-4347.

¹The annual data since 1963 are taken from the CPS. Data in 1949 and 1959 are obtained from the 1950 and 1960 Census. See Appendix A for the details. The pattern of the skill premium in Figure 1 has been widely documented in the literature. For example, Figure 1 in Acemoglu (2003) shows a similar pattern for both skill premium and relative supply of skilled labor.
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 Krusell, Ohanian, Rios-Rull, and Violante 2000, hereafter KORV), skill-biased technological change induced by international trade (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 with a smooth secular increase in the relative demand for skilled labor combined with the observed fluctuations in the rate of growth of relative supply of skilled labor can potentially explain the dynamics of the skill premium. They especially attribute the fluctuations in the relative supply of skilled labor to the demographic change. They also argue that the accelerating skill-biased technological change in the 1980s, has contributed to the rising college wage premium since 1980 via raising relative demand of skilled labor.

This paper contributes to the literature 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? We develop a general equilibrium overlapping generations model 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 to unskilled labor. Third, following Greenwood, Hercowitz, and Krusell (1997, hereafter GHK), We assume the existence of investment-specific technological change, i.e., a technological change on investment goods.

The contribution of the paper is threefold. First, in contrast to the partial equilibrium setup by KORV (2000), it uses a general equilibrium framework to evaluate the relative importance of the above two exogenous factors in shaping the skill premium. Second, this paper provides a microfoundation for the skill-biased technological change, i.e., the investment-specific technology change, and evaluates the effects of this particular type of skill-biased technological change on the evolution of the skill premium. Third, this paper endogenizes the choice of college entrance. It therefore can shed light on the impact of two factors on the evolution of educational attainment and hence the relative supply of skilled labor.

We 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 in the U.S. over the postwar period. We find that demographic change dwarfs ISTC before the late 1960s and does contribute to the decline in the skill premium in the 1970s. However, after the late 1970s ISTC takes over and drives the dramatic increase in the skill premium. In addition, since the model endogenizes the college entry decision as well, we can also investigate the quantitative effect of these two exogenous forces on the evolution of the college enrollment rate. We find that ISTC can explain about 35% of the increase in the enrollment rate for the period 1951-2000, while demographic change does not have a significant effect on the enrollment rate over time.

In this model, ISTC affects the skill premium and college enrollment rate through a simple economic
mechanism. 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 and encourages individuals to go to college.

Demographic change affects the skill premium and college enrollment rate through a different channel. Growing birth cohorts (“baby boom”) change the age structure in the economy and skew it toward younger (college-age 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 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. Therefore, the total effect of demographic change on the skill premium and the enrollment rate is ambiguous and has to be investigated quantitatively.

In quantitative terms, before the late 1960s the U.S. had undergone dramatic population growth (see Figures 2 and 3), while the change in ISTC was only moderate (see Figure 9). 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 that of 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 such as KORV (2000). Compared to KORV (2000), our extension lies in two dimensions. First, we embed ISTC and capital-skill complementarity in a dynamic general equilibrium framework. Second, we endogenize the relative supply of skilled labor by modelling the college attendance choice. Therefore, the model is able to capture the dynamic interaction between the skill premium and the relative supply of skilled labor. As far as we know, this is the first paper to use a general equilibrium model with ISTC and endogenous college entrance decision to analyze the evolution of the skill premium in the U.S.

In spirit, this paper is also close to Heckman, Lochner, and Taber (1998). They develop and estimate a general equilibrium overlapping generations 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 microfoundation for the source of skill-biased technological change (i.e., ISTC) as in this paper. They also do not ask the research question about the quantitative decomposition of the impacts of ISTC and demographic change on

---

2 Restuccia and Vandenbroucke (2010) examine the quantitative contribution of changes in the return to schooling, which imply skill-biased technological change, in explaining the evolution of educational attainment in the U.S. from 1940 to 2000 in a similar model with discrete schooling choice. Their model, however, does not include capital accumulation and capital-skill complementarity and hence cannot address the effect of ISTC on the skill premium via physical capital accumulation.

3 Several 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 Kurucu (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 microfoundation 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.
the skill premium.\textsuperscript{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 to allow for a better model of educational attainment. The model also includes a time-varying cohort size which can be used to quantify the effect of demographic change and allows for a better analysis of labor supply. Both elements are 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. Section 3 presents the economic model of the decision to go to college, describes the market environment, and defines the general equilibrium in the model economy. Section 4 shows how to parameterize the model economy. Section 5 provides calibration results for the pre-1951 steady state and shows some comparative static results. 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 ISTC and demographic change on the skill premium. Section 7 discusses the limitation of the benchmark model in replicating the college attendance rate and conducts a counterfactual experiment to isolate the effect of ISTC only in a general equilibrium framework by inputting the relative supply of skilled labor exogenously. Section 8 provides several sensitivity analyses. Finally, Section 9 concludes.

2 Stylized Facts

This section summarizes the data pattern regarding the college choice that determines the supply of skilled labor. See Appendix A for the source and construction of the data. 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 an 18-year lag version of U.S. fertility growth; that is, it reflects the baby boom and baby bust.\textsuperscript{5}

Figure 3 measures the college-age population. We report the population of age 18 to 21 in the U.S. since 1955. These series follow a pattern similar to that in Figure 2. The baby boom pushed the college-age population up until the fertility rate reached its peak around 1960, corresponding to the peak of the college-age 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

\textsuperscript{4}As a notable exception, Lee and Wolpin (2010) use a two-sector structural model similar to the one in Lee and Wolpin (2006) to account for wage and employment changes in the U.S. from 1968 to 2000. They include exogenous factors on the labor supply side such as demographic and population changes and the demand side factors such as skill-biased technical change, capital-skill complementarity and changes in relative product-market prices. They evaluate quantitatively the effects of these factors on wage inequality, education premium, occupation and industry wage and employment difference, male-female difference of labor market outcomes and the effects on sectoral outputs. Their model, however, does not include asset accumulation. Therefore, it cannot address the interaction between demographic change and asset holdings as I emphasized above. In other words, their model probably underestimates the effects of demographic change. As shown in Figure 7 in their paper, without the counter effect from demographic change, the model predicts a skill premium that is increasing even in the 1970s.

\textsuperscript{5}Cohort size has been increasing since 1995 because the baby boomers’ children reached college age around the mid-1990s.
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 were a decade of depressed college enrollment, and it was not until 1985 that the enrollment rate exceeded the level in 1968. Starting in 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 Figures 3 and 4, we can see that the dramatic increase in the relative supply of skilled labor in the 1970s is due to a combination of a rising college-age population and a rising college enrollment rate in the 1960s. The demographic change affects only the cohort size of the college-age population. The higher enrollment rate shifts the proportion of the college-age population into the skilled labor pool. 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 decision to go to college 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 us 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 and returns to skill. In order to understand the evolution of the 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 going to college. In Figure 5 we 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 have increased dramatically.

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

1. The skill premium rose during the 1950s and 1960s, fell in the 1970s, 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 pattern similar to that of the skill premium, as do tuition payments.

The stylized facts about the skill premium observed in Figure 1 are the target of this paper. To answer the quantitative question raised in Section 1, we will take the demographic change in Figure 2 and the measured investment-specific technological change in Figure 9 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, we 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.\(^6\)

3.1 Demographics

The economy is populated by overlapping generations. Individuals enter the economy when they are 18 years old and finish high school, which we call the birth cohort and model as age \(j = 1\). We assume individuals work up to age \(J\), which is the maximum life span. The model period is one year. To distinguish between the age of a cohort and calendar time, we use \(j\) for age, and \(t\) for calendar time. For example, \(N_{j,t}\) is the population size of the age-\(j\) cohort at time \(t\).

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

\[
N_{1,t} = (1 + n_t)N_{1,t-1}. \tag{1}
\]

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

\[
\eta_{j,t} = \frac{N_{j,t}}{N_t} = \frac{N_{j,t}}{\sum_{i=1}^{J} N_{it}}. \tag{2}
\]

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 the 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

\[
u(c_{j,t+j-1}) = \frac{c_{j,t+j-1}^{1-\sigma}}{1-\sigma}. \tag{3}\]

The parameter \(\sigma\) 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.

\(^6\)He (2011) uses a partial equilibrium version of the model (without the production side) to study the quantitative effect of change in the female college wage premium on the change in female college enrollment rate.
### 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. We 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 $p$. We assume that an individual who enters college will successfully graduate from college. 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 simplicity, we assume there is no college dropout and no unemployment.

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

$$
\begin{align*}
    c_{j,t+j-1} + p_{t+j-1} + a_{j,t+j-1} & \leq (1 + r_{t+j-1})a_{j-1,t+j-2} \quad \forall j = 1, 2, 3, 4 \\
    c_{j,t+j-1} + a_{j,t+j-1} & \leq (1 + r_{t+j-1})a_{j-1,t+j-2} + w_{t+j-1}^c \quad \forall j = 5, \ldots, J \\
    c_{j,t+j-1} & \geq 0, a_{0,t-1} = 0, a_{f,t+f-1} \geq 0,
\end{align*}
$$

where $\{\epsilon_j^c\}_{j=1}^J$ 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

$$
\begin{align*}
    c_{j,t+j-1} + a_{j,t+j-1} & \leq (1 + r_{t+j-1})a_{j-1,t+j-2} + w_{t+j-1}^h \quad \forall j = 1, \ldots, J \\
    c_{j,t+j-1} & \geq 0, a_{0,t-1} = 0, a_{f,t+f-1} \geq 0.
\end{align*}
$$

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

### 3.4 Schooling Choice

Next, we would like to explicitly model an individual’s schooling choice. In order to generate a positive enrollment rate in the model, we 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), we assume that different individuals within each birth cohort are endowed with different levels of the disutility cost of schooling. We index people by their disutility level $i \in [0, 1]$, and the associated disutility cost that individual $i$ bears is represented by $\chi(i)$. We assume $\chi'(i) < 0$.\(^8\) The cumulative distribution function (CDF) of the disutility cost is denoted by $F$, $F(i_0) = \int \chi(i) \, di$.

\(^7\)Notice 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.

\(^8\)Navarro (2007) finds that ability is the main determinant of this “psychic” cost, and it plays a key role in determining schooling decisions. High-ability individuals face a very low disutility cost, while low-ability individuals face a large disutility cost of attending college. Therefore, we can also view $i$ as the index of individuals’ “learning ability."
where savings. We might need a social planner to do so. That increases significantly the complexity of the model and the computation.

For the capital-skill complementarity, we require skilled labor combination and unskilled labor is production technology exhibits constant returns to scale. The elasticity of substitution between the capital-

We 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), we adopt an aggregate production function with capital-skill complementarity as follows:

\[
Y_t = A_t F(K_t, S_t, U_t) = A_t \mu U_t^\theta + (1-\mu)(\lambda(K_t)^\rho + (1-\lambda)S_t^\theta)^{\frac{\theta}{\rho}} \quad (8)
\]

where \( A_t \) is the level of total factor productivity \( \text{TFP} \). We also have \( 0 < \lambda, \mu < 1 \), and \( \rho, \theta < 1 \). This production technology exhibits constant returns to scale. The elasticity of substitution between the capital-skilled labor combination and unskilled labor is \( \frac{1}{\theta-\rho} \) and the one between capital and skilled labor is \( \frac{1}{1-\theta} \).

For the capital-skill complementarity, we require \( \frac{1}{\theta-\rho} < \frac{1}{1-\theta} \), which means \( \rho < \theta \).

\text{Griliches (1969) provides evidence from U.S. manufacturing data that skill is more complementary to capital than to 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.}

\text{The difference between our production function and the one in KORV (2000) is that we do not distinguish between structures \( (K_s) \) and equipment \( (K_e) \) as in KORV (2000), so the capital \( K \) in our model is just the total capital stock which is the sum of capital equipment and structures. The reason why we use \( K \) instead of differentiated \( K_s \) and \( K_e \) is in our decentralized OLG setting the capital is provided through individuals’ savings. Individuals cannot distinguish equipment and structures from their savings. We might need a social planner to do so. That increases significantly the complexity of the model and the computation.}
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), we 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.

Defining \( \tilde{K}_{t+1} \equiv \frac{K_{t+1}}{q_t} \), it is easy to show that this economy is equivalent to the model with the following production function and capital accumulation

\[
Y_t = A_t [\mu U_t^\theta + (1 - \mu)(\lambda(B_t \tilde{K}_t) + (1 - \lambda)S_t^\gamma)]^{1/\theta}
\]

\[
\tilde{K}_{t+1} = (1 - \tilde{\delta})\tilde{K}_t + X_t
\]

with \( \tilde{\delta} = 1 - (1 - \delta)\frac{2\gamma}{\theta} \) and \( B_t = q_{t-1} \). This transformation maps changes in ISTC into the changes in the capital productivity level \( B_t \).\(^{11}\) It simplifies the computation of the model. From now on, we refer to this transformed version of the model as the benchmark model.

Based on the transformation, the representative firm’s profit maximization implies the first-order conditions as follows:

\[
r_t = \lambda(1 - \mu)A_t B_t^\rho H_t(\lambda(B_t \tilde{K}_t) + (1 - \lambda)S_t^\gamma)\tilde{K}_t^{\gamma - 1} - \tilde{\delta},
\]

\[
w_t^c = (1 - \mu)(1 - \lambda)A_t H_t(\lambda(B_t \tilde{K}_t) + (1 - \lambda)S_t^\gamma)\tilde{K}_t^{\gamma - 1}S_t^{\gamma - 1},
\]

\[
w_t^h = \mu A_t H_t U_t^{\theta - 1},
\]

where \( H_t = [\mu U_t^\theta + (1 - \mu)(\lambda(B_t \tilde{K}_t) + (1 - \lambda)S_t^\gamma)]^{\frac{\theta - 1}{\gamma}}.\)

Dividing (10) by (11), we derive the expression for the skill premium:

\[
\frac{w_t^c}{w_t^h} = \frac{(1 - \mu)(1 - \lambda)}{\mu} [\lambda \frac{B_t \tilde{K}_t}{S_t}]^\rho + (1 - \lambda)]^{\frac{\theta - 2}{\gamma}} [\tilde{S}_t]^{\theta - 1}.
\]

Log-linearizing (12), and using a “hat” to denote the rate of change between time \((\hat{X} = \frac{\Delta X}{X})\), we obtain (ignoring time subscripts for convenience)

\[
(\frac{\hat{w}^c}{\hat{w}^h}) = \lambda(\theta - \rho)(\frac{B\hat{K}}{S})^\rho [\hat{B} + \hat{K} - \hat{S}] + (\theta - 1)[\hat{S} - \hat{U}].
\]

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 the “relative quantity effect” in KORV (2000). The other term \( \lambda(\theta - \rho)(\frac{B\hat{K}}{S})^\rho [\hat{B} + \hat{K} - \hat{S}] \) 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 trade-off between these two effects.

\(^{11}\) This transformation is motivated by GHK (1997). See Appendix B in GHK (1997) for a similar transformation for their benchmark economy.
The transformation above maps ISTC $q_t$ into the changes in the capital productivity level $B_t$. Therefore, increases in $q_t$ will be transformed into increases in $B_t$. As shown in equation (13), when $B_t$ increases, through the capital-skill complementarity effect, it will raise the skill premium. ISTC thus is also skill-biased. The paper therefore provides a micro-foundation of unmeasured skill-biased technological change.

Finally, the resource constraint in the economy is given by

$$C_t + P_t + X_t = Y_t,$$

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

### 3.6 The Competitive Equilibrium

The model above is a standard OLG setting with discrete schooling choices. We assume that individuals have perfect foresight about the paths of exogenous changes $\{n_t\}$ and $\{q_t\}$. Suppose an individual $i$ born at time $t$ has already made the schooling decision $s_{i,t}$. Conditional on this choice, we can present her utility maximization problem in terms of a dynamic programming representation.

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

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

subject to (4)-(5).

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

$$V^h_{t+j-1}(a_{j-1,t+j-2},j) = \max_{\{c_{j,t+j-1},a_{j,t+j-1}\}}\{u(c_{j,t+j-1}) + \beta V^h_{t+j}(a_{j,t+j-1},j + 1)\}$$

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

$$s_{i,t} = c \quad \text{if } V^c_t(a_{0,t-1} = 0,1) - DIS(i) > V^h_t(a_{0,t-1} = 0,1),$$

$$s_{i,t} = h \quad \text{if } V^c_t(a_{0,t-1} = 0,1) - DIS(i) < V^h_t(a_{0,t-1} = 0,1),$$

$$s_{i,t} = \text{indifferent} \quad \text{if } V^c_t(a_{0,t-1} = 0,1) - \chi(i) = V^h_t(a_{0,t-1} = 0,1).$$

Based on the individual’s dynamic program and the schooling choice criterion above, the definition of the competitive equilibrium in this model economy is as follows.

**Definition 1** Let $\mathcal{A} = R$, $\mathcal{S} = \{c,h\}$, $\mathcal{J} = \{1,2,\ldots,J\}$, $\mathcal{D} = [0,1]$, and $\mathcal{T} = \{1,2,\ldots,T\}$. Given the age structure $\{n_{i,t}\}_{i=1}^J$, a competitive equilibrium is a sequence of individual value functions $V^*_t : \mathcal{A} \times \mathcal{J} \to R$; individual consumption decision rules $C^*_t : \mathcal{A} \times \mathcal{J} \to R_+$; individual saving decision rules $A^*_t : \mathcal{A} \times \mathcal{J} \to A$.

---

12A perfect 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 the TFP growth rate to investigate its impact on the postwar Japanese saving rate. Their sensitivity analysis shows that alternative expectations hypotheses do not significantly change the quantitative results.
for \( s \in S \) and \( t \in T \); an individual \( i \)'s period 1 schooling choice \( s^*_i,t \) for \( s \in S \), \( i \in D \), and \( t \in T \); 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 \( \{w^c_i, w^b_i, r_t\}_{t=1}^T \); and a sequence of measures of individual distribution over age and assets \( \lambda^s_t : A \times J \rightarrow R_+ \) for \( s \in S \) and \( t \in T \) such that:

1. Given prices \( \{w^c_i, w^b_i, r_t\} \), the individual decision rules \( C^s_t \) and \( A^s_t \) solve the individual dynamic problems (14) and (15).

2. Optimal schooling choice \( s^*_i,t \) is the solution to the schooling choice criterion in (16) for each individual \( i \).

3. Prices \( \{w^c_i, w^b_i, r_t\} \) are the solutions to the firm's profit maximization problem (9)-(11).

4. The time-variant age-dependent distribution of individuals choosing \( s \) follows the law of motion

\[
\lambda^s_{t+1}(a', j + 1) = \sum_{a:a' \in A^s_t(a,j)} \lambda^s_t(a, j).
\]

5. Individual and aggregate behaviors are consistent

\[
K_t = \sum_j \sum_a \sum_s \eta_{j,t} \lambda^s_t(a, j) A^s_t(a, j - 1),
\]

\[
S_t = \sum_j \sum_a \eta_{j,t} \lambda^s_t(a, j) \epsilon^c_j,
\]

\[
U_t = \sum_j \sum_a \eta_{j,t} \lambda^s_t(a, j) \epsilon^b_j.
\]

6. The goods market clears

\[
\sum_{j=1}^J \sum_a \sum_s \eta_{j,t} \lambda^s_t(a, j) C^s_t(a, j) + \sum_{j=1}^4 \sum_a \eta_{j,t} \lambda^s_t(a, j) p_{j,t} + X_t = Y_t,
\]

or

\[
C_t + P_t + X_t = Y_t.
\]

When ISTC and demographic change both stabilize at some constant level, 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-dependent but time-invariant. Therefore, we can define the stationary competitive equilibrium accordingly.

4 Parameterization

In this section, we 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.\textsuperscript{13} It is worth emphasizing that the evolution of the skill premium over the 1951-2000 period is not a target of calibration but the goal for evaluating the model's performance.

\textsuperscript{13}We choose the U.S. economy from 1947 to 1951 as the initial steady state based on the observations that both the ISTC and the demographic changes were quite stable during this time period.
4.1 Cohort-Specific Skill Premium

The skill premium data we report in Figure 1 are 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 $\left\{ \frac{w_{t+j-1}^{i}}{w_{t+j-1}^{i+1}} \right\}_{j=1}^{J}$. In order to understand the mechanism of the schooling decision for each cohort, and more important, as will be shown later, to help pin down the distribution of the disutility cost $\chi(i)$, we need to find the data counterpart of this cohort-specific skill premium.

We 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. (We choose to end the sample in 1991 because of the quality of the data—the 1991 cohort has only 12-year HSG wage and 8-year CG wage data.) In order to make our results comparable to the literature, we 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, we look only at high school graduates (HSG) and college graduates (CG). The wage here is the annualized wage and salary earnings and the personal consumption expenditure deflator from NIPA is used to convert all wages to constant 2002 dollars.

Since the CPS is not a panel data set, theoretically speaking, we cannot track specific cohorts from it. However, since it is a repeated cross-sectional data set, we can use a so-called “synthetic cohort construction method” to construct a proxy of a cohort’s specific skill premium. Using this method repeatedly for each birth cohort, we 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). We use an econometric method to predict the mean wage at those specific age points and extrapolate the missing data. We predict them by either second- or third-order polynomials, or a conditional Mincer equation as follows

$$\log[HSG\text{wage}(age)] = \beta_0^h + \beta_1^h \text{experience}_h + \beta_2^h \text{experience}_h^2 + \epsilon^h, \text{ experience}_h=\text{age}-18$$

$$\log[CG\text{wage}(age)] = \beta_0^c + \beta_1^c \text{experience}_c + \beta_2^c \text{experience}_c^2 + \epsilon^c, \text{ experience}_c=\text{age}-22.$$ 

The criterion is basically the goodness of fit. We 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, Figure 6 shows a complete life cycle wage profile of HSGs and CGs for the 1975 cohort by using the prediction from third-order polynomials.

4.2 Distribution of the Disutility Cost

The distribution of disutility cost $\chi(i)$ 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 schooling choice criterion embodied in (16) actually sheds some light on how to compute the distribution of the disutility.

---

14 He (2011) uses the same approach to construct the female cohort-specific skill premium and studies its quantitative impact on female college entry decision.
cost. Note that the person \( i^* \) who is indifferent between going to college or not has

\[
V^c_t(a_{0,t-1} = 0, 1) - \chi(i^*) = V^h_t(a_{0,t-1} = 0, 1),
\]

that is, her disutility cost is exactly the difference between two conditional value functions. Since the disutility cost is a decreasing function of index \( i \), individuals with disutility index \( i > i^* \) go to college. Therefore, for a specific cohort \( t \), if we calculate the difference between two conditional value functions

\[
V^c_t(a_{0,t-1} = 0, 1) - V^h_t(a_{0,t-1} = 0, 1),
\]

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 the disutility cost. In this way, we can pin down one point on the CDF of the disutility cost. Applying this procedure to different cohorts will give us a picture of how the disutility cost is distributed.\(^{15}\)

Figure 7 shows the CDF of the disutility cost that is used in the benchmark model. Estimating this CDF function involves a fixed-point algorithm. We first guess the interest rate \( r \) under the calibrated discount parameter \( \beta \) and the calibrated preference parameter \( \sigma \). For each cohort born at time \( t \), we normalize the 18-year-old HSG wage (which is \( w_{t,1}^h \) in the model) to one and input the normalized cohort-specific lifetime wage profiles for both HSGs and CGs from the data constructed in Section 4.1. We go through the backward induction of the Bellman equation as described in Section 3.6 to obtain the value function difference

\[
V^c_t(a_{0,t-1} = 0, 1) - V^h_t(a_{0,t-1} = 0, 1)
\]

and hence the cut-off disutility cost for every cohort \( t \) from 1948 to 1991. By plotting them on \( x \)-axis against enrollment rate data in the same time range (1948-1991) on \( y \)-axis, we have 44 points on the possible CDF of the disutility cost. By assuming that the disutility costs follow a normal distribution, we then estimate the CDF function.\(^{16}\) Later in the computation of the stationary equilibrium, during each iteration when we obtain factor prices \( \{w^c, w^h, r\} \), we can conduct backward induction of a Bellman equation loop to obtain the conditional value functions. Feeding the difference between these two functions into the estimated CDF, we get the corresponding enrollment rate. We then check if the equilibrium interest rate \( r \) is the same as the guess. If it is not, we have to change the guess on \( r \) and redo the procedure. The iteration stops when the initial guess on \( r \) used in calculating CDF converges to the equilibrium interest rate in the steady state.

### 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 \( n \) that is used in the initial steady state is calculated as the average growth rate of the HSG cohort size from 1948 to 1951, which is 0%.

\(^{15}\)Here we assume the distribution of the disutility cost is stationary.

\(^{16}\)Heckman, Lochner, and Taber (1998) also assume that the “nonpecuniary benefit of attending college” is normally distributed. A more flexible Beta distribution yields a very similar estimated CDF as a normal distribution within the reasonable range of the disutility cost.
4.4 Preferences and Endowments

We 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).

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

\[
\varepsilon^h_j = \frac{HSGwage_j}{HSGwage}, \quad \varepsilon^c_j = \frac{CGwage_j}{CGwage}, \quad \forall j = 1, \ldots, 48.
\]

The result is shown in Figure 8. Both proﬁles exhibit a clear hump shape and reach a peak around age 55. Also notice that \( \varepsilon^c_j = 0, \forall j = 1, \ldots, 4 \), since we assume CGs never work while in school.

Tuition for the 1951 cohort is the real TFRB charges from 1951 to 1954 as shown in Figure 5. We divide them by the data of real labor income of age 18 unskilled workers in 1951 and thus convert these four-year tuitions into four ratios. The ratios are then inputted into the model and are multiplied by the model-generated real labor income of age 18 unskilled labor \( w^h_{1} \) to convert back to the model counterpart of the tuitions.

4.5 Production 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 that 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. ISTC \( q \) is also normalized to one. We set the depreciation rate of capital \( \delta \) to 0.069 by following İmrohoroğlu, İmrohoroğlu, and Joines (1999), who calculate this parameter from annual U.S. data since 1954. Since ISTC stabilizes in the initial steady state, the transformed depreciation rate \( \tilde{\delta} \) is equal to \( \delta \).

4.6 ISTC

Following GHK (1997) and KORV (2000), in the benchmark 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, we can use the relative price of capital to identify ISTC \( q \). We take the price index of personal nondurable 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. We then divide these two sequences to obtain the data counterpart of \( q \). Finally we normalize the level of \( q \) in 1951 to be one. Figure 9 shows the natural logarithm 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%).
Table 1: Parameter Values in the Benchmark Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value and Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J$</td>
<td>maximum life span</td>
<td>48, corresponding to age 65 in real life</td>
</tr>
<tr>
<td>${\varepsilon^s_j}_{j=1}^J, s = c, h$</td>
<td>age efficiency profiles</td>
<td>1962-2003 CPS, and 1950, 1960 census</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>CRRA coefficient</td>
<td>1.5, Gourinchas and Parker (2002)</td>
</tr>
<tr>
<td>$\theta$</td>
<td>elasticity b/w $U$ and $K$</td>
<td>0.401, KORV (2000)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>elasticity b/w $S$ and $K$</td>
<td>-0.495, KORV (2000)</td>
</tr>
<tr>
<td>$\delta$</td>
<td>depreciation rate</td>
<td>0.069, Imrohorghlu et al. (1999)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>discount rate</td>
<td>1.027</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>share of $K$ in production</td>
<td>0.645</td>
</tr>
<tr>
<td>$\mu$</td>
<td>share of $U$ in production</td>
<td>0.415</td>
</tr>
<tr>
<td>$sd$</td>
<td>scale factor of disutility cost</td>
<td>2.90</td>
</tr>
</tbody>
</table>

This leaves four parameter values to be calibrated: the subjective discount rate $\beta$, the income share of capital in the capital-skilled labor combination $\lambda$, the income share of unskilled labor $\mu$, and the scale factor of the disutility cost $sd$ (see Appendix B for details).\footnote{As shown in Section 4.2, since we use the estimated CDF of the disutility cost along with the computed value function difference to determine the enrollment rate in the simulation, there is a scale difference between the CDF that is estimated from the data and the value function difference that is computed based on the parameter values in the model. Scale factor $sd$ is used to take care of this scale difference.} We calibrate these four parameters so that the model can replicate, as closely as possible, four moment conditions in the data for the period 1947-1951. These four moment conditions are:

1. Average capital-output ratio 2.67 from 1947 to 1951 (NIPA data).
2. Average income share of labor 72.43% from 1947 to 1951 (NIPA data).
3. Average skill premium 1.4556 in 1949 (census data).
4. Average college enrollment rate 41.54% from 1947 to 1951.

This exercise ends up with $\beta = 1.027$, $\lambda = 0.645$, $\mu = 0.415$, and $sd = 2.90$. Table 1 summarizes the parameters used in the model.

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

5 Steady-State Results

5.1 Initial Steady State

In this section, we 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 2.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Model</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{w^c}{w^h}$</td>
<td>1.4543 (construction)</td>
<td>1.4556 (1949 census data)</td>
</tr>
<tr>
<td>$e$</td>
<td>41.61% (construction)</td>
<td>41.54% (1947-51 average)</td>
</tr>
<tr>
<td>$K/Y$</td>
<td>2.72 (construction)</td>
<td>2.67 (1947-51 average)</td>
</tr>
<tr>
<td>$(w^hU + w^cS) / Y$</td>
<td>72.17% (construction)</td>
<td>72.43% (1947-51 average)</td>
</tr>
<tr>
<td>$C/Y$</td>
<td>80.65%</td>
<td>79.57% (1947-51 average)</td>
</tr>
<tr>
<td>$X/Y$</td>
<td>18.80%</td>
<td>20.17% (1947-51 average)</td>
</tr>
<tr>
<td>$r$</td>
<td>3.33%</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Macro Aggregates in the Benchmark Economy: Initial Steady State

<table>
<thead>
<tr>
<th>$n$ (%)</th>
<th>$w^c/w^h$</th>
<th>$e$ (%)</th>
<th>$S/U$ (%)</th>
<th>$BK/S$</th>
<th>rel. supply effect</th>
<th>K-S comple. effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (benchmark)</td>
<td>1.4543</td>
<td>41.61</td>
<td>67.99</td>
<td>5.58</td>
<td>1.5408</td>
<td>1.4240</td>
</tr>
<tr>
<td>4.06</td>
<td>1.5105</td>
<td>41.96</td>
<td>61.70</td>
<td>5.30</td>
<td>1.5408</td>
<td>1.4240</td>
</tr>
<tr>
<td>-1.57</td>
<td>1.4412</td>
<td>41.52</td>
<td>69.74</td>
<td>5.68</td>
<td>1.4318</td>
<td>1.4631</td>
</tr>
</tbody>
</table>

Table 3: Effect of Population Growth on Steady State

The simulations show that the model does well in matching the data. It matches our targets—skill premium ($\frac{w^c}{w^h}$), enrollment rate ($e$), capital-output ratio ($K/Y$), and labor income-output ratio ($\frac{(w^hU + w^cS)}{Y}$) by construction. Additionally, several key macro aggregate ratios, such as the consumption-output ratio ($C/Y$) and the investment-output ratio ($X/Y$) are also in line with the U.S. average data. The risk-free real interest rate is 3.33%.

5.2 Comparative Static Experiments

In this section, we 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 $B$ in the steady state. In other words, we compare steady states between different economies with different growth rates of $n$ and $B$, respectively, while keeping other parameters unchanged as in the benchmark case. We summarize the corresponding results in Tables 3 and 4, respectively. In Table 3, 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, skewing it toward younger cohorts. Keeping the enrollment rate unchanged, more individuals from the college-age cohort stay in college. Meanwhile, more people from the college-age 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 shown in Table 3.\(^{18}\) When $n$

\(^{18}\)To formalize this idea, consider a two period OLG model with young and old agents, in which becoming skilled takes one period. Assuming the enrollment rate $e$ is the constant propensity to go to college among young agents and $n$ is the population.
increases to around 4%, the relative supply of skilled labor $S/U$ decreases by 6.3%. 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. People accumulate fewer assets during their early working years. A shift toward younger cohorts in the demographic structure thus decreases the incentive to accumulate assets in the economy. As a result, the capital-output ratio ($K/Y$) decreases from 2.72 in the benchmark case to 2.52 in the $n = 4.06\%$ case. It also leads to a decrease in the effective capital-skilled labor ratio ($BK/S$). Then, through the capital-skill complementarity effect, it tends to decrease the skill premium. We can further disentangle the relative quantity and capital-skill complementarity effect as in equation (12). Suppose that $S/U$ changes from the benchmark value 67.99% to the one in the $n = 4.06\%$ case, while we keep $BK/S$ the same as in the benchmark case. Equation (12) would imply that the skill premium increases from 1.4543 in the benchmark case to 1.5408. On the other hand, suppose $S/U$ remains unchanged, while we change the effective capital-skilled labor ratio $BK/S$ from its benchmark value 5.58 to its value under $n = 4.06\%$. Then equation (12) would imply that the skill premium decreases from its benchmark value to 1.4240 via only the capital-skill complementarity effect. Quantitatively, the impact of the demographic change on the relative supply of skilled labor dominates that on the relative demand for skilled labor through capital-skill complementarity. Thus, the skill premium increases to 1.5105, which is closer to the one implied by the pure relative supply effect. 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 to accumulate assets. These two impacts again tend to offset each other. Quantitatively, a change in $n$ from 0% to -1.57% slightly decreases the skill premium. Again this is in the same direction as the pure relative supply effect, which shows that a change in $n$ affects the relative supply of skilled labor more significantly.

Since an increase in $n$ raises the skill premium, it also increases the benefits of going to college. The experiment shows that under the $n = 4.06\%$ case, the college enrollment rate increases from 41.61% in the benchmark case to 41.96%. On the other hand, a decrease of $n$ from 0% to -1.57% lowers the enrollment rate from 41.61% to 41.52%. Compared to the impact on the skill premium, the effect of demographic change on the college enrollment rate seems very insignificant.

Next, we show the effect of a permanent change in $q$ on the steady state. As shown in Figure 9, ISTC $q_t$ has increased from 1 in 1951 to 3.89 in 2000. In other words, ISTC has been increasing almost four times over this period. As mentioned in Section 3.5, we can map the sequence of $q_t$ to the changes in the capital productivity level $B_t$, which translates into a change of capital productivity $B_t$ from its initial steady-state value of one in 1951 to 3.73 in 2000. Suppose that the U.S. economy reaches the steady state again after 2000. Keeping other things equal, Table 4 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, $B_tK_t$. Since capital is complementary to skilled labor, increases in effective capital also raise the demand for skilled labor. As evidence, when $B$-
increases from one to 3.73, the effective capital-skilled labor ratio \((BK/S)\) increases from 5.58 to 20.47, almost 4 times higher. The increase in \(BK/S\) alone would raise the skill premium from 1.4543 to 2.2133. However, a rising skill premium gives individuals a stronger incentive to go to college. The enrollment rate increases from 41.61% to 48.33% and hence raises the relative supply of skilled labor from 67.99% to 89.23%. The change in the relative supply of skilled labor alone would reduce the skill premium from 1.4543 to 1.2351.

The quantitative results in Table 4 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 1.8805 in the model, which is quite close to 1.8357 in the data.

### 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 assess the model’s performance on matching 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, İmrohoroğlu, and İmrohoroğlu (2006) and Conesa and Kreuger (1999), we compute the model along a transition path from the initial pre-1951 steady state toward 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, we 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 \(\{n_t\}_{t=1951}^{2000}\). We also feed in the normalized tuition payments \(\{p_t\}_{t=1951}^{2000}\). We then assume that capital-specific technological change keeps increasing after 2000 at the average growth rate of the period 1951-2000 until 2020, and it stabilizes at this constant level until 2050. For simplicity, we also assume that after 2000 there is no demographic change and that the tuition payment is constant at the 2000 level. Since we want to focus on the effect of ISTC, the neutral TFP change has been normalized to unity for all time periods.

<table>
<thead>
<tr>
<th>(B)</th>
<th>(w^e/w^h)</th>
<th>(e) (%)</th>
<th>(S/U) (%)</th>
<th>(BK/S)</th>
<th>rela. supply effect</th>
<th>K-S comple. effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (benchmark)</td>
<td>1.4543</td>
<td>41.61</td>
<td>67.99</td>
<td>5.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.73 (2000 level)</td>
<td>1.8805</td>
<td>48.33</td>
<td>89.23</td>
<td>20.47</td>
<td>1.2351</td>
<td>2.2133</td>
</tr>
</tbody>
</table>

Table 4: Effect of Investment-Specific Technological Change on Steady State

---

19 As described in Section 4.4 for the steady state, the four-year tuitions faced by each cohort are normalized by dividing them by the age 18 unskilled worker’s annual wage of the same cohort in the data. The four ratios then are inputted into the model and are multiplied by the model-generated real labor income of age 18 unskilled labor \(w_{18}^e\) to convert back to the model counterpart of the tuitions.

20 The results are actually not sensitive to the timing of the final steady state. See Section 7.4 for details.
through the transition.\footnote{See Section 7.2 for the sensitivity analysis when the variable neutral TFP change is allowed.} We 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 10. Notice that since the skill premium and the college enrollment rate in 1951 have been used for calibration, we can only tell the performance of the model by comparing the model-generated variables with the data after 1951.

In Figure 10, the simulated skill premium from the benchmark economy captures the increasing trend over the period 1951-2000. From 1951 to 2000, the data show that the skill premium increases from 1.4556 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 1.7774 and the average growth rate is 0.43%. In other words, the model captures about 85% of the increase in the level of the skill premium over the period 1951-2000. Focusing on 1951 to 1960, the average growth rate of the skill premium in the data is 0.91% (we have only two data points), while in the model it is -0.08%. The model does not capture the increase through that decade. From 1963, we have annual data for the skill premium so the comparison between the data and the model’s performance is more accurate. From 1963 to 1970, the data show the skill premium increases from 1.4427 to 1.5047 at an average growth rate of 0.63%, while the model predicts an increase from 1.4075 to 1.5608 at an average growth rate of 0.67%. From 1971 to 1980, the skill premium first decreases sharply but it has started to increase since 1975. The average growth rate is almost zero at 0.07%. The model, however, overshoots the data and misses this decline by predicting an almost flat skill premium over this period. The average growth rate is 0.06%. The skill premium starts to increase dramatically beginning in 1981 in the data. From 1981 to 2000, the skill premium increases from 1.4730 to 1.8357 at an average growth rate of 1.00%, while the model predicts an increase from 1.5696 to 1.7774 at an average growth rate of 0.63%. The model captures 56% of the increase in the skill premium during these two decades. Overall, for the three episodes in the “N” shape of the skill premium from 1963 to 2000, the model captures the increasing trend of the changes in the skill premium for the period 1963-1970 and 1981-2000. But it fails to replicate the declining part of the skill premium and significantly overshoots the data in the 1970s.

The model also raises the enrollment rate from 41.61% to 49.11% for the period 1951-2000, while the enrollment rate in the data increases from 41.54% to 63.3%, as shown in Figure 11. In other words, the benchmark model can explain about 35% of the increase in the enrollment rate during this period. Matching the college enrollment rate is not the target of the current model and understanding why the enrollment rate has changed over time is a deeper goal, which is beyond the scope 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 changes (e.g., the GI Bill and the Vietnam War) and the shifts in social norms that have especially affected women’s decision to go to college (see Goldin 2006 and He 2011).\footnote{A possible reason why the model does not capture the enrollment rate data well is the model does not allow college drop-out and/or two-year college program. Individuals who enroll in the model will finish college in four years. The data counterpart of the enrollment rate in the model therefore should be college completion rate instead of college enrollment rate. The growth in college degree conferred has been much slower than the growth in college enrollment (see Turner 2004), which allows the model to have some potential in matching the enrollment rate data better.} However, with such a highly stylized model in which the main determinant of college attendance is the net present value of future wages, two exogenous driving forces can still capture more than one-third of the increase in the college enrollment rate from 1951 to 2000. Of course, a natural question would be: which driving force has a bigger impact on
the skill premium and the college enrollment rate? We leave that to the following exercise.

6.2 Counterfactual Decomposition

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

We first shut down the investment-specific technological change by setting $B_t = 1$, so the only exogenous force remaining is the demographic change. The results are shown in Figure 12 with the legend “Demographic only.” Figure 12 shows that the model fits the skill premium data fairly well until 1982. And more important, it generates the declining skill premium during the 1970s. However, it cannot capture the dramatic increase since 1980 as shown in the data. More specifically, from 1963 to 1970, the model generates a 0.86% average growth rate in the skill premium, which can explain 138% of the average growth rate in the data for this period. From 1971 to 1980, the data show that the skill premium decreases first and then increases with an average growth rate close to zero over the whole decade, while the model generates an average growth rate of -0.17%. More important, the skill premium in the model decreases from 1.4994 in 1971 to 1.4791 in 1980, which fits the data much better than the benchmark case. However, from 1981, the model predicts a slight decrease in the skill premium (-0.02% per year), in contrast to the dramatic increase shown in the data.

When ISTC is shut down, the model generates little variation in the enrollment rate. As shown in Figure 12 with the legend “Demographic only,” from 1951 to 2000, the model predicts that the enrollment rate slightly decreases from 41.61% to 40.99%, 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. This confirms the results obtained from comparing steady states in Section 5.2.

Next, we shut down the demographic change by setting $n_t = 0$. What remains is only the investment-specific technological change. The results are shown in Figure 12 with the legend “ISTC only.” Overall, the “ISTC only” case shows that the skill premium keeps increasing over the period 1963-2000 at an average growth rate of 0.55%, which captures about 81% of the average growth rate in the data over the same period. Compared to the average growth rate of 0.11% in the “Demographic only” case, it is obvious ISTC is much more important in driving the skill premium over the period 1963-2000. Focusing on the three episodes, from 1963 to 1970, the average growth rate of the skill premium in the model is 0.66%, which is smaller than the average growth rate of 0.86% in the “Demographic only” case. From 1971 to 1980, the data show that the skill premium first decreases and then increases with an average growth rate close to zero, while the model goes in the other direction to predict a growth rate of 0.31% on average. However, after 1981, the “ISTC only” model generates a 0.63% average growth rate in the skill premium over the period 1981-2000, which can explain about 63% of the increase in the skill premium in the data. As shown in Table 5, compared to the “Demographic only” case, the increase in the skill premium is entirely driven by the ISTC after 1980.

Figure 13 also shows the impact of ISTC alone on the college enrollment rate. In the model, the enrollment rate increases from 41.61% in 1951 to 49.13% in 2000, almost the same as in the benchmark model. The increase in the enrollment rate over the period, therefore, is mainly driven by the ISTC.

By putting all three cases: the “Benchmark” model, the “Demographic only” case, and the “ISTC only”
<table>
<thead>
<tr>
<th>Period</th>
<th>Data (%)</th>
<th>Benchmark (%)</th>
<th>Demographic (%)</th>
<th>ISTC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1963-2000</td>
<td>0.68</td>
<td>0.64</td>
<td>0.11</td>
<td>0.55</td>
</tr>
<tr>
<td>1963-1970</td>
<td>0.63</td>
<td>1.51</td>
<td>0.86</td>
<td>0.66</td>
</tr>
<tr>
<td>1971-1980</td>
<td>0.07</td>
<td>0.06</td>
<td>-0.17</td>
<td>0.31</td>
</tr>
<tr>
<td>1981-2000</td>
<td>1.00</td>
<td>0.63</td>
<td>-0.02</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Table 5: Average Annual Growth Rate of the Skill Premium: Model vs. Data

case together in Figure 12, we can see that the “Demographic only” outcome is very close to the benchmark outcome until 1966; 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 and drives the increase in the skill premium. This conclusion is also confirmed in Table 5, which compares the average annual growth rate of the skill premium in the data and in the three model cases for different periods.

Since the effective capital-skilled labor ratio ($BK/S$) and the relative supply of skilled labor ($S/U$) are the two major determinants of the skill premium in the model (see equations 12 and 13), we also present simulations of both ratios in Figures 14 and 15, respectively. Figure 14 shows that the $BK/S$ ratio in the “Benchmark” case closely tracks that ratio in the “Demographic only” case 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 changes during that time. From 1959 to 1960, $n_t$ drops from 14.2% to 5.7%. This decrease in the growth rate of HSGs’ cohort size leads 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 1966, the benchmark $BK/S$ has closely followed the one in the “ISTC only” case. This suggests that ISTC has become the major driving force affecting this ratio. And it is the drastic rise in the capital-skilled labor ratio since the late 1970s, through capital-skill complementarity, that has been driving the rising skill premium. For purposes of comparison, we also report the data counterpart of this ratio from KORV (2000). The aggregate capital is the sum of the quality-adjusted real stock of equipment and the real stock of structures from KORV (2000). We normalize the ratio to be equal to the ratio in the benchmark model for 1963 for an easier comparison. Notice that KORV data end in 1992. Overall, one can see that the benchmark model fits the data fairly well except that the model underestimates the data since 1981.

Figure 15 shows the relative supply of skilled labor in the benchmark and the two decomposition cases. The “Demographic only” case predicts a decreasing $S/U$ from 1951 to 1976 and an 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

---

23 Jeong, Kim and Manovskii (2008) argue that the large entry of inexperienced baby boomers in the labor market during the 1970s reduced the experience per worker and led to large increase in the return to experience; while skill-biased technological change played a very insignificant role in determining the return of experience.
between these two decomposition cases, as it decreases until 1966, then begins to increase and converges to the ratio in the “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. For purposes of comparison, we also report the data of relative supply of skilled labor in the figure. The data show that the relative supply of skilled labor has increased from 28.69% in 1951 to 75.95% in 2000, around 2.6 times. The benchmark model, however, predicts an increase from 67.99% to 81.93%.24 In other words, the benchmark model only predicts about 30% of the increase in the relative supply of skilled labor due to its difficulty in replicating the college enrollment rate.

7 The Experiment with Inputted Labor Supply

As mentioned in the introduction, the benchmark model extends KORV (2000) into a general equilibrium setup by adding two more layers on top of their framework. The first layer embeds ISTC in a general equilibrium framework and hence endogenizes the capital accumulation in KORV’s production technology. The second one endogenizes the college entry decision and hence endogenizes the relative supply of skilled labor. From Figures 14 and 15, it is clear that the model has done a better job in the first layer than the second one. Given the existence of these two layers in the model and the difficulty in replicating the college enrollment rate, it is important to understand the impact of each layer on driving the results separately. Therefore, in this section, we conduct a counterfactual experiment to further decompose these two layers. In this experiment, instead of endogenizing the school choice as in the benchmark model, we input exogenously the relative supply of skilled labor from the data as in Figure 1 into the model. In other words, we shut down the second layer in the model and focus only on the first one. We then recalibrate the model and compute the transition path. The simulated skill premium is reported in Figure 16 with the legend “Benchmark.”

Replicating the relative supply of skilled labor by construction, the model predicts that the skill premium actually decreases from 1.4569 in 1951 to 1.1363 in 2000. It is not surprising because the original benchmark model with endogenous college attendance choice significantly underestimates the relative supply of skilled labor and hence the “relative quantity effect,” which has a negative impact on the skill premium as shown in equation (13). We also redo the decomposition exercise as in Section 6.2. Since we have fixed labor supply here, the “ISTC only” case is able to isolate the impact of changing age structure (via variable \( n_t \)) on capital accumulation as described qualitatively in Section 5.2, because the only difference between the “Benchmark” and the “ISTC only” case is that we have a variable growth rate of cohort size \( n_t \) in the former case. As shown in Figure 16, for most years before 1988, the “ISTC only” case generates a slightly higher skill premium than the benchmark case. This confirms the mechanism mentioned before: when baby boomers enter into the labor market, the age structure shifts toward younger cohorts. The young individuals, however, accumulate less capital in the earlier stage of their life cycle. An increase in \( n_t \) hence decreases aggregate capital in the economy because of this shift in age structure. It in turn decreases the skill premium via the “capital-skill

24The reason why the benchmark model predicts the relative supply of skilled labor of 67.99%, which is much higher than that number in the data, is because in the initial steady state we have to match the college enrollment rate of 41.54% in the data. And since it is a steady state, we have the same enrollment rate for each cohort. While in the data, for generations before 1951, they had much lower college enrollment rates than that in 1951, which translate into a lower relative supply.
complementarity” effect. Our exercise here shows the mechanism, although is in the right direction, is not quantitatively important. What is quantitatively important is that when we shut down the ISTC in the model, as shown in the figure, the “Demographic only” case generates a much worse skill premium compared to the data. The difference between the “Benchmark” and “Demographic only” case isolates the quantitative impact of ISTC alone on the skill premium when the labor supply is fixed. The finding in the endogenous labor supply scenario is confirmed here: ISTC has no significant impact on the skill premium until 1966. Then we see that its impact speeds up over time and dominates after 1980.

Since the relative supply is fixed in the model, the only endogenous term that determines the change in skill premium in equation (13) is the effective capital-skilled labor ratio $BK/S$. Figure 17 compares this ratio in three cases with the data from KORV (2000). For purposes of comparison, we normalize the ratio from the data in 1963 to be equal to the ratio in the current benchmark model for the same year. Since the relative supply of skilled labor is higher here compared to that in the original benchmark model as in Section 6.2, the divergence between the data and the benchmark model is wider than that in Figure 14. But the model is still able to capture a significant part of the increase in $BK/S$ data over time. If we shut down the ISTC, which is the engine of the capital accumulation, with the exogenous increasing relative supply of skilled labor, $BK/S$ decreases significantly as indicated in the “Demographic only” case.

Comparing Figures 12 and 16, one can see that the relatively better performance of the original benchmark model in terms of matching the skill premium in the data comes from its underestimation of the relative quantity effect. When the relative supply of skilled labor is inputted exogenously from the data, the capital-skill complementarity effect is not strong enough to drive up the skill premium as in the data. A possible reason could lie in the two key elasticity parameters $\theta$ and $\rho$. KORV (2000) estimate these two parameters in a partial equilibrium framework when four inputs in the production function, $S$, $U$, equipment $K_e$, and structures $K_s$, are all exogenously inputted from the data. Compared to KORV (2000), the current paper adopts a general equilibrium framework. Therefore all inputs are endogenous. We also have total capital stock $K$ instead of differentiated $K_e$ and $K_s$ in the current model, which makes the elasticity parameters harder to map to the ones in KORV (2000). The values of $\theta$ and $\rho$ affect both physical and human capital accumulation. As shown in equation (13), the values of parameter $\theta$ and $\rho$ obviously also affect the magnitude of change in the skill premium. When $\theta$ is higher, keeping other things equal, it magnifies the capital-skill complementarity effect and shrinks the relative quantity effect simultaneously. How to estimate these two key parameters in a general equilibrium framework is beyond the scope of the current paper. However, in an unreported exercise, if we keep $\rho = -0.495$, but pick $\theta = 0.79$ which is the value used in He and Liu (2008), in the current benchmark model with exogenously inputted labor supply, the model predicts that the skill premium increases from 1.4561 in 1951 to 2.1782 in 2000. It actually overshoots the data. With $\theta = 0.79$ and $\rho = -0.495$, in equation (13), the weight before the term for the “capital-skill complementarity” effect $\theta - \rho$ is much higher than its value under the benchmark parameter values, and the weight before the term for the “relative quantity” effect $\theta - 1$ is lower (recall that $\theta$ has to be less than 1). Therefore, it is not surprising that it generates a much higher and rising skill premium compared to the benchmark case in Figure 16.25

25 Another possible reason why the benchmark model with exogenously inputted labor supply significantly underpredicts the evolution of the skill premium is because the parameter values of $\theta$ and $\rho$ that KORV (2000) estimate is based on the quality-adjusted capital equipment, not the total capital stock as used in the current paper. ISTC on equipment only is much faster than ISTC on total capital stock. For example, data from Cummins and Violante (2002) show that ISTC on equipment
8 Sensitivity Analysis

In this section, we show that the paper’s results are robust to the alternative values of key parameters, including neutral technological change, increase in life expectancy, and the timing of the final steady state. For each experiment, we recalibrate four parameters: $\beta$, $\lambda$, $\mu$, and $sd$, while keeping other parameters unchanged as in Table 1. Therefore, the difference in the skill premium generated by each experiment and the benchmark model quantifies the impact of the change embodied in the experiment.

8.1 Elasticity of Substitution

Next, since the two elasticity parameters in the production function, $\theta$ and $\rho$, are the key parameters in the model, we verify the sensitivity of the analysis to different values of these two parameters. Specifically, we 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. We then recalibrate the model according to these new values. All of the comparative static and transition path results are similar to the benchmark results shown here. For example, the skill premium increases from 1.4543 to 1.7417 for the 1951-2000 period.

8.2 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 (13), which determines the dynamics of the skill premium, we should not expect that including TFP growth would change the results significantly. As a robustness check, we carry on an experiment that allows TFP to grow at a rate of 0.15% along the transition path. Including the neutral TFP change slightly increases the skill premium. From 1951 to 2000, the skill premium increases from 1.4534 to 1.8449, while in the benchmark model, it increases from 1.4543 to 1.7774.

8.3 Increasing Life Expectancy

Demographic changes were not only happening in terms of fluctuations in the fertility rate. In 1950 life expectancy at birth in the U.S. was only 68.2 years. It was 77 years in 2000. To take into account the effects of this demographic dimension, we do the following experiment. In an extended model with retirement and Social Security (people retire at age 66 and live up to age 100), we first use the survival probability taken from the U.S. Life Table for the 1949-1951 period to calibrate the model. It predicts that the skill premium in the initial steady state is 1.4562. We then keep all of the other parameters unchanged except that now the survival probability is replaced by the data taken from the U.S. Life Table for the 1999-2001 period. only has increased from 1 in 1951 to 8.10 in 2000, more than double than ISTC on total capital stock. If we input the data sequence of ISTC on equipment only into the model, even with the same parameter values of $\theta$ and $\rho$ as used in the current paper, one would expect the “capital-skill complementarity” effect is much stronger than in Figure 17, which will bring the model simulation much closer to the data. Of course, how to distinguish capital equipment and capital structure in the OLG setting is a challenging question.

24
Although the survival probability has increased substantially over these 50 years, which leads to the increase in life expectancy, the model predicts that the skill premium changes only slightly, to 1.4218. Given that the survival probability for the period 1999-2001 is the upper bound of the changing survival probabilities for the transition period, this experiment implies that increasing life expectancy would not change our results significantly.26

8.4 Timing of the Final Steady State

Finally, we test the robustness of the model to the timing of the final steady state. In the current model, we 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 significantly affect the results. For example, if the model reaches the final steady state right after 2000, e.g., 2005, we still obtain results almost identical to those in the benchmark model. The skill premium increases from 1.4543 to 1.7632 for the 1951-2000 period.

Figure 18 summarizes the simulated transition path for the skill premium in the different experiments mentioned above.

9 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, less attention has been paid to investigating the relative importance of each driving force to the evolution of the skill premium. In this paper, we establish and compute a general equilibrium overlapping generations model with endogenous schooling choice to answer an important quantitative question: what percentage of the change in the skill premium for the postwar 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 for 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.

We calibrate the model to match the 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

26Restuccia and Vandenbroucke (2010) find that the effect of changing life expectancy on educational attainment is not quantitatively important.
from 1951 to 2000, we 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 contributes to the decline in the skill premium in the 1970s. However, ISTC has taken over and has been driving the dramatic increase in the skill premium since the early 1980s, explaining about two-thirds of the increase in the skill premium since 1981.

Due to the simple structure in modeling college entry and attrition decision, the current paper has done a less successful job in replicating the entire increase in the college enrollment rate as observed in the data. To understand what missing feature is crucial in capturing the enrollment rate data poses an interesting question for future research.
References


Appendix A. Data

The skill premium used in Figure 1 is the ratio of the real mean annualized wage of CGs and HSGs from the CPS for 1962-2003 and the 1950 and 1960 Census. Following the literature such as Katz and Autor (1999) and Acemoglu (2003), the data counterpart of the relative供应 of skilled labor \((S/U)\) is the ratio of annual hours worked of CG equivalents and HSG equivalents, where college equivalents=CG+0.5×workers with some college, and HSG equivalents=HSG+0.5×workers with some college. The data are taken from the CPS for 1962-2003 and the 1950 and 1960 Census. See Section 4.1 for further details on sample selection.

Data on HSG cohort size 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 population data for 18-21 years old in Figure 3 come from different sources. Data for 1970-2000 are from NCES: *DES* 2002, Table 15. Data for 1960-1969 are from NCES: *DES* 1995. Data for 1955-1959 are from the *Standard Education Almanac* 1968, Table 1.

College enrollment rates of HSG for 1960-2001 in Figure 4 are from NCES: *DES* 2002, Table 183. Data for 1948-1959 are calculated by the author. To construct them, first we 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 recent HSG, we use the overlapped years 1960-1965 to calculate the average difference between our calculation and the true data, then adjust our calculation for the 1948-1959 period according to this difference.

Data on average TFRB charges in Figure 5 are constructed as follows. First, we obtain data for estimated Average Charges to Full-Time Resident Degree-Credit Undergraduate Students between 1956-57 and 1966-67 from the *Standard Education Almanac* 1969, Table 120; 1967-68 to 1973-74 from the *Standard Education Almanac* 1981-82, p. 231-232; 1974-75 to 1983-84 data from the *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. Data for 1948-1955 are from the *Standard Education Almanac* 1968, Table 102: “Estimated Costs of Attending College, Per Student: 1931-1981.” To make it consistent with the data after 1955, we use the overlapped 1956 data to adjust. Second, we focus only on public or private four-year institutions. We obtain the TFRB charges for those institutions. Third, we calculate the enrollment share of public and private four-year institutions. For the 1948-1964 data, we obtain the total fall enrollment in degree-granting institutions by control of institution (private vs. public) from NCES: *DES* 2002 Table 172, noticing that it applies to all higher education institutions. Then, from Table 173, we have total fall enrollment in degree-granting institutions by control and type of institution from 1965 to 2000. Fourth, we 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 data for the cohort-specific skill premium is in the text. (See Section 4.1.)
11 Appendix B. Algorithm to Compute the Stationary Equilibrium

Given the parameter values as shown in the text, we 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 $j$, $\sum_a \lambda^c(a,j) = e_0$ and $\sum_a \lambda^h(a,j) = 1 - e_0$, so by equations (19) and (20) we have

   $S_0 = e_0 \cdot \sum_j \eta_j e_j^c$,

   $U_0 = (1 - e_0) \cdot \sum_j \eta_j e_j^h$.

Given all the inputs, from the firm’s FOCs (9)-(11), we can compute the interest rate $r$ and wage rates $w^c$ and $w^h$.

3. Discretize the asset level $-b \leq a \leq a_{\text{max}}$ (make sure that the borrowing limit $b$ and the maximum asset $a_{\text{max}}$ will never be reached). Given prices $\{w^c, w^h, r\}$, feed in the normalized tuition data. By using backward induction (remember that $a_J = 0$), we can solve the conditional value function $V^s(a,1)$ for $s = h, c$, and therefore obtain the cut-off disutility cost $\chi(i^*) = (V^c(a = 0,1) - V^h(a = 0,1))/sd$, where $sd$ is the scale factor of the disutility cost, which is calibrated to replicate the enrollment rate data.

4. From the estimated CDF function of the disutility cost in Figure 7, corresponding to $\chi(i^*)$, we obtain the new enrollment rate $e_1$. Check the convergence criterion ($|e_0 - e_1|/e_0 \leq \text{tol}_e$). 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 (17), compute the age-dependent distributions by forward recursion. Then use these distributions $\lambda^s(a,j)$ and the age shares $\eta_j$ to compute the per capita (next period) capital stock $K_1$ as follows

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

   where $n$ is the growth rate of cohort size. Check the convergence criterion ($|K - K_1|/K \leq \text{tol}_K$) 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

\[ C = \sum_{j=1}^{J} \sum_{a} \sum_{s} \eta_{j} \lambda^{s}(a,j)C^{s}(a,j), \]
\[ P = \sum_{j=1}^{4} \sum_{a} \eta_{j} \lambda^{c}(a,j)p_{j}, \]
\[ X = (n + \delta)K. \]

7. Finally, check if the market clearing condition given by equation (21) holds. If it does, stop.

12 Appendix C. Algorithm to Compute the Transition Path

In this paper we follow Chen, İmrohoroğlu and İmrohoroğlu (2006) (also see Conesa and Kreuger 1999) in computing a transition path from the initial pre-1951 steady state toward a final steady state. In this way, we view 1952-2000 as a part of the transition path. Defining notation, we use \( t = 1 \) for 1951, \( t = T \) for the final steady state, and \( t = 2, \ldots, T - 1 \) for the transitional period. We 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 \( f_{n}^{T} \), transformed capital-specific technological change \( f_{B}^{T} \), and the tuition payments \( f_{p}^{T} \).

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

Figure 2: High School Graduates Cohort Size
Figure 3: College Age Population

Standard Education Almanac (SEA), 1968
Figure 4: College Enrollment Rate of High School Graduates

Source: DES 2002 and calculated by the author
Figure 5: Average Real TFRB charges

Source: DES, SEA and calculated by the author
Figure 6: Life-cycle HSG Wage Profile: 1975 Cohort

Source: CPS 1962-2003
Figure 7: CDF of Disutility Cost
Figure 8: Age Efficiency Unit Profile

Figure 9: Investment-Specific Technological Change (Log Units)

Source: Cummins and Violante (2002)
Figure 10: Skill Premium: Model vs. Data
Figure 11: College Enrollment Rate: Model vs. Data
Figure 12: Skill Premium: Decomposition
Figure 13: College Enrollment Rate: Decomposition
Figure 14: Capital-Skill Labor Ratio: Decomposition
Figure 15: Relative Supply of Skilled Labor: Decomposition
Figure 16: Skill Premium in Model with Exogenously Inputted Labor Supply
Figure 17: Capital-Skill Labor Ratio in Model with Exogenously Inputted Labor Supply
Figure 18: Skill Premium: Sensitivity Analysis