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

by Hui He†
Department of Economics
University of Hawaii at Manoa

Working Paper No. 10-16R
September 17, 2010

Abstract

This paper documents a dramatic increase in the college enrollment rate of women from 1955 to 1980 and asks a quantitative question: to what extent can such increase be accounted for by the change in the female cohort-specific college wage premium? I develop and calibrate an 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 attendance decision.

Key Words: Female College Enrollment rate, College Wage Premium, Life-cycle
JEL Classification: E24, J24, J31, I21

---

*This paper is a substantially revised version of an earlier paper listed as UHM Working Paper No. 09-12. This paper is based on the third chapter of my dissertation at the University of Minnesota. I thank Kelley 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.
† Mailing address: Department of Economics, University of Hawaii at Manoa, 2424 Maile Way, Saunders Hall 542, Honolulu, HI 96822 (e-mail: huihe@hawaii.edu)

1 Introduction

Female college attainment in the United States has changed dramatically over the last 50 years. In 1955, only 34.7% of college students were women. This ratio increased to 56.3% in 2001 (National Center for Educational Statistics, *Digest of Education Statistics* 2003, Table 174). The major reason behind this dramatic increase in female college attainment is 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% in 1955; however, it has been increasing since then. In 1980, this rate increased to 51.8%. In 2002, 68.4% of female high school graduates went to college.[1]

A well-known phenomenon in the US labor market is the rising college wage premium over the past 50 years (except that it decreased in the 1970s, Katz and Murphy 1992, Katz and Autor 1999). Data from the Current Population Survey (CPS) show that this is also true for females. The ratio of the annual mean wage of female college graduates to 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. The ratio, however,

[1] College enrollment rates for the 1960-2002 period are taken from the National Center for Educational Statistics, *Digest of Education Statistics* 2003, Table 186. The data for 1955-1959 were calculated by the author. See He (2009) Appendix A for the data construction.

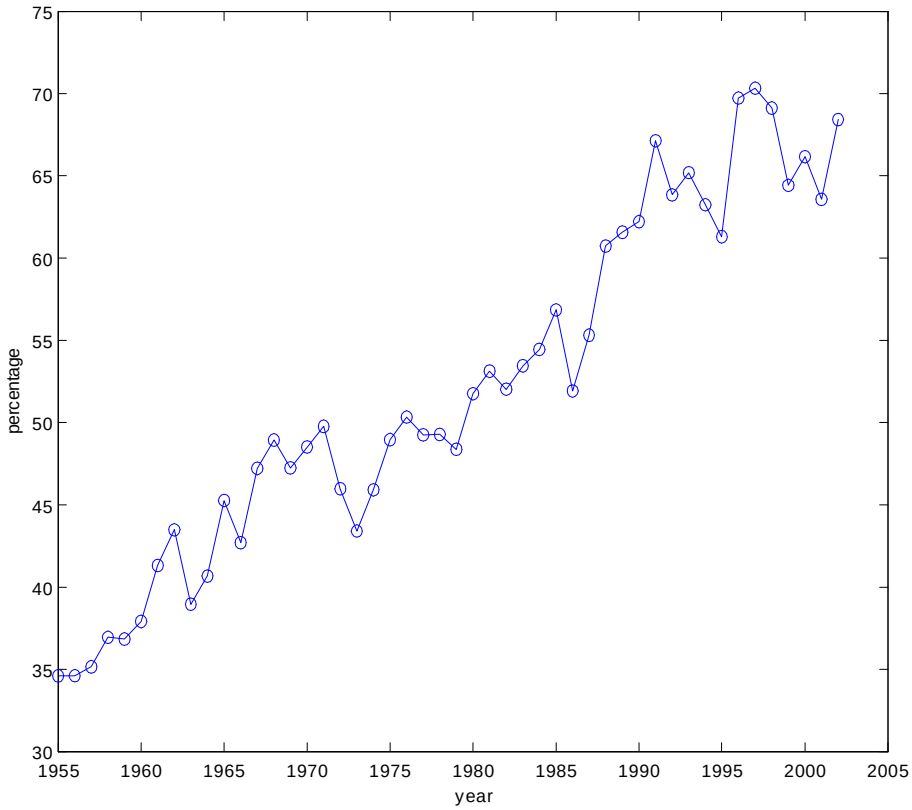

Figure 1: Female college enrollment rate of recent high school graduates

has increased dramatically since 1980. In 2001, it was 1.91. This data pattern of the aggregate wage premium also affects the cohort-based life-cycle female wage premium as shown in Section 3.1.

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 cohort-specific life-cycle college wage premium*? In order to answer this question, this paper develops and calibrates a discrete time overlapping generations model with endogenous college-entry decision. Different cohorts of women enter into the economy at age 18 and face the decision whether or not to go to college. The 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 between the rising female college enrollment rate and 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 been increasing for most of the time 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 works quite well in capturing the rising female college enrollment rate during the period 1955-1980. The rising college wage premium is the major driving force behind the substantial increase in women's college attainment. The results also suggest that the change in expectations of future employment opportunity and earnings among young women may have played an important role in driving the enrollment rate since the early 1970s.

This paper contributes to a large empirical literature on studying female college enrollment.[2] 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. These papers do not examine the time trend of the female college enrollment rate. Anderson (2002), however, tries to answer the same question by looking at different cohorts over time. She finds that an

[2] Most of this literature focuses on studying the gender difference in college attainment. 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, Card and Lemieux 2001) and they cannot be easily captured in the model, this paper ignores the male side and focuses only on female college-entry decisions.

important component 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 people's educational 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 a key role in explaining this reversal.

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.[3] 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 why the ratio of male to female college graduates (sex college attainment ratio, SCAR)

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

was so high in the 1970s. These papers, however, do not study 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 can account for a substantial fraction of the reduction in the college attainment gender gap. However, changes in marital status and marital sorting go in the direction of reducing the college attainment of women. However, she compares the SCAR only in 1976 and 1990 and does not examine the time path of the changes in college education attainment. Ge and Yang (forthcoming) 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, changes in parental education, and changes in the marriage market on changes in college attainment 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. 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, in spirit, is quite close to Ge and Yang (forthcoming), although it puts less emphasis on the effects of marriage and fertility but has a relatively rich structure in consumption and savings. The results show

that the change in life-cycle earnings is a key factor in accounting for the changes in the female college enrollment rate for the 1955-1980 period, which is the period Ge and Yang (forthcoming) do not cover.[4]

The reminder of the paper is organized as follows. Section 2 presents a simple model of the college attendance decision. Section 3 describes the data and analyzes some findings from the data. 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 the one used in He (2009). 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, which can be used either for consumption or for investment. There is no uncertainty in the model. Individuals have perfect foresight.

[4] Restuccia and Vandenbroucke (2010) use a model similar to the one in this paper to investigate the increase of 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 the data.

2.1 Demographics

The economy is populated by overlapping generations of finite-lived women with total measure one. Women enter into the economy (or are "born") with zero initial assets when they are age 18, which is the common age of high school graduates. I call them the birth cohort and model age as $j = 1$. They then live and work up to age J. To distinguish between the age of a cohort and calendar time, I use j for age, and t for calendar time.

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

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

where $c_{j,t+j-1}$ is consumption for the age-j woman at time $t + j - 1$. σ is the coefficient of relative risk aversion. $\frac{1}{\sigma}$ is the intertemporal elasticity of substitution. Since leisure does not enter into the utility function, each woman will supply all of

her labor endowment, which is normalized to be one.

2.3 Budget Constraints

A woman chooses go to college or not at the beginning of the first period. I use $s \in \{c, h\}$ to indicate this choice. If an individual chooses $s = h$, she ends up with a high school diploma and goes on the job market to work as an unskilled laborer and earns high school graduate wage sequence $\{w_j^h\}_{j=1}^J$. Or she can 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 college graduate wage sequence $\{w_j^c\}_{j=1}^J$. For the sake of simplicity, I assume there is no college dropout and no unemployment.[5]

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

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

In the first four periods, she pays tuition $tuition_{t+j-1}$, consumes $c_{j,t+j-1}$, and saves

[5] Ge and Yang (forthcoming) make the same assumption.

$a_{j,t+j-1}$. After graduation, she earns wage $w^c_{j,t+j-1}$ 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 pay tuition during the first four periods, and they pay back the loans later.[6]

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

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

2.4 Schooling Choice

I assume that different individuals within the birth cohort are endowed with different levels of ability. And ability affects only individuals' disutility cost of schooling.[7] The disutility cost of schooling is a strictly decreasing function of ability. Higher ability implies lower disutility cost.

Individuals are indexed by their ability level $i \in [0, 1]$. The CDF of the ability

[6] Cameron and Taber (2004) find no evidence that access to borrowing is an important component of schooling decisions.

[7] This is a common assumption in the literature. Ge and Yang (forthcoming) and Restuccia and Vandenbroucke (2010) have a similar assumption. Navarro (2007) empirically finds learning ability is the main determinant of this "psychic" cost and it plays a key role in determining schooling decisions.

distribution is denoted by F, $F(i_0) = \Pr(i \leq i_0)$. $DIS(i)$ represents the ability-related disutility cost for individual i. Notice that $DIS(i) \geq 0$ and $DIS'(i) < 0$. An individual i born at time t thus has the discounted lifetime utility

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

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 $DIS(i)$ 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. I use $UTIL^c$ to denote the discounted lifetime utility for college graduates; $UTIL^h$ denotes the discounted lifetime utility derived from physical consumption for high school graduates. $UTIL^c - UTIL^h$ represents the utility gain from attending college. Obviously, individual i will choose to go to college if $DIS(i) < [UTIL^c - UTIL^h]$, not to go if $DIS(i) > [UTIL^c - UTIL^h]$, and is indifferent if $DIS(i) = [UTIL^c - UTIL^h]$.

Since the borrowing constraint does not exist, the model implies

$$UTIL_t^c - UTIL_t^h \gtreqqless 0 \text{ iff } NPV_t \gtreqqless 0,$$

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 $w_{j,t+j-1}^c = 0, \forall j = 1, ...4$, students never work when they stay in college, and I can further decompose NPV into three components

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

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 $\{\frac{w_{j,t+j-1}^c}{w_{j,t+j-1}^h}\}_{j=1}^{J}$ is going to affect people's schooling decision. Other things being equal, an increase in the lifetime college wage premium raises the benefits of schooling, hence NPV. Higher NPV

induces a higher utility gain from schooling $UTIL_t^c - UTIL_t^h$. Given the stationary distribution of the disutility cost, a higher utility gain from schooling means more likely $DIS(i) < [UTIL^c - UTIL^h]$, which implies a higher enrollment rate.

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

$$DIS(i^*) = [UTIL^c - UTIL^h].$$

Therefore, women with ability $i < i^*$ will choose not to go to college, while women with $i > i^*$ will choose to go.[8] The enrollment rate thus is equal to the probability when $i > i^*$. 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 $i^{*\prime}$. Since $\Pr(i > i^{*\prime}) > \Pr(i > i^*)$, more women go to college. A higher life-cycle earnings differential thus encourages college attendance.

[8] Ge and Yang (forthcoming) and Restuccia and Vandenbroucke (2010) also have a similar threshold mechanism to determine the schooling choices in their model.

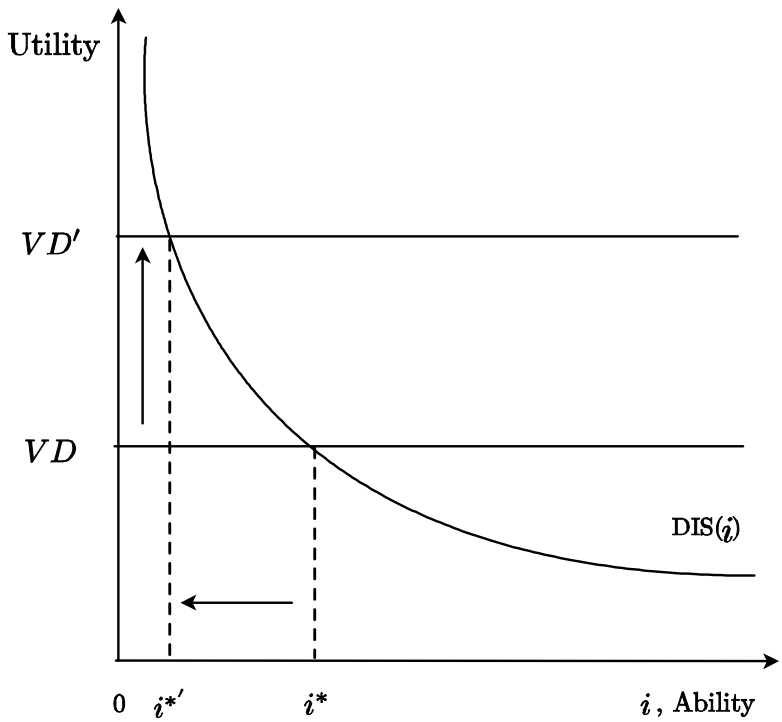

Figure 2: The determination of the college enrollment rate

2.5 Dynamic Programming Representation

For purposes of computation, it is easier to write an individual's schooling decision problem in terms of dynamic programming language. Let $V^c_{t+j-1}(a_{j-1,t+j-2}, j)$ denote the value function of an age-j woman with asset holding $a_{j-1,t+j-2}$ at the beginning of time $t+j-1$ who chooses to go to college at age $j=1$. It is the solution to the dynamic problem

$$V^c_{t+j-1}(a_{j-1,t+j-2}, j) = \max_{\{c_{j,t+j-1}, a_{j,t+j-1}\}} \{u(c_{j,t+j-1}) + \beta V^c_{t+j}(a_{j,t+j-1}, j+1)\} \tag{6}$$

subject to the budget constraint (2).

For a woman who chooses not to attend college, the corresponding value function is given by

$$V^h_{t+j-1}(a_{j-1,t+j-2}, j) = \max_{\{c_{j,t+j-1}, a_{j,t+j-1}\}} \{u(c_{j,t+j-1}) + \beta V^h_{t+j}(a_{j,t+j-1}, j+1)\} \tag{7}$$

subject to budget constraint (3).

Individuals solve their perfect foresight dynamic problem by using backward induction. Back to age 1 at time t, a woman with ability index i will make her schooling

decision $s_{i,t}$ based on the criteria below

$$
\begin{aligned}
s_{i,t} &= c \quad \text{if } V_t^c(a_{0,t-1}=0,1) - DIS(i) > V_t^h(a_{0,t-1}=0,1), \\
s_{i,t} &= h \quad \text{if } V_t^c(a_{0,t-1}=0,1) - DIS(i) < V_t^h(a_{0,t-1}=0,1), \\
s_{i,t} &= \text{indifferent} \quad \text{if } V_t^c(a_{0,t-1}=0,1) - DIS(i) = V_t^h(a_{0,t-1}=0,1).
\end{aligned} \tag{8}
$$

3 Data

I use the March Current Population Survey (CPS) from 1962 to 2003 to construct the data counterparts in the model. I choose the sample restrictions to follow those used in Eckstein and Nagypál (2004) except I further restrict the data to include only high school graduates (HSG hereafter) between ages 18 and 65 and college graduates (CG) between ages 22 and 65 in the sample. As Eckstein and Nagypál's paper, I restrict my attention to full-time full-year (FTFY) workers. The wage here is the annualized wage and salary earnings. I use the personal consumption expenditure deflator from NIPA to convert all wages in terms of constant 2002 dollars.

3.1 Cohort-Specific Wage Premium

In the model, women in different cohorts make the educational decision based on the expected earnings differential specific to their cohort. The perfect foresight assumption allows me to use *actual* observed future earnings in the CPS as the measure of expected future earnings. Since the CPS is a repeated cross-sectional data set, I use a so-called "pesudo-cohort construction method" to construct the cohort-specific expected wage profiles.[9] For example, the 1962 cohort's (18-year-old HSG in 1962) lifetime (18-65 years old) female HSG wage profile $\{w^h_{j,1961+j}\}_{j=1}^{48}$ is constructed as follows: I take 18-year-old female HSGs in 1962, calculate their mean wage, then 19-year-old female HSGs in 1963, calculate the mean wage, then 20-year-old female HSGs in 1964, 21-year-old female HSGs in 1965, and so on, until I reach 58-year-old female HSGs in 2002, which is the end year of my CPS data set.

I use a similar approach to construct the 1962 cohort's female CG wage profile $\{w^c_{j,1961+j}\}_{j=1}^{48}$. But I start from 1966 because if someone from the 1962 cohort chooses to go to college, she needs to spend four years in college. She graduates in 1966 and starts to earn CG wages from that year. Therefore, I take 22-year-old female CGs in 1966, calculate their mean wage, then calculate the mean wage for 23-year-old

[9] It is a pesudo-cohort because the CPS is not a panel data set. It does not track people over their lifetimes. Heckman, Lochner, and Todd (2003) use a similar method to estimate the cohort-based return to schooling.

female CGs in 1967, and so on.

Figure 3 shows the life-cycle wage profiles for six cohorts. They are the 1955, 1960, 1965, 1970, 1975 and 1980 cohorts. For each cohort, the wage profile of CGs is significantly higher than that of HSGs. Two facts about the life-cycle wage profiles need to be mentioned here: (1) Earnings rise with age, but at a decreasing rate; (2) Earnings increase faster for more educated workers, which implies CGs have a steeper hump-shaped (or increasing but concave) wage profile than that of HSGs. Notice that for the 1955 and 1960 cohorts, the late-age earnings for CGs become quite noisy. This is due to the smaller sample size for CGs at the later age.

The college wage premium over the life-cycle $\{\frac{w^c_{j,t+j-1}}{w^h_{j,t+j-1}}\}_{j=1}^{48}$ exhibits some interesting patterns for these cohorts. Due to data availability, I calculate the wage premium only from age 22 to age 40. The average college wage premium from age 22 to age 40 for the 1955 cohort was 1.45. For the 1960 cohort, it was 1.48. It then decreased significantly to 1.38 for the 1965 cohort and 1.39 for the 1970 cohort. For these two cohorts, the compressed college wage premium in the 1970s significantly reduced their earnings differential at their prime age when the CG wage profile is in a stage of steep ascent. In contrast, the rising college wage premium starting from 1980 helped to increase the average college wage premium from age 22 to 40 for the 1975 and 1980 cohorts to 1.56 and 1.63, respectively.

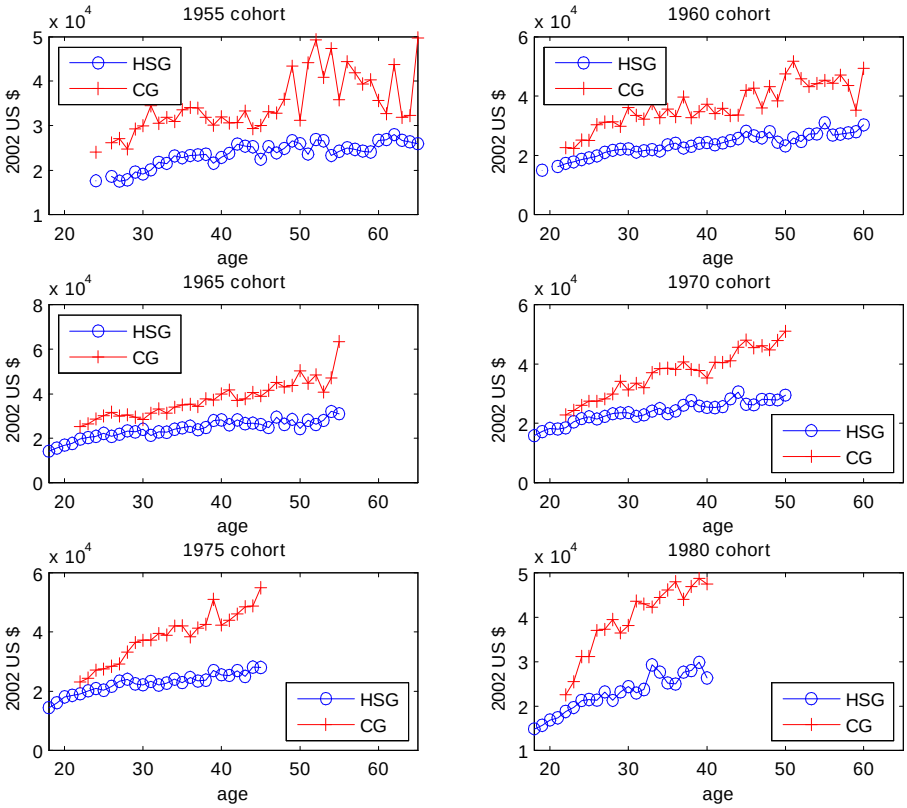

Figure 3: Life-cycle wage profiles for six cohorts

Following the pesudo-cohort construction method, I am able to construct life-cycle wage profiles for HSGs and CGs from the 1955 to the 1980 cohort. However, 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). I use the econometric method to predict the mean wage at that specific age to extrapolate the missing data. I predict them by either second- or third-order polynomial specification or a conditional Mincer equation as follows:

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

My extrapolation stops after the 1980 cohort because after this cohort, a lack of data points creates trouble; hence, I do not have a reliable prediction.[10] By filling in the missing data, eventually I obtain complete cohort-specific life-cycle wage profiles for HSGs and CGs from the 1955 to the 1980 cohort.

These cohort-specific life-cycle wage profiles provide the information needed in the first two terms of equation (5). To fully understand the higher education choice

[10] The 1980 cohort has life-cycle wage profiles 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.

over time, I also need information about tuition, which is the direct cost of college education as in the third term of equation (5). In Figure 4 I report the real tuition, fees, room and board (TFRB) per student charged by an average four-year institution in terms of constant 2002 dollars.[11] TFRB increased over time except in the 1970s, when it became stable. Starting in 1980, real TFRB increased dramatically. Different cohorts face different TFRB charges based on the years during which they attended college.

4 Results

In this section, the economic model in Section 2 is calibrated to generate the female college enrollment rate. The calibration strategy I employ here is to adopt the common values widely used in the literature for the preference parameters. For the model-specific disutility parameter b (see equation (9) below), I calibrate it to match the enrollment rate data of the initial cohort (HSGs in 1955). Under these calibrated parameter values, I then input the data of cohort-specific life-cycle wage profiles and real TFRB in Section 3 into the model. Given that I have enough information about individuals' budget constraints, I am able to solve their dynamic programming problem as in (6) and (7). Finally, I use the criterion in (8) to determine

[11] See He (2009) for the explanation to construct this data sequence.

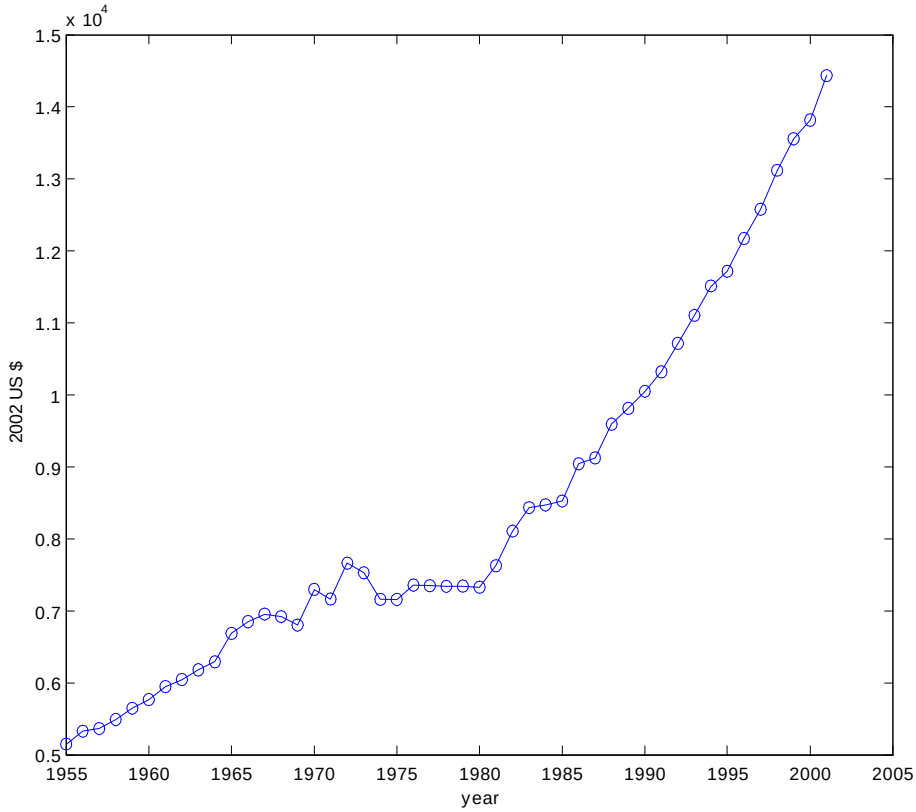

Figure 4: Real tuition, fees, room and board charge per student

the enrollment rate in different years.

4.1 Calibration

The value of discount factor β is taken to be 0.96 to match the interest rate r, which is set to 4%.[12] The value of the CRRA coefficient σ is 2, which is widely used in the life-cycle literature.

For simplicity, I assume that the ability level i is uniformly distributed among women and the ability-related disutility cost takes the form

$$DIS(i) = b(\frac{1}{i} - 1). \tag{9}$$

I assume that this cost function is also time invariant. For the lowest ability individual ($i = 0$), $DIS(i) = \infty$, so she will never go to college. On the other hand, for the highest ability individual ($i = 1$), $DIS(i) = 0$. 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. Given this functional form, as is shown in Section 2, the college enrollment rate at time t is determined by the threshold level i_t^* as in

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

Parameters	Description	Value
β	discount factor	0.96
σ	CRRA coefficient	2
b	scale factor of disutility cost function	7.28
r	real interest rate	4%

Table 1: Parameter values of the model

the following equation

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

Since ability level is uniformly distributed and $DIS'(i) < 0$, the enrollment rate at time t is equal to $\Pr(i > i_t^*) = 1 - F(i_t^*) = 1 - i_t^*$.

I calibrate the scale factor of disutility cost function b to match the female college enrollment rate data in 1955, which is 34.6%. This ends up with $b = 7.28$.

Table 1 summarizes the parameter values used in the model.

4.2 Findings

Based on the parameter values shown in Table 1, at year t I solve the birth cohort's dynamic problem using a standard numerical method to obtain the difference in value function $VD_t = V_t^c(a_{0,t-1} = 0, 1) - V_t^h(a_{0,t-1} = 0, 1)$. This difference determines a unique threshold level i_t^* and hence the corresponding college enrollment rate at year

t (e_t)

$$e_t = \Pr(i > i_t^*) = 1 - i_t^*.$$

Figure 5 compares the female college enrollment rates from the model with those in the data from 1955 to 1980. The model replicates the rising trend of college enrollment rates for females very well. In the data, the enrollment rate increased from 34.6% in 1955 to 51.8% in 1980; in the model, it increased from 34.6% to 52.0%.

Looking at a detailed comparison, from 1955 to 1960, the model predicts that the enrollment rate increased from 34.6% to 39.2%; in the data it increased from 34.6% to 38.0%. The model actually overshoots the data for this period. It implies that some 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 quite in line with the data. However, from 1967 to 1972, the model significantly underpredicts the college enrollment rate for females. In 1967, the enrollment rate was 47.2% in the data, while the model predicts a rate of only 40.6%, which was 6.6% lower than the data. Similarly, in 1970, the college enrollment rate was 48.5% in the data, while it was 43.2% in the model.

One possible reason for the underprediction of the female college enrollment rate in the model during that period is that the model does not capture the changing

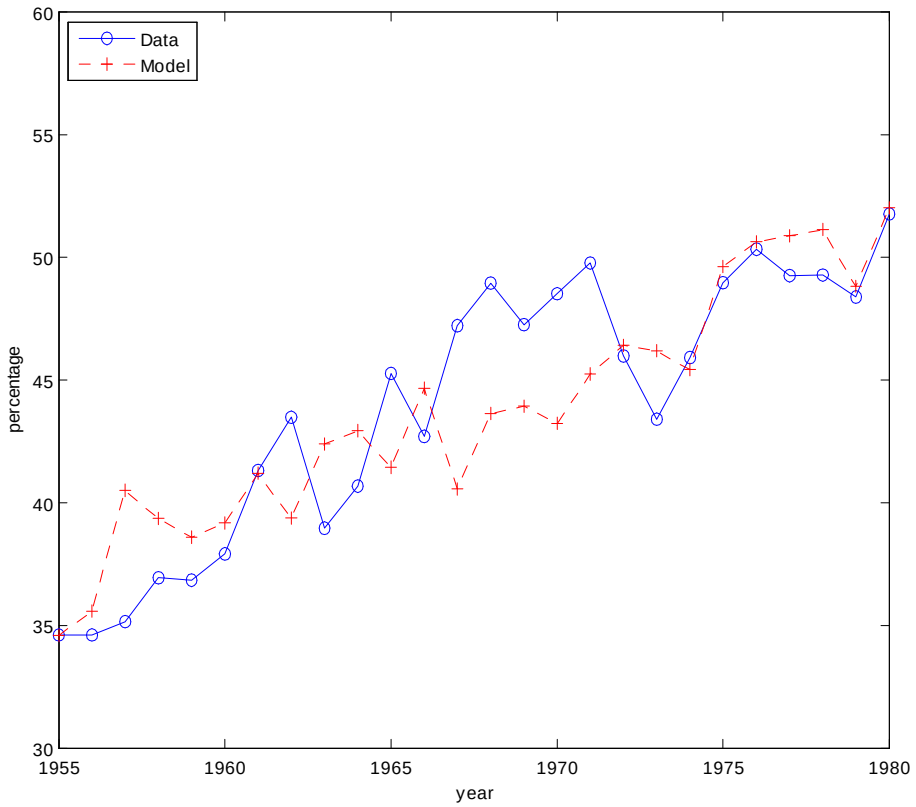

Figure 5: Female college enrollment rate: model vs. data

social norms and expectations about the role of work among young women in the late 1960s and early 1970s. As evidence of this changing social norm, Astin, Oseguera, Sax, and Korn (2002) report the results of the Astin Freshman Survey, which is a national sample of college freshmen, the vast majority of whom were 18 years old. In the survey, the freshmen agreed or disagreed with the statement: "The activities of married women are best confined to the home and family." The fraction of female freshmen disagreeing with this statement increased dramatically from 59% in 1967 to 83% in 1973. The fraction, however, has been quite stable since then.

Rising expectations of future employment certainly encouraged girls to go to college. Based on the observation from the Astin Freshmen Survey and their own data, Goldin, Katz, and Kuziemko (2006) estimate that the change in expectations about future labor participation would account for a 4.8 to a 5.7 percentage point increase in the female college graduation rate from 1968 to 1979. The college enrollment rate in the current model from 1967 to 1972 on average is below that in the data by 5 percentage points. Since in the model the college enrollment rate is equal to the college completion rate, changes in expectations among young women could capture the entire difference between the model's prediction and the data.[13]

Since 1972, the model has done a very good job of replicating the data. In the

[13]The counterfactual experiment in Section 5.2 provides further evidence for changing expectations of future earnings among young women since the late 1960s.

data, the female college enrollment rate increased from 46.0% in 1972 to 51.8% in 1980. The model counterpart was from 46.5% to 52%. Girls who graduated from high school around that time had already witnessed a drastic increase in the female labor force participation rate and formed their expectations accordingly; therefore, it is not surprising to see that the decision to go to college is entirely driven by economic concerns. The higher college wage premium for females since 1980 has raised the benefits of attending college as shown in the first term of equation (5). It was a significant factor in encouraging girls to go to college.

To summarize, the results show that the human capital investment model works quite well to capture the rising female college enrollment rate from 1955 to 1980. The results also suggest that the changing expectations of future employment opportunity among young women may also play an important role in driving this enrollment rate especially during the late 1960s and early 1970s.

5 Counterfactual Experiments

In this section, I run a series of counterfactual experiments to quantify the effects of changing tuition costs, the assumption of perfect foresight, time-varying disutility costs, and the distribution of disutility costs on the female college enrollment rate.

5.1 Fixed Tuition Cost

In order to quantify the effects of changing tuition costs over the target period, I fix tuition costs at the level of the 1955 cohort. Therefore, the 1956-1980 cohorts face the same tuition costs as the 1955 cohort. 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 over the period. Therefore, the direct cost of schooling apparently is not a significant factor in determining women's college entrance behavior.[14]

5.2 Naive Expectations

To test whether the benchmark results are sensitive to the assumption of perfect foresight, I change the assumption to naive expectations, which means 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, for the women in the 1970 cohort, I do not input their cohort-specific wage profiles in the model. I rather use the cross-sectional wage profiles for ages 18-65 in 1970 to proxy their life-cycle wage profiles. Due to data availability, I have annual cross-sectional wage profiles from 1961 to 2002 from the CPS. In order to compare this with the benchmark

[14] Ge and Yang (forthcoming) obtain similar results.

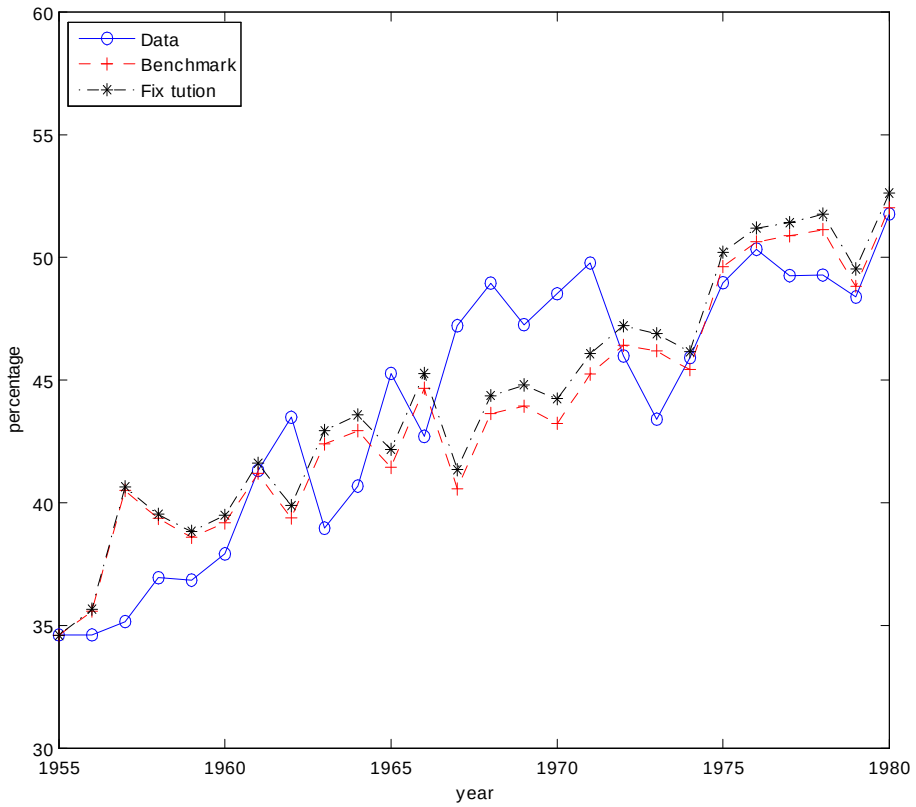

Figure 6: No change in tuition cost

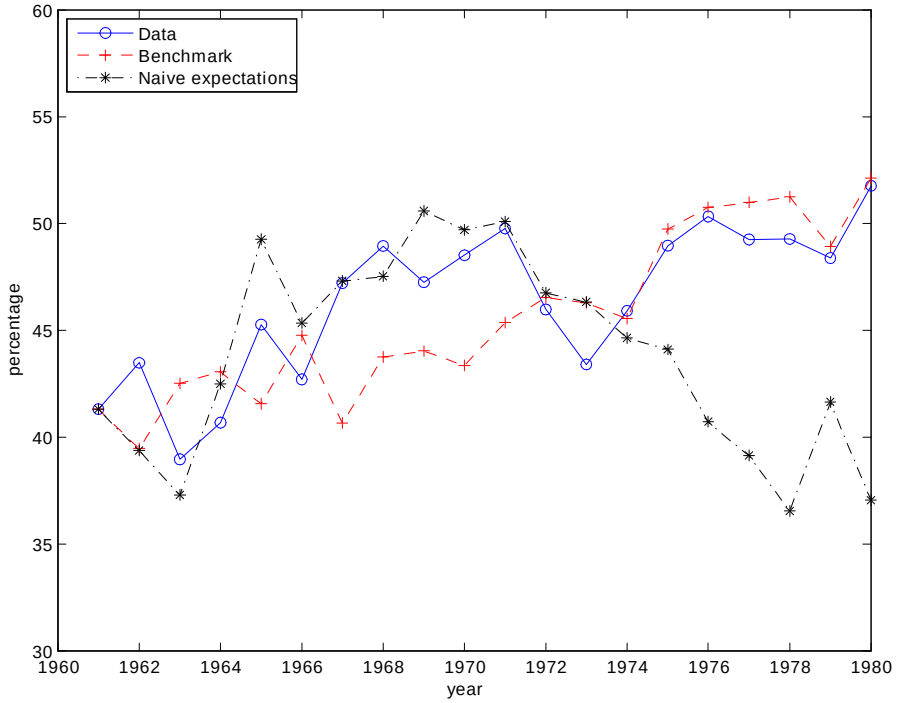

Figure 7: Naive expectations vs. perfect foresight

case, I input the profiles for the period 1961-1980 in the model. The results are presented in Figure 7.[15]

Figure 7 shows a very interesting pattern. The model with naive expectations actually matches the data better from 1961 to 1972. It especially fits the data from 1966 to 1972 when the benchmark model significantly underpredicts the data. However, It predicts that the enrollment rate decreased from 46.3% in 1973 to 37.1%

[15]For the purposes of comparison, I calibrate scale factor b differently to match the enrollment rate in 1961 in the benchmark model and the model with naive expectations respectively.

in 1980, while in the data it increased from 43.4% to 51.8%. The benchmark model with perfect foresight, however, matches the data for the period 1973 to 1980. The reason the model with naive expectations works better in the 1960s is that the cross-sectional wage premium increased in that decade, while for the perfect foresight benchmark model, cohorts in the 1960s faced a decreasing wage premium in the early stage of their life-cycle, which reduced the incentive to go to college. However, the decreasing college wage premium in the 1970s reduces the cross-sectional earnings differential more severely than the life-cycle ones because for the latter the rising wage premium since 1980 compensates for the loss in the 1970s. For the model with naive expectations, individuals are short sighted and they feel the compressed wage premium only when they make the decision to attend college. For the benchmark model with perfect foresight, the rising college wage premium since 1980 significantly raised the future earnings differential for college-age women in the 1970s and hence encouraged them to go to college.

The finding in Figure 7 suggests that around the early 1970s there is a dramatic change in the way women form their expectations of future earnings. It moves from myopic to more forward looking. 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 7 provide evidence of the changing expectations from a different angle.

5.3 Changing Disutility Costs

In the third experiment, I investigate the importance of the assumption of stationary distribution of the disutility cost $DIS(i)$. In particular, I ask the question: given the cohort-specific wage profiles, in order to exactly match the female college enrollment rate in the data, how should the scale factor of disutility cost b change over time?

Figure 8 shows the level of b for each year from 1955 to 1980 that matches the female college enrollment rate in the data. The trend line is slightly decreasing over the years. However, the t-statistics of the OLS regression coefficient are not significant.[16] We cannot reject the assumption that b is constant over time.

[16] The OLS regression results are (standard errors are in parentheses)

$$b = \underset{(26.8220)}{28.3390} - \underset{(0.0136)}{0.0122} * year.$$

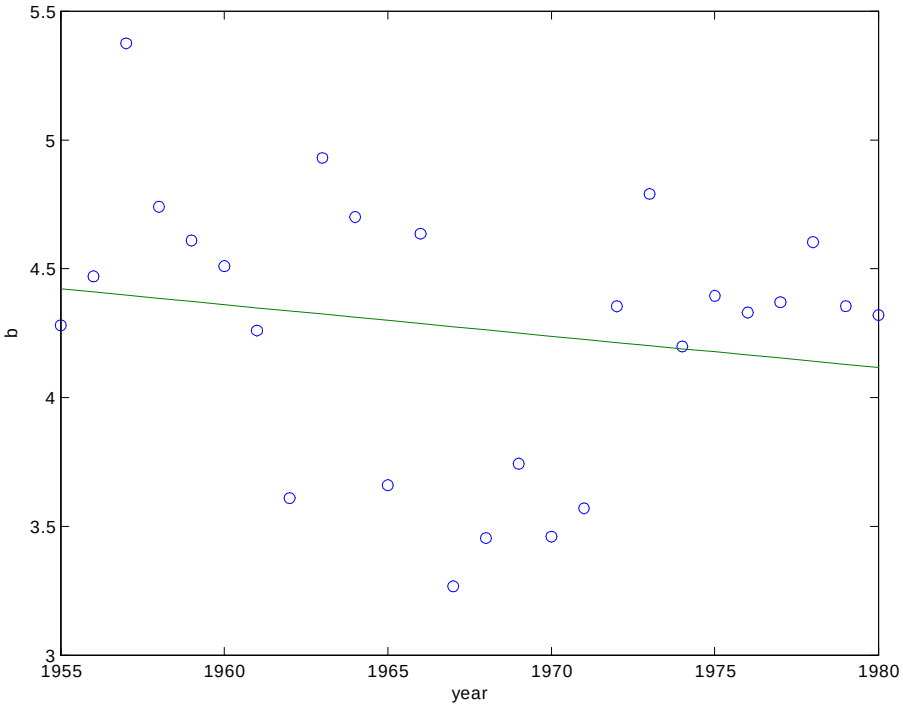

Figure 8: Changing scale factor of the disutility cost

5.4 Distribution of Disutility Costs

In the benchmark model, we make a specific assumption about the distribution of ability i in the function of the disutility cost. To test if the results are robust to the choice of the distribution of ability, I follow Restuccia and Vandenbroucke (2010) to consider a more general Beta distribution of ability. $i \sim Beta(A, C)$, where $i \in [0, 1]$ and A and C are the two positive parameters governing the shape of the distribution. The uniform distribution I assume in the benchmark model is just a special case of the Beta distribution with $A = C = 1$. Now there are three parameters in the function of disutility cost $DIS(i)$: b, A, and C. I calibrate these three parameters to match the college enrollment rates in three years over the period 1955-1980.[17]

Figure 9 shows that with this more general distribution of ability, the results are still quite close to the ones in the benchmark model. The model results are robust to the alternative distributional assumption.

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. The experiments justify the assumption that the disutility cost of schooling is quite time-invariant.

[17] In the current version, I calibrate b, A, and C to match the enrollment rate data in 1955, 1961 and 1974, which are the three years for which the benchmark model fits the data best. The results, however, are not quite sensitive to the years I pick.

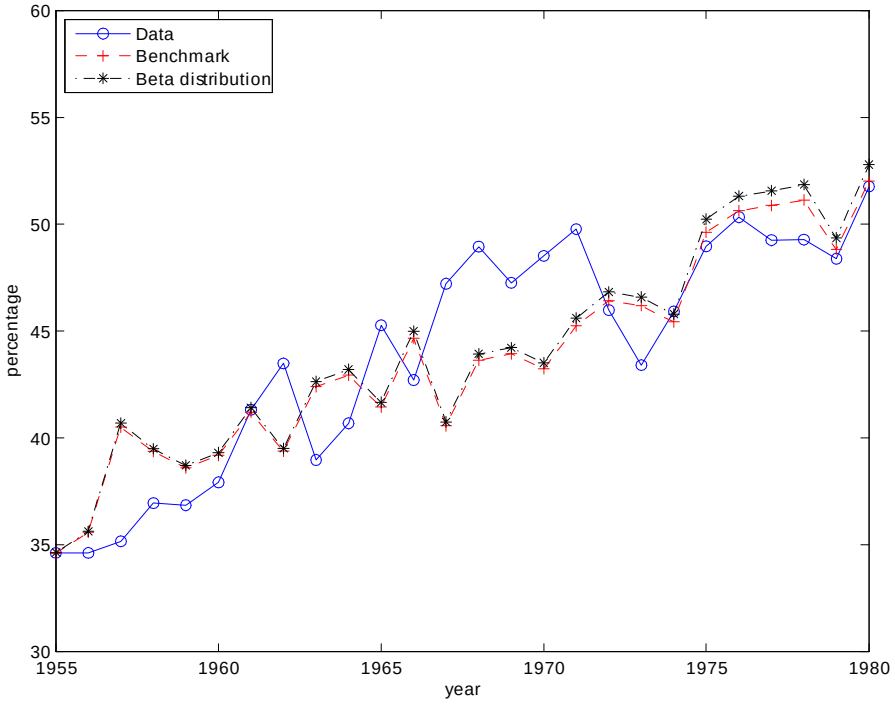

Figure 9: Beta distribution of ability

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.

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 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 works quite well in capturing the rising female college enrollment rate during this period. The rising college wage premium is the major driving force. The results also suggest that the change in expectations of future employment opportunity and earnings among young women may have played an important role in driving the enrollment rate since 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] Education may also affect women's fertility

[18] See Ge and Yang (forthcoming) and Ge (2010).

and marriage decisions. This paper does not address these issues. However, it 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 model surely will 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

[1] Anderson, P. M. (2002), "Where the Boys No Longer Are: Recent Trends in U.S. College Patterns," Department of Economics, Dartmouth College, working paper.

[2] Astin, A. W., L. Oseguera, L. J. Sax and W. S. Korn (2002), *The American Freshman: Thirty-Five Year Trends, 1966–2001.* Los Angeles, CA: Higher Education Research Institute, Graduate School of Education and Information Studies, University of California, Los Angeles.

[3] 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.

[4] 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.

[5] 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.

[6] Cameron, S. V. and C. Taber (2004), "Estimation of Educational Borrowing Constraints Using Returns to Schooling," *Journal of Political Economy* 112(1): 132-182.

[7] 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.

[8] Charles, K. K. and Ming-Ching Luoh (2003), "Gender Differences in Completed Schooling," *Review of Economics and Statistics*, 83(3): 559-577.

[9] 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.

[10] 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.

[11] Ge, S. (2010), "Women's College Choice: How Much Does Marriage Matter?" Department of Economics, Virginia Tech, Working Paper.

[12] Ge, S. and F. Yang (forthcoming), "Accounting for the Gender Gap in College Attainment," *Economic Inquiry*.

[13] 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.

[14] He, H. (2009), "What Drives the Skill Premium: Technological Change or Demographic Variation?" Department of Economics, University of Hawaii, Working Paper.

[15] Heckman, J. J., L. J. Lochner and P. E. Todd (2003), "Fifty Years of Mincer Earnings Regressions," NBER Working Paper 9732.

[16] 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.

[17] 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.

[18] Katz, L. F.and D. H. Autor (1999), "Changes in the Wage Structure and Earnings Inequality," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, vol. 3A, North-Holland, 1999, 1463-1555.

[19] Navarro, S. (2007): "Using Observed Choices to Infer Agent's Information: Reconsidering the Importance of Borrowing Constraints, Uncertainty and Preferences in College Attendance," manuscript, University of Wisconsin at Madison.

[20] Olivetti, C. (2006), "Changes in Women's Hours of Market Work: The Role of Returns to Experience," *Review of Economic Dynamics*, 9: 557-587.

[21] Restuccia, D. and G. Vandenbroucke (2010), "The Evolution of Education: A Macroeconomic Analysis," Working Paper.

[22] Rios-Rull, J., and V. Sanchez-Marcos (2002), "College Attainment of Women." *Review of Economic Dynamics*, 5(4): 965–998.

[23] Sanchez-Marcos, V. (2008), "What accounts for the increase in college attainment of women?" *Applied Economic Letters*, 15(1): 41-44.