

Why Have Girls Gone to College? A Quantitative Examination of the Female College Enrollment Rate in the United States: 1955-1980*

Hui He

*Department of Economics, University of Hawai'i at Manoa
2424 Maile Way, Saunders Hall 542, Honolulu, HI 96822, U.S.A.
E-mail: huihe@hawaii.edu*

This paper explores the extent to which the increase in the college enrollment rate of women in the U.S. from 1955 to 1980 can be accounted for by the change in the female college wage premium. I develop and calibrate a dynamic overlapping generations model with discrete schooling choice. I find that changes in the life-cycle earnings differential can explain the increase in the female college enrollment rate very well. Young women's changing expectations of future earnings may also play an important role in driving their college entry decision.

Key Words: Female college enrollment rate; College wage premium; Life-cycle.
JEL Classification Numbers: E24, J24, J31, I21.

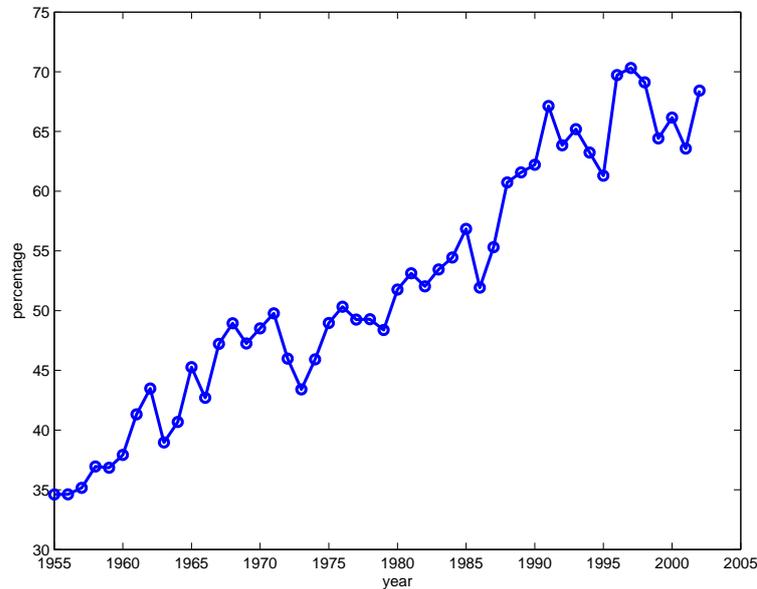
1. INTRODUCTION

Female college attainment in the United States has changed dramatically over the last 50 years. In 1955, only 34.7 percent of college students were women. By 2001, this ratio increased to 56.3 percent. (National Center for Educational Statistics, *Digest of Education Statistics* 2003, Table 174). The main driver of this large increase in female college attainment was the rising college enrollment rate of women over the past five decades. As shown in Figure 1, the female college enrollment rate of recent high school graduates (individuals age 16 to 24 who graduated from high school or completed a GED during the preceding 12 months) was only 34.6 percent

* This paper is based on the third chapter of my dissertation at the University of Minnesota. I thank Kelly Bedard, Michele Boldrin, Zvi Eckstein, and participants at the 2007 North American Summer Meeting of the Econometric Society for helpful comments and suggestions. All errors are my own.

in 1955; however, it has been increasing since then. In 1980, this rate increased to 51.8 percent. And by 2002, 68.4 percent of female high school graduates went to college.¹

FIG. 1. Female college enrollment rate of recent high school graduates



A well-known phenomenon in the U.S. labor market is the rising college wage premium over the past 50 years (except when it decreased in the 1970s, as shown in Katz and Murphy 1992 and Katz and Autor 1999). Data from the Current Population Survey (CPS) show that this trend is also true for females. The ratio of the annual mean wage of female college graduates to that of high school graduates increased from 1.44 in 1963 to 1.51 in 1969. It then decreased from 1.49 in 1970 to 1.38 in 1980, however, has since increased dramatically to 1.91 in 2001.

This paper investigates the connection between these two phenomena by asking a quantitative question: *to what extent can the changes in the female college enrollment rate from 1955 to 1980 be explained by the changes in the female college wage premium?*² In order to answer this question, this paper

¹College enrollment rates for the 1960 to 2002 period are taken from the National Center for Educational Statistics, *Digest of Education Statistics* 2003, Table 186. The data from 1955 to 1959 were calculated by the author.

²The period from 1955 to 1980 is chosen because the cohort-based estimates of life-cycle wage profiles used in the paper are subject to a serious missing data problem before the 1955 cohort and after the 1980 cohort due to the CPS sample period. See Section 3.1 for further discussion.

develops and calibrates a discrete time overlapping generations model with endogenous college-entry decision. Different cohorts of women enter the economy at age 18 and face the decision of whether or not to go to college. This decision is based on the comparison of their expected future wage differentials, their forgone wages during their college years, their tuition payments and their idiosyncratic disutility costs, which capture the non-pecuniary costs of a college education. The decision, in turn, determines their consumption, savings and wages over the life-cycle until age 65.

The economic mechanism linking the rising female college enrollment rate to the rising female college wage premium is intuitive: the increasing wage premium raises the expected wage differentials, which lead to higher benefits of a college education. Since the female college wage premium has generally increased since 1955, one would expect to find that more and more women go to college.

Inputting the cohort-specific life-cycle wage profiles into the model, I find that the model captures the rising female college enrollment rate during the period 1955-1980 quite well. The rising college wage premium is the major force driving the substantial increase in women's college enrollment. The results also suggest that the change in expectations of future earnings (towards more forward-looking) among young women may have also played an important role in driving the enrollment rate in the late 1960s and early 1970s.

This paper contributes to a large empirical literature on female college enrollment.³ Averett and Burton (1996) study how one cohort (those ages 14 to 21 in 1979) responded to the jump in the college wage premium after 1980. They find that the effect of the college wage premium for women is small and statistically insignificant. Jacob (2002) finds that higher returns to college education and the greater non-cognitive skills among women account for nearly 90 percent of the gender gap in the college attendance rate in 1988. While these papers do not examine the time trend of the female college enrollment rate, Anderson (2002) tries to answer the same question by analyzing different cohorts over time. She finds that an important driver of the increase in female enrollment over time is the behavior of older women who enrolled less frequently than men when young, but who later made up for this lack of higher education. Charles and Luoh (2003) argue that not only the expected earnings differential but also the anticipated *dispersion* of future earnings determine individuals's education-

³Most of this literature focuses on the gender difference in college enrollment. Since it is well known that non-economic factors, such as the Korean and Vietnam wars (through the GI Bill and the military draft), had a significant impact on male college attendance during the 1955-1980 period (Bound and Turner 2002 and Card and Lemieux 2001) and they cannot be easily captured in the model, this paper ignores males and focuses only on female college-entry decisions.

al investment decisions. Using the CPS data, they show that the dispersion of future earnings for college-educated women has decreased over the past three decades. Goldin, Katz and Kuziemko (2006) document the reversal of the college gender gap and argue that the relatively greater economic benefits of college education and the relatively lower non-pecuniary costs of college attendance for women play key roles in explaining this reversal. Becker, Hubbard and Murphy (2010) also argue that women have lower nonmonetary costs (stronger “noncognitive skills”) associated with attending college than men, and therefore, are more likely to attend and complete college. Different from the empirical approach taken in most of the literature, the current paper develops a dynamic overlapping generations model in which individuals optimally choose their college-entry decision to quantify the effect of rising wage premium on the female college enrollment rate over time.

This work is more closely related to a growing literature that employs structural models to *quantitatively* decompose the driving force behind the female college-entry decision and educational attainment.⁴ Ge (2010) structurally estimates a dynamic choice model of college attendance, labor supply and marriage and finds that marriage benefits from college attendance are important in determining young women’s schooling choice. Rios-Rull and Sanchez-Marcos (2002) develop an overlapping generations model with endogenous schooling, marriage and fertility choice to study the high ratio of male to female college graduates (sex college attainment ratio, SCAR) in the 1970s. These papers, however, do not analyze the changes in college education attainment over time. Sanchez-Marcos (2008) builds on Rios-Rull and Sanchez-Marcos (2002) to quantify the reduction in the SCAR. She finds that observed changes in earnings and fertility account for a substantial amount of the reduction in the college attainment gender gap. Changes in marital status and marital sorting, conversely, reduce the college attainment of women. However, she only compares the SCAR in 1976 and 1990 and does not examine the time path of the changes in college education attainment. Ge and Yang (2010) develop a discrete choice model of college-entry decisions with a rich structure in marriage and fertility to quantify the effects of changes in relative earnings, parental education and the marriage market on changes in college attainment (measured by the fraction of individuals that have completed college education among each specific group) by gender from 1980 to 1996. They find that the increasing gap in earnings between college and high school graduates and increasing parental education have important effects on the increase in college attainment for both genders but cannot explain the reversal of the gender gap.

⁴There is a growing literature that uses a dynamic model to study the rising female labor supply. See Attanasio, Low and Sanchez-Marcos (2008), Olivetti (2006) and Fernandez, Fogli and Olivetti (2004).

Changes in the marriage market via the increasing probability of divorce are crucial in explaining the relative increase in female college attainment. The current paper is close to Ge and Yang (2010) in spirit. However, it puts less emphasis on the effects of marriage and fertility and has a relatively richer structure in consumption and savings.⁵ The results show that the change in life-cycle earnings is a key factor accounting for the changes in the female college *enrollment* rate from 1955 to 1980, a period not covered in Ge and Yang (2010). Restuccia and Vandenbroucke (2010), on the other hand, use a model similar to that herein to investigate the increase in the educational attainment (measured by average years of schooling) of white males from 1940 to 2000. They conclude that changes in return to schooling can account for the entire increase in educational attainment in their data.⁶ This paper confirms their findings for female college enrollment behavior.

The remainder of the paper is organized as follows. Section 2 presents a simple model of the college attendance decision. Section 3 describes the data and the parameterization of the model. Section 4 presents the results of the benchmark model. Section 5 conducts several counterfactual experiments. Finally, Section 6 concludes.

2. MODEL

In this section, I present the economic model that will be used later for calibration. The framework is similar to that in He (2009).⁷ It is a discrete time overlapping generations (OLG) model in which individuals make the schooling choice in the first period. There is only one good in the economy, which can be used for either consumption or investment. There is no uncertainty in the model and individuals have perfect foresight.⁸

⁵Ge and Yang (2010) do not explicitly model individuals' consumption and savings decisions.

⁶Unlike the current paper, Restuccia and Vandenbroucke (2010) do not directly input the estimated life-cycle wage profiles from the data in the model. Rather, they calibrate the skill-specific productivity parameters that determine the trend of wage premium.

⁷He (2009) develops an overlapping generations general equilibrium model with endogenous discrete schooling choice. The production technology features capital-skill complementarity as in Krusell et al. (2000). Using the model, the 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. The current paper has a similar setting for the consumer's problem and schooling choice. However, it is silent on the production side, taking the wage profiles as given from the data to investigate their effects on schooling choice.

⁸This paper focuses on the effect of the *trend* of the college wage premium on the female college enrollment rate. An important stylized fact in the U.S. labor market over the past half century is not only the trend of the college wage premium, but also the rising within-group wage inequality. The paper could benefit from adding uncertainty or heterogeneity into individuals' wage profiles. However, that is beyond the scope of the

2.1. Demographics

The economy is populated by overlapping generations of finite-lived women with total measure one. Women enter the economy (or are “born”) with zero initial assets at age 18, the common age of high school graduates. I call this the birth cohort and model age as $j = 1$. They live and work until age J . The model period is one year. To distinguish age from calendar time, I use two time subscripts for each economic variable. For example, $c_{j,t+j-1}$ denotes consumption for an age- j woman (who is born at time t) at time $t + j - 1$.

2.2. Preferences

Each individual female born at time t wants to maximize her discounted lifetime utility

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

The period utility function is assumed to take the CRRA form

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

σ is the coefficient of relative risk aversion and $\frac{1}{\sigma}$ is the intertemporal elasticity of substitution. Since leisure does not enter into the utility function, each woman supplies all of her labor endowment, which is normalized to one.

2.3. Budget Constraints

A woman decides whether or not to go to college at the beginning of the first period. The choice is denoted by $s \in \{c, h\}$. If an individual chooses $s = h$, she ends up with a high school diploma, goes on the job market to work as an unskilled laborer and earns the high school graduate wage sequence $\{w_j^h\}_{j=1}^J$. She can alternatively choose $s = c$, spend the first four years in college as a full-time student and pay the tuition. After that, she goes on the labor market to find a job as a skilled worker and earns the college graduate wage sequence $\{w_j^c\}_{j=1}^J$. For simplicity, I assume there is no college dropout and no unemployment.⁹

current paper and I refer readers to Charles and Luoh (2003) and Heathcote, Storesletten and Violante (2010).

⁹College education could be a risky investment. Allowing the probability of dropout and unemployment reduces the return to schooling and thus might lower the incentive to go to college. See Garriga and Keightley (2007) for an overlapping generations model with endogenous enrollment, time-to-degree and dropout behavior.

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

$$\begin{aligned}
c_{j,t+j-1} + tuition_{t+j-1} + a_{j,t+j-1} &\leq (1 + r_{t+j-1})a_{j-1,t+j-2} & (2) \\
&\forall j = 1, 2, 3, 4 \\
c_{j,t+j-1} + a_{j,t+j-1} &\leq (1 + r_{t+j-1})a_{j-1,t+j-2} + w_{j,t+j-1}^c \\
&\forall j = 5, \dots, J \\
c_{j,t+j-1} &\geq 0, a_{0,t-1} = 0, a_{J,t+J-1} \geq 0.
\end{aligned}$$

In the first four periods, she pays tuition $tuition_{t+j-1}$, consumes $c_{j,t+j-1}$ and saves $a_{j,t+j-1}$. After graduation, she earns wage $w_{j,t+j-1}^c$ at age j and consumes and saves subject to what she earns and accumulates. Notice that there is no borrowing constraint in this economy. Since they do not have any initial assets, college students need to borrow money for consumption and to pay tuition during the first four periods, and they pay back the loans later.¹⁰

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

$$\begin{aligned}
c_{j,t+j-1} + a_{j,t+j-1} &\leq (1 + r_{t+j-1})a_{j-1,t+j-2} + w_{j,t+j-1}^h & (3) \\
&\forall j = 1, \dots, J \\
c_{j,t+j-1} &\geq 0, a_{0,t-1} = 0, a_{J,t+J-1} \geq 0.
\end{aligned}$$

2.4. Schooling Choice

I assume that different individuals within each birth cohort are endowed with different levels of ability and that ability affects only individuals' disutility cost of schooling.¹¹ In particular, the disutility cost of schooling is a strictly decreasing function of ability and, therefore, higher ability implies lower disutility cost.

Individuals in each cohort are indexed by their ability level $i \in [0, 1]$. The CDF of the ability distribution is denoted by F , $F(i_0) = \Pr(i \leq i_0)$. $\chi(i)$ represents the time-invariant, ability-related disutility cost for individual i .

¹⁰Since the model does not include heterogeneity in individuals' wealth distribution and wage profiles, given the zero initial asset, with the borrowing constraint, the enrollment rate in each birth cohort would be zero. Cameron and Taber (2004) also find no empirical evidence that access to borrowing is an important component of schooling decisions.

¹¹This is a common assumption in the literature. Ge and Yang (2010) and Restuccia and Vandenbroucke (2010) make a similar assumption. Navarro (2007) empirically finds that learning ability is the main determinant of this "psychic" cost and that it plays a key role in determining schooling decisions. An alternative (and more complicated) modelling strategy would be allowing ability levels to affect earnings. See Hendricks and Schoellman (2009) for an analysis of the relationship between the evolution of wages and abilities.

Note that $\chi(i) \geq 0$ and $\chi'(i) < 0$. An individual i born at time t thus has the discounted lifetime utility

$$\sum_{j=1}^J \beta^{j-1} u(c_{j,t+j-1}) - \mathcal{I}_i \chi(i), \quad (4)$$

where

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

She maximizes her lifetime utility subject to the budget constraints (2) or (3) conditional on her educational choice. If an individual chooses to go to college, she has to bear the idiosyncratic disutility cost. Notice that the disutility cost, $\chi(i)$, does not enter into the budget constraint; therefore, everyone with the same educational achievement from the same birth cohort has the same lifetime utility derived from physical consumption. V^c denotes the discounted lifetime utility for college graduates born at time t ; V^h denotes the discounted lifetime utility derived from physical consumption for high school graduates born at time t . $V_t^c - V_t^h$ represents the utility gain from attending college. Individual i will choose to go to college if $\chi(i) < [V_t^c - V_t^h]$, not to go if $\chi(i) > [V_t^c - V_t^h]$, and is indifferent if $\chi(i) = [V_t^c - V_t^h]$. We thus have the following criteria of schooling choice for a woman with ability index i born at time t :

$$\begin{aligned} s_{i,t} &= c & \text{if } V_t^c - \chi(i) > V_t^h, \\ s_{i,t} &= h & \text{if } V_t^c - \chi(i) < V_t^h, \\ s_{i,t} &= \text{indifferent} & \text{if } V_t^c - \chi(i) = V_t^h. \end{aligned} \quad (5)$$

Since the borrowing constraint does not exist, the model implies

$$V_t^c - V_t^h \geq 0 \text{ iff } NPV_t \geq 0,$$

where

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

Here, NPV stands for the net present value of higher education. Since $w_{j,t+j-1}^c = 0, \forall j = 1, \dots, 4$, students never work while they are in college,

and NPV can be further decomposed into three components:

$$NPV_t = \sum_{j=5}^J \frac{w_{j,t+j-1}^c - w_{j,t+j-1}^h}{\prod_{i=2}^j (1 + r_{t+i-1})} - \sum_{j=1}^4 \frac{w_{j,t+j-1}^h}{\prod_{i=2}^j (1 + r_{t+i-1})} \quad (6)$$

$$- \sum_{j=1}^4 \frac{tuition_{t+j-1}}{\prod_{i=2}^j (1 + r_{t+i-1})}.$$

The first term represents the benefits of schooling: college graduates can earn more through the earnings differential. The second term represents the opportunity cost of schooling: it is the present value of four years of forgone wages for college students. The third term is the present value of tuition paid during the college years, which represents the direct cost of schooling. From this representation it is very clear how the cohort-specific lifetime college wage premium, $\left\{ \frac{w_{j,t+j-1}^c}{w_{j,t+j-1}^h} \right\}_{j=1}^J$, will affect people's schooling decision. Other things being equal, an increase in the lifetime college wage premium raises the benefits of schooling, and hence, NPV . Higher NPV induces a higher utility gain from schooling, $V_t^c - V_t^h$. Given the stationary distribution of the disutility cost, a higher utility gain from schooling makes it more likely that $\chi(i) < [V_t^c - V_t^h]$, which implies a higher enrollment rate.

The basic intuition of this model is depicted in Figure 2. The x-axis measures ability i . Women are ranked from zero to one by their ability. The disutility cost, $\chi(i)$, is a decreasing function of the ability index i . VD represents the utility gain from attending college $V^c - V^h$. The cut-off ability (or indifference level), i^* , is determined by

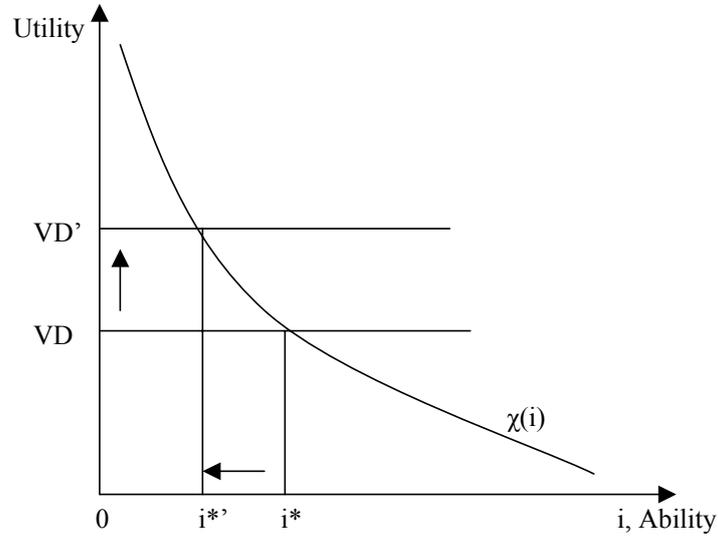
$$\chi(i^*) = [V^c - V^h].$$

Therefore, women with ability $i < i^*$ will choose not to go to college, while women with ability $i > i^*$ will choose to go. The enrollment rate is thus equal to the probability that $i > i^*$. If the college wage premium increases over the life-cycle, so does the NPV . Therefore, the utility gain, VD , increases to VD' , and this will decrease the cut-off point to $i^{*'}$. Since $\Pr(i > i^{*'}) > \Pr(i > i^*)$, more women go to college. A higher life-cycle earnings differential thus encourages college attendance.

3. DATA AND CALIBRATION

In this section, I first show how the life-cycle wage profile data are constructed. I then describe the parameterization of the model.

FIG. 2. The determination of the college enrollment rate



3.1. Cohort-Specific Wage Premium

The March Current Population Survey (CPS) from 1962 to 2003 is employed to construct the wage profiles used as inputs in the model. Sample restrictions and the definition of education levels are set to follow those in Eckstein and Nagypál (2004), except I further restrict the data to include only high school graduates (HSG hereafter, those who completed exactly 12 years of schooling) between ages 18 and 65 and college graduates (CG hereafter, those who completed 16 years of schooling) between ages 22 and 65 in the sample. Also following Eckstein and Nagypál (2004), I restrict my attention to full-time full-year (FTFY) workers.¹² 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.

In the model, women make the educational decision based on the expected earnings differential specific to their cohort. The perfect foresight assumption allows for the use of *actual* observed future earnings in the CPS

¹²This restriction is consistent with the model assumption of fixed labor supply and absence of unemployment. However, during the time period analyzed female labor supply (both intensive and extensive margin) increased significantly. Adding the labor force participation decision to the model might further strengthen the effect of the rising college wage premium on the female enrollment rate. I leave that extension for future research.

as the measure of expected future earnings. Since the CPS is a repeated cross-sectional data set, a so-called “pseudo-cohort construction method” is used to construct the cohort-specific expected wage profiles.¹³ For example, the 1962 cohort’s (18-year-old HSG in 1962) lifetime (18-65 years old) female HSG wage profile, $\{w_{j,1961+j}^h\}_{j=1}^{48}$, is constructed as follows: calculate the mean wage of 18-year-old female HSGs in 1962, that of 19-year-old female HSGs in 1963, that of 20-year-old female HSGs in 1964, that of 21-year-old female HSGs in 1965 and so on, ending with 58-year-old female HSGs in 2002, which is the end year of the CPS data set employed.

A similar approach is used to construct the 1962 cohort’s female CG wage profile, $\{w_{j,1961+j}^c\}_{j=1}^{48}$. However, this process starts from 1966 because if someone from the 1962 cohort chooses to go to college, she spends four years in college. She graduates in 1966 and starts earning CG wages from that year. Therefore, this cohort’s lifetime wage profile is constructed by calculating the mean wage of 22-year-old female CGs in 1966, that of 23-year-old female CGs in 1967 and so on.

This method is used to construct life-cycle wage profiles for HSGs and CGs for the 1955 to the 1980 cohorts. However, this procedure leads to two problems. First, due to the time range of the CPS data, I do not have a complete life-cycle wage profile for any cohort. For example, some cohorts miss the later age data points (cohorts after 1961) and some miss the early age data points (cohorts between 1955 and 1960). Therefore, an econometric method is used to predict the mean wage at that specific age to extrapolate the missing data. They are predicted by either second- or third-order polynomial specification or a conditional Mincer equation as follows:

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

Extrapolation stops after the 1980 cohort because after this cohort, a lack of data points prevents reliable prediction.¹⁴ Filling in the missing data produces complete cohort-specific life-cycle wage profiles for HSGs and CGs for all cohorts from 1955 to 1980. The second problem caused

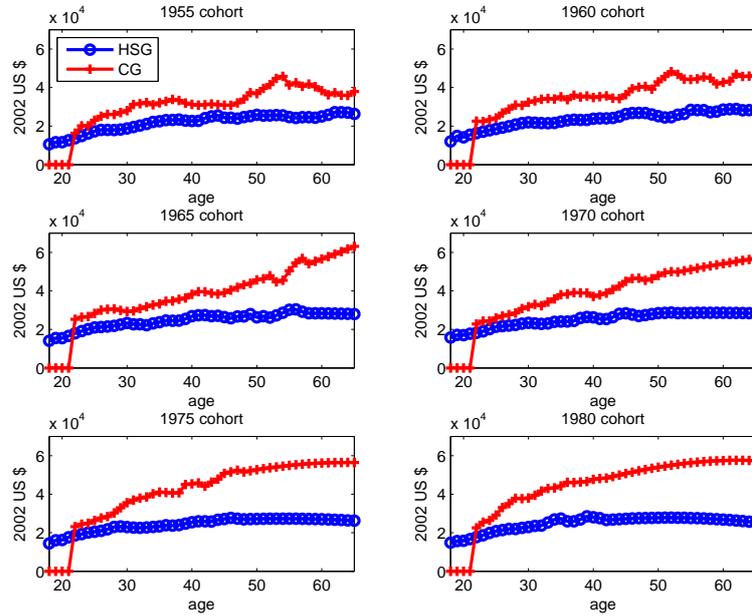
¹³It is a pseudo-cohort because the CPS is not a panel data set, it does not track people over their lifetime. Heckman, Lochner and Todd (2003) use a similar method to estimate the cohort-based return to schooling.

¹⁴The 1980 cohort has a life-cycle wage profile only up to age 40 from the CPS data. Heckman, Lochner and Todd (2003) also notice this problem and stop in 1983 for their cohort-based estimates.

by the pseudo-cohort method is that lower female labor force participation in the 1950s and 1960s creates noisy estimates of wage data, especially for older female workers. A three-year moving average is thus applied to smooth the set of complete life-cycle wage profiles.

Figure 3 shows the life-cycle wage profiles for six selected cohorts: 1955, 1960, 1965, 1970, 1975 and 1980 cohorts. For each cohort, the wage profile of CGs is significantly higher than that of HSGs. Two features of the life-cycle wage profiles should be noted: (1) Earnings rise with age, but at a decreasing rate; and (2) Earnings increase faster for more educated workers, which implies that CGs have a steeper hump-shaped (or increasing but concave) wage profile than HSGs.

FIG. 3. Life-cycle wage profiles for six cohorts

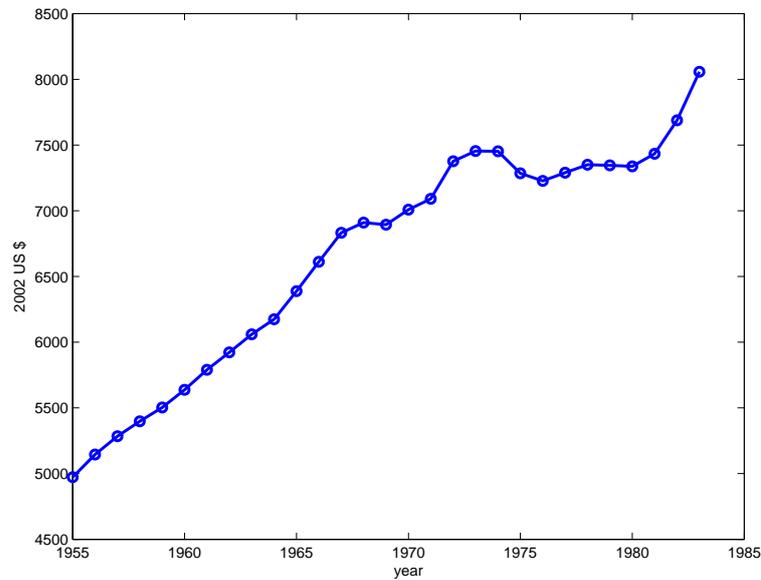


The college wage premium over the life-cycle, $\left\{ \frac{w_{j,t+j-1}^c}{w_{j,t+j-1}^h} \right\}_{j=5}^{48}$, exhibits interesting patterns for these cohorts. The average college wage premium from age 22 to age 65 for the 1955 cohort was 1.46. For the 1960 cohort, it was 1.54. It continued to increase to 1.62 for the 1965 cohort, but then stabilized at 1.62 for the 1970 cohort. This is because the compressed college wage premium in the 1970s significantly reduced the earnings differential at the prime age when the CG wage profile was in a stage of steep ascent. In contrast, the rising college wage premium starting from 1980 helped to

raise the average life-cycle college wage premium for the 1975 and 1980 cohorts to 1.81 and 1.85, respectively.

These cohort-specific life-cycle wage profiles provide the information needed in the first two terms of equation 6. To fully understand the higher education choice over time, tuition information is needed. This direct cost of college education is represented by the third term of equation 6. Figure 4 reports the real tuition, fees, room and board (TFRB) per student charged by an average four-year institution in constant 2002 dollars.¹⁵ TFRB increased over time except the 1970s. Different cohorts faced different TFRB charges based on the years during which they attended college. For example, the 1955 cohort paid the tuition from 1955 to 1958.

FIG. 4. Real tuition, fees, room and board charge per student



¹⁵Average TFRB charges data are constructed as follows. First, data about estimated average charges to full-time resident degree-credit undergraduate students in public and private four-year institutions are attained from various issues of *Standard Education Almanac*. Second, total fall enrollment in degree-granting institutions by control of institution (private vs. public) is obtained from National Center for Educational Statistics, *Digest of Education Statistics 2002*. Average TFRB charges of public and private four-year institutions are then weighted by enrollment share. Finally, the three-year moving average method is used to smooth the data.

3.2. Calibration

The value of discount factor β is taken to be 0.96 to match the interest rate r , which is set to four percent.¹⁶ The value of the CRRA coefficient σ is two, which is widely used in the life-cycle literature.

For simplicity, I assume that the ability level i is uniformly distributed among women over $[0, 1]$ and the ability-related disutility cost takes the form

$$\chi(i) = b\left(\frac{1}{i} - 1\right). \quad (7)$$

For the lowest ability individual ($i = 0$), $\chi(i) = \infty$, so she will never go to college. On the other hand, for the highest ability individual ($i = 1$), $\chi(i) = 0$. Since the present value of the life-cycle wage profile of CGs is higher than that of HSGs (see Figure 3), she will certainly choose to go to college. This functional form thus guarantees the simulated enrollment rates will be between zero and one.

The scale factor of disutility cost function b is calibrated to match the female college enrollment rate data in 1955, which is 34.6 percent. This results in $b = 4.37$.

Table 1 summarizes the parameter values used in the model.

TABLE 1.

Parameter values of the benchmark model

Parameters	Description	Value
β	discount factor	0.96
σ	CRRA coefficient	2
b	scale factor of disutility cost function	4.37
r	real interest rate	4%

4. RESULTS

In this section, the economic model in Section 2 is computed numerically to generate the female college enrollment rates. Under the calibrated parameter values as in Table 1, the life-cycle wage profile data for cohorts from 1955 to 1980 (as shown in Figure 3) and the real TFRB data (as shown in Figure 4) are inputted into the model. Since the data contain enough information about individuals' budget constraints, at any year t from 1955 to 1980, the birth cohort's dynamic programming problem can be solved backward to obtain the conditional value functions at age 1, V_t^c and V_t^h

¹⁶An unreported experiment shows that the benchmark results are very robust to different combinations of β and r .

for this cohort (born at time t). Given the disutility cost in equation 7, the college enrollment rate at time t (e_t) is determined by the threshold level i_t^* according to the following equation

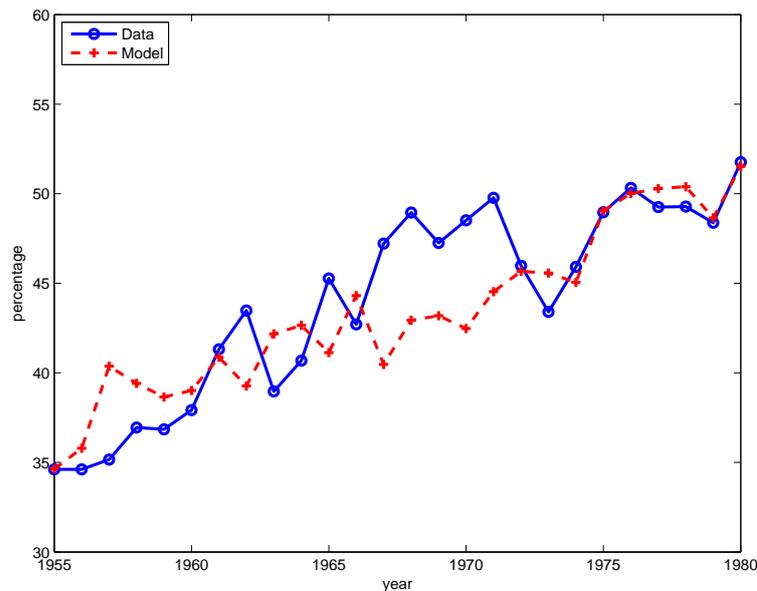
$$\chi(i_t^*) = V_t^c - V_t^h.$$

Since the ability level is uniformly distributed and $\chi'(i) < 0$, we have $e_t = \Pr(i > i_t^*) = 1 - F(i_t^*) = 1 - i_t^*$.

Figure 5 compares the female college enrollment rates from the model with those in the data from 1955 to 1980. Since the female enrollment rate in 1955 is used for calibration, one should evaluate the model performance by comparing the model-generated female enrollment rates for the period from 1956 to 1980 with the data.

Overall, the model replicates the rising trend of college enrollment rates for females very well. In the data, the enrollment rate increased from 34.6 percent in 1956 to 51.8 percent in 1980; in the model, it increased from 35.8 percent to 51.5 percent. The benchmark model captures about 91 percent of the increase of female college enrollment rate over the entire period from 1956 to 1980.

FIG. 5. Female college enrollment rate: model vs. data



From 1955 to 1960, the model predicts that the enrollment rate increased from 34.6 percent to 39.0 percent; in the data it increased from 34.6 percent to 37.9 percent. Thus, the model actually overshoots the data for this

period and implies that a factor other than the life-cycle earnings differential was deterring women from going to college in that period. From 1960 to 1966, the prediction from the model is relatively aligned with the data. However, from 1967 to 1971, the model significantly underpredicts the college enrollment rate for females. In 1967, the enrollment rate was 47.2 percent in the data while the model predicts a rate of only 40.5 percent, a 6.7% differential. Similarly, in 1971, the college enrollment rate was 49.8 percent in the data, while only 44.5 percent in the model.

One possible reason for the significant underprediction of the female college enrollment rates by the model during the period from 1967 to 1971 could be the assumption of perfect foresight used in the benchmark model. Women in the model know their actual life-cycle wage profiles. As discussed in Section 3.1, the compressed college wage premium in the 1970s significantly reduced the earnings differential for the cohorts in the late 1960s and early 1970s and thus discouraged college enrollment. Women in reality, however, might not have perfect foresight of their future earnings and hence more likely tend to make their college-entry decisions based on the current observed cross-sectional wages (across ages). Therefore, for female HSGs in the late 1960s with short-sighted expectations, the decreased college wage premium starting in 1970 has less impact than it would for women in the benchmark model. In Section 5.2, this conjecture is tested by changing the model's assumption of perfect foresight.

The model replicates the data since 1972 very well. In the data, the female college enrollment rate increased from 46.0 percent in 1972 to 51.8 percent in 1980. The model counterpart was from 45.7 percent to 51.5 percent. The higher college wage premium for females since 1980 has raised the benefits of attending college, as shown in the first term of equation 6. It was a significant factor in encouraging girls to go to college.

5. COUNTERFACTUAL EXPERIMENTS

In this section, I run a series of counterfactual experiments to test the robustness of the benchmark results to changing tuition costs, an alternative assumption of expectations and a more general distribution of disutility costs.

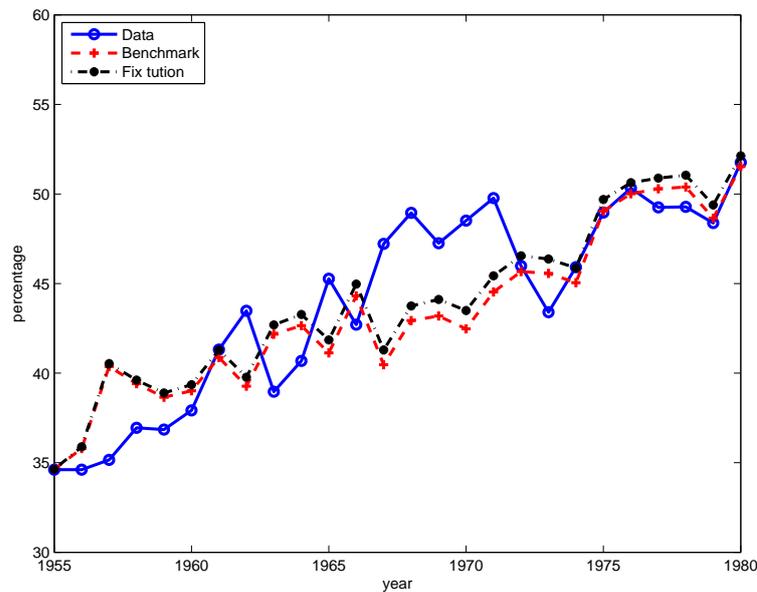
5.1. Fixed Tuition Cost

In order to quantify the effects of changing tuition costs over the target period, tuition costs are held fixed at the level of the 1955 cohort. Therefore, the 1956 to 1980 cohorts face the same tuition costs for their four-year college education as the 1955 cohort (i.e. real TFRB from 1955 to 1958). Figure 6 shows the results. Compared to the benchmark case, when the tuition cost is fixed at the level for the 1955 cohort, the female college

enrollment rate increases only slightly (on average 0.6 percent per year) over the period from 1955 to 1980. Therefore, the direct cost of schooling does not appear to be a significant factor in determining women's college entrance behavior.

The tuition cost is only a small fraction of annual earnings. For example, the average annual TFRB cost for the 1955 cohort only accounts for about 24 percent of the average annual wage of HSGs for the same cohort. Moreover, the growth of real TFRB is outpaced by the growth of the earnings differential over the life-cycle. As shown in Figure 4, real TFRB for a four-year education increased from \$4973 in 1955 to \$8338 in 1984, an average increase of \$179 per year. On the other hand, the average annual earnings differential between CGs and HSGs over the life-cycle was \$10394 for the 1955 cohort. It increased to \$12360 for the 1960 cohort, and to \$14662 for the 1965 cohort. Although stabilized at \$14966 for the 1970 cohort, it increased dramatically to \$19116 for the 1975 cohort and \$20770 for the 1980 cohort. Looking back to equation 6, this suggests that the third term is almost negligible compared to the first term. Therefore, it is not surprising that the tuition cost has virtually no significant effect on college enrollment rates. This exercise thus confirms that the rising female college enrollment rate is largely driven by the rising college wage premium and not by changing tuition costs over time.

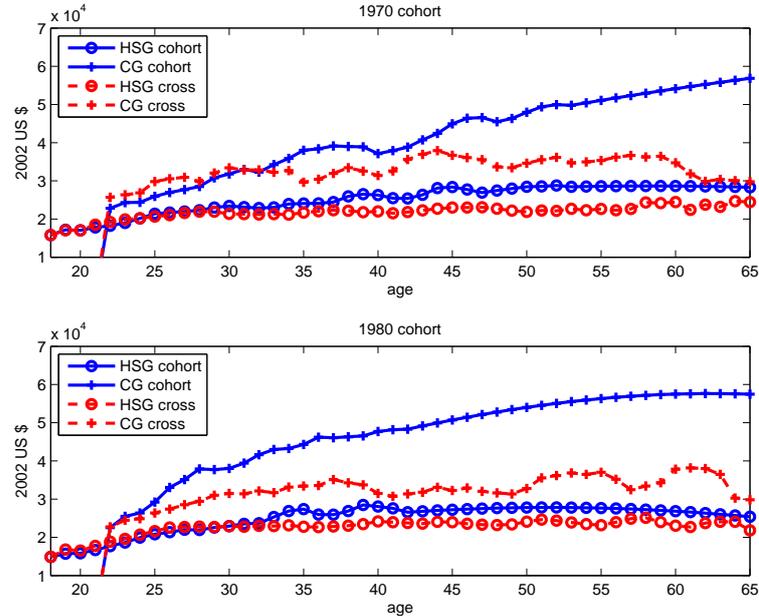
FIG. 6. Fixed tuition cost



5.2. Myopic Expectations

To test whether the benchmark results are sensitive to the assumption of perfect foresight, that assumption is changed to one of myopic expectations, meaning individuals can forecast their future earnings based only on the observed *cross-sectional* earnings at the time they are making the college entry decision. For example, the cohort-specific wage profiles for the women in the 1970 cohort are not inputted as shown in Figure 3. Instead the cross-sectional wage profiles for ages 18 to 65 in 1970 are used to proxy their life-cycle wage profiles. Figure 7 shows the difference between cohort-specific wage profiles for the benchmark model with perfect foresight expectations (blue and solid line) and the cross-sectional wage profiles as the input for the model with myopic expectations (red and dashed line) for the 1970 and 1980 cohorts, respectively. Notice that the cohort-specific wage profiles are lower than cross-sectional profiles for CGs of the 1970 cohort from age 22 to early 30s. This is exactly because the compressed wage premium in the 1970s, although reflected in the forward-looking cohort-based life-cycle wage profiles, did not affect the cross-sectional wage profiles prior to 1970. On the other hand, the dramatically rising wage premium since 1980 significantly raised the cohort-specific wage profiles for CGs of the 1980 cohort. The earnings differential was much higher under the cohort-based wage profiles from age 22 for that cohort.

FIG. 7. Wage profiles: cohort-based vs. cross-section



Due to CPS data availability, one can only obtain the complete cross-sectional wage profiles for ages from 18 to 65 since 1961. Therefore, the period for comparison between the benchmark model and the model with myopic expectations is from 1961 to 1980. First, the tuition data from 1961 to 1984 and the cohort-specific wage profiles for cohorts from 1961 to 1980 are inputted in the model. b is recalibrated to 4.29 to match the 41.32 percent female enrollment rate in 1961 (keeping other parameter values from Table 1). In other words, the benchmark model is redone, only changing the starting point from 1955 to 1961. The resulting female college enrollment rates are reported as the “Benchmark” case in Figure 8. The model is then solved again, this time changing the data input to the cross-sectional wage profiles for cohorts from 1961 to 1980. Here, $b = 2.91$ is recalibrated to match enrollment rate data in 1961, keeping the data of real TFRB and other parameter values unchanged. The results are reported as the “Myopic expectations” case in the same figure.

FIG. 8. Myopic expectations vs. perfect foresight

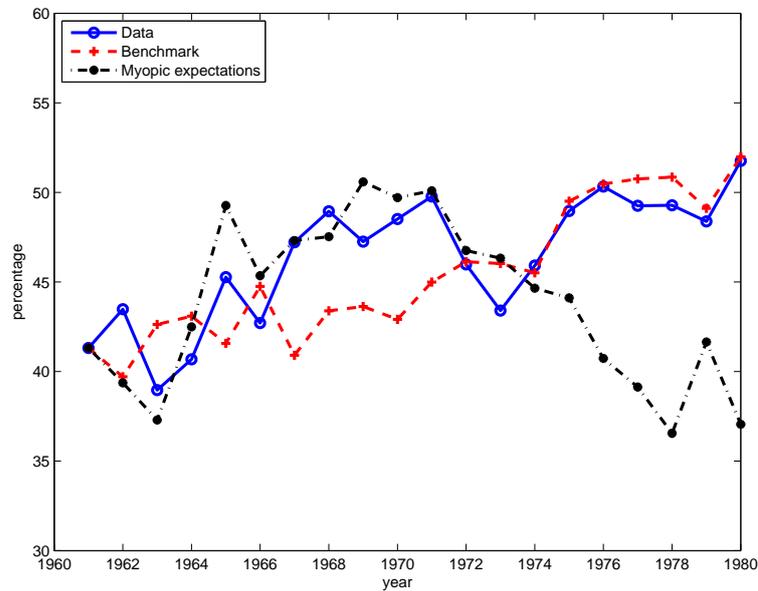


Figure 8 shows a very interesting pattern. The model with myopic expectations actually matches the data better from 1961 to 1972. It especially fits the data very well from 1966 to 1972, the period during which the benchmark model significantly underpredicts the data. However, the myopic expectations model predicts that the enrollment rate decreased from 46.8 percent in 1973 to 37.1 percent in 1980, while in the data it increased

from 43.4 percent to 51.8 percent. The benchmark model with perfect foresight, however, matches the data for the period 1973 to 1980. The reason the model with myopic expectations works better in the late 1960s and early 1970s is that, as shown in Figure 7 for the 1970 cohort, the compressed college wage premium in the 1970s did not impact the cross-sectional wage premium prior to the early 1970s cohorts, while for the perfect foresight benchmark model, cohorts in the late 1960s and early 1970s faced a decreasing wage premium in the early stage of their life-cycle, which reduced the incentive to go to college. However, as shown in Figure 7 for the 1980 cohort, the decreasing college wage premium in the 1970s reduced the cross-sectional earnings differential more severely than the life-cycle ones because for the latter, the rising wage premium since 1980 compensated for the loss in the 1970s. The benefits of college education are much lower for individuals who are short-sighted than for individuals with perfect foresight. The model with myopic expectations thus predicts a sharp drop in enrollment rates.

Figure 8 suggests that around the early 1970s, there might have been a dramatic change in the way women form their expectations of future earnings. Cohorts prior to 1970 simply did not expect a fall in the college wage premium. They started to learn about the decreasing wage premium in the early 1970s and moved from myopic to more forward-looking expectations. Interestingly, the timing coincides with the finding of Goldin, Katz and Kuziemko (2006), who claim that “rapidly changing expectations among young women concerning their future life-cycle labor force participation started in the late 1960s.” They argue that this change might be due to increasing female labor force participation rates, the legality and widespread acceptance of the “pill,” and the resurgence of feminism. The results in Figure 8 provide evidence of the changing expectations from a different angle.

5.3. Distribution of Disutility Costs

In the benchmark model, a specific assumption is made about the distribution of ability level i in the function of the disutility cost. To test if the results are robust to the choice of the distribution of ability, à la Restuccia and Vandenbroucke (2010), a more general distribution of ability, $i \sim \text{Beta}(A, C)$, is considered, where $i \in [0, 1]$ and A and C are the two positive parameters governing the shape of the distribution. Beta distribution is chosen for two reasons. First, the domain of a beta distribution is $[0, 1]$, which is identical to the domain for the ability level in the benchmark model. Second, beta distribution is a very flexible probability distribution. Depending on the values of two shape parameters, the probability density function (pdf) can be U-shaped, bell-shaped, strictly increasing or decreasing and it is not necessarily symmetric. The uniform distribution assumed

in the benchmark model is just a special case of the beta distribution with $A = C = 1$. Therefore, we can allow the distribution be quite different from the benchmark one to impose more discipline on the robustness check. With this new distribution, there are now three parameters in the function of disutility cost $\chi(i)$: b , A , and C . These three parameters are calibrated to match the college enrollment rates in 1955, 1961 and 1972, which are the three years that the benchmark model fits the data best.¹⁷ This requires solving three non-linear equations. For example, for 1961, the dynamic programming problem for the 1961 cohort is solved to obtain the value function difference $V_{1961}^c - V_{1961}^h$. The threshold ability level i_{1961}^* is then determined by

$$\chi(i_{1961}^*) = b\left(\frac{1}{i_{1961}^*} - 1\right) = V_{1961}^c - V_{1961}^h,$$

which gives $i_{1961}^*(b) = \frac{1}{\frac{V_{1961}^c - V_{1961}^h}{b} + 1}$ as a function of parameter b . The female college enrollment rate in 1961 is thus given by $e_{1961} = \Pr(i > i_t^*) = 1 - \text{cdf_beta}(i_{1961}^*(b), A, C)$. The calibration requires

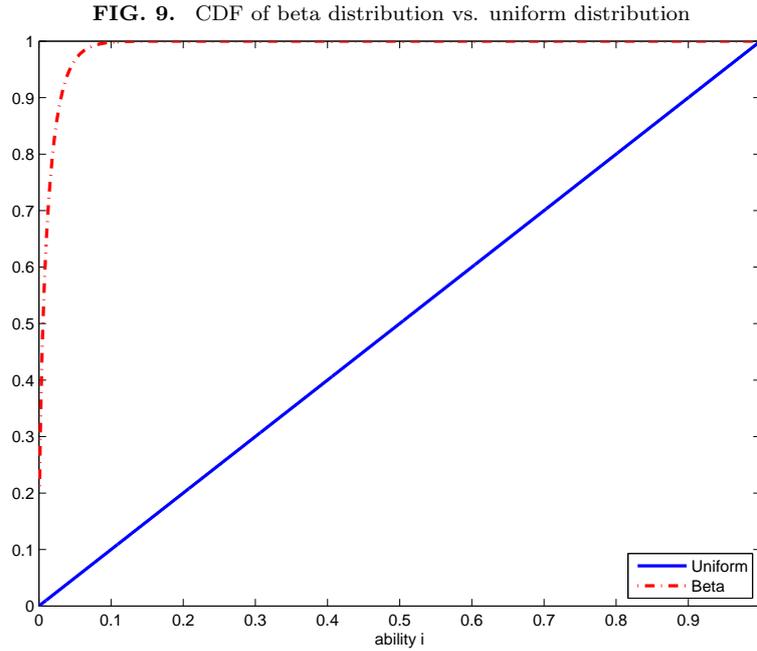
$$1 - \text{cdf_beta}(i_{1961}^*(b), A, C) = 41.32\% \text{ (data in 1961)}.$$

There are three such equations for 1955, 1961, and 1972 to solve three unknowns b , A and C . The calibration results in $b = 0.0248$, $A = 0.5368$, and $C = 46.1526$. The cumulative distribution function (cdf) of the beta distribution under the values of A and C is plotted in Figure 9 together with the cdf of uniform distribution, which is a straight line.

Although the cdf of the beta distribution is very different from that of the benchmark uniform distribution, Figure 10 shows that with this more general and flexible distribution of ability, the results are still surprisingly close to those of the benchmark model. This demonstrates that the model results are robust to the alternative distributional assumption. The assumption of uniform distribution used in the benchmark case is not critical in determining the results.

To summarize, the counterfactual experiments show that the rising tuition costs over time have little quantitative influence on the benchmark results. The results are also robust to a more general distribution of the disutility cost. Finally, the experiment testing different assumptions of expectations provides interesting evidence on young women's changing expectations of future earnings over the period 1955 to 1980.

¹⁷The results are not sensitive to the years that are chosen. For example, selecting either 1955, 1958 and 1961 or 1955, 1966 and 1980 produces similar paths of enrollment rates for the period from 1955 to 1980.

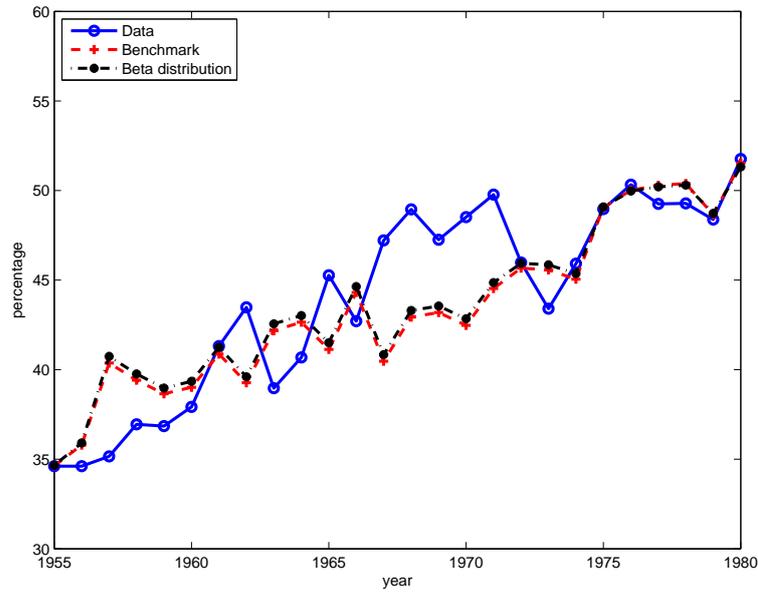


6. CONCLUSION

This paper develops a discrete time overlapping generations model with an endogenous college-entry decision. The decision is based on the cost-benefit analysis implied by the standard human capital investment theory. Two key features are the exogenous choice-dependent life-cycle wage profiles and an idiosyncratic disutility cost of a college education. Using this model, I quantitatively examine the driving force behind the dramatic increase in the female college enrollment rate from 1955 to 1980. I find that the model captures the rising female college enrollment rate during this period quite well. The rising college wage premium is the major driving force. The results also suggest that the change in expectations of future earnings among young women may have played an important role in driving the enrollment rate in the late 1960s and early 1970s.

The recent literature shows that the marriage market may be an important determinant in women's schooling decision.¹⁸ On the other hand, education may also affect women's fertility and marriage decisions. This paper does not address these issues. It, however, would be an interesting extension to include endogenous marriage and fertility choices in the current model to analyze the interaction among these choices. This extended

¹⁸See Ge and Yang (2010) and Ge (2010).

FIG. 10. Female college enrollment rate: beta distribution of ability

model will surely provide a platform for understanding not only the changes in women's college-entry decisions, but also the evolution of the marriage rate and fertility decisions over time. I leave that for future research.

REFERENCES

- Anderson, P. M., 2002. Where the Boys No Longer Are: Recent Trends in U.S. College Patterns. Mimeo, Department of Economics, Dartmouth College.
- Attanasio, O., H. Low, and V. Sanchez-Marcos, 2008. Explaining Changes in Female Labour Supply in a Life-Cycle Model. *American Economic Review* **98**(4), 1517–1552.
- Averett, S. L. and M. L. Burton, 1996. College Attendance and the College Wage Premium: Differences by Gender. *Economics of Education Review* **15**(1), 37–49.
- Becker, G., W. Hubbard, and K. Murphy, 2010. The Market for College Graduates and the Worldwide Boom in Higher Education of Women. *American Economic Review* **100**(2), 229–233.
- Bound, J. and S. Turner, 2002. Going to War and Going to College: Did World War II and the G.I. Bill Increase Educational Attainment for Returning Veterans? *Journal of Labor Economics* **20**(4), 784–815.
- Cameron, S. V. and C. Taber, 2004. Estimation of Educational Borrowing Constraints Using Returns to Schooling. *Journal of Political Economy* **112**(1), 132–182.
- Card, D. and T. Lemieux, 2001. Going to College to Avoid the Draft: The Unintended Legacy of the Vietnam War. *American Economic Review* **91**(2), 97–102.
- Charles, K. K. and Ming-Ching Luoh, 2003. Gender Differences in Completed Schooling. *Review of Economics and Statistics* **83**(3), 559–577.

- Eckstein, Z. and E. Nagypál, 2004. The Evolution of U.S. Earnings Inequality: 1961–2002. *Federal Reserve Bank of Minneapolis Quarterly Review* **28(2)**, 10–29.
- Fernandez, R., A. Fogli, and C. Olivetti, 2004. Mothers and Sons: Preferences Formation and Female Labor Force Dynamics. *Quarterly Journal of Economics* **119(4)**, 1249–1299.
- Garriga, C., and M. Keightley, 2007. A General Equilibrium Theory of College with Education Subsidies, In-School Labor Supply, and Borrowing Constraints. Federal Reserve Bank of St. Louis Working Paper, No. 2007-051A.
- Ge, S., 2010. Women’s College Choice: How Much Does Marriage Matter? *Journal of Labor Economics*, forthcoming.
- Ge, S. and F. Yang, 2010. Accounting for the Gender Gap in College Attainment. *Economic Inquiry*, forthcoming.
- Goldin, C., L. F. Katz, and I. Kuziemko, 2006. The Homecoming of American College Women: The Reversal of the College Gender Gap. *Journal of Economic Perspectives* **20(4)**, 133–156.
- He, H., 2009. What Drives the Skill Premium: Technological Change or Demographic Variation? Department of Economics, University of Hawai’i, Working Paper, No. 09-11.
- Heathcote, J., K. Storesletten, and G. Violante, 2010. The Macroeconomic Implications of Rising Wage Inequality in the United States. *Journal of Political Economy* **118(4)**, 681–722.
- Heckman, J. J., L. J. Lochner, and P. E. Todd, 2003. Fifty Years of Mincer Earnings Regressions. NBER Working Paper, No. 9732.
- Hendricks, L. and T. Schoellman, 2009. Student Abilities During the Expansion of U.S. Education, 1950–2000. Working Paper.
- Jacob, B. A., 2002. Where the Boys Aren’t: Non-Cognitive Skills, Returns to School and the Gender Gap in Higher Education. *Economics of Education Review* **21(6)**, 589–598.
- Katz, L. F. and K. M. Murphy, 1992. Changes in Relative Wages, 1963–1987: Supply and Demand Factors. *Quarterly Journal of Economics* **107(1)**, 35–78.
- Katz, L. F. and D. H. Autor, 1999. Changes in the Wage Structure and Earnings Inequality. In *Handbook of Labor Economics*, Vol. 3A. O. Ashenfelter and D. Card (eds.) Amsterdam: North-Holland.
- Krusell, P., L. E. Ohanian, J. V. Rios-Rull, and G. L. Violante, 2000. Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis. *Econometrica* **68(5)**, 1029–1053.
- Navarro, S., 2007. Using Observed Choices to Infer Agent’s Information: Reconsidering the Importance of Borrowing Constraints, Uncertainty and Preferences in College Attendance. Mimeo, Department of Economics, University of Wisconsin at Madison.
- Olivetti, C., 2006. Changes in Women’s Hours of Market Work: The Role of Returns to Experience. *Review of Economic Dynamics* **9(4)**, 557–587.
- Restuccia, D. and G. Vandenbroucke, 2010. The Evolution of Education: A Macroeconomic Analysis. Working Paper.
- Rios-Rull, J. and V. Sanchez-Marcos, 2002. College Attainment of Women. *Review of Economic Dynamics* **5(4)**, 965–998.
- Sanchez-Marcos, V., 2008. What accounts for the increase in college attainment of women? *Applied Economic Letters* **15(1)**, 41–44.