Accounting for the Gender Gap in College Attainment*

Suqin Ge† Fang Yang‡

Virginia Tech SUNY-Albany

February 25, 2009

Abstract

One striking phenomenon in the U.S. labor market is the reversal of the gender gap in college attainment. Females have outnumbered males in college attainment since 1987. We develop a discrete choice model of college entry decisions to study the effects of changes in relative earnings, changes in parental education, and changes in the marriage market on time series observations of college attainment by gender. We find that the increase in the relative earnings between college and high school individuals and the increasing parental education have important effects on the increase in college attainment for both genders but cannot explain the reversal of the gender gap. Declining marriage rates decrease returns to college for females less than those for males, and thus is crucial in explaining the reversal of the gender gap in college attainment.

Keywords: gender, education, marriage, women, intergenerational schooling persistence

JEL Classification: J24, J16, I20

---

*We thank Michele Boldrin, Betty Daniel, Mariacristina De Nardi, Zvi Eckstein, John Jones, Michael Sattinger, and seminar participants at Virginia Tech, 2008 AEA meetings, University of Virginia, SUNY-Albany, 2008 North American Summer Meeting of the Econometric Society, Xiamen University, W.E. Upjohn Institute for Employment Research, and the Philadelphia Fed for helpful comments and suggestions. We are grateful to our discussant, Raquel Fernandez, at 2008 AEA meetings. Ge acknowledges the AEA/CSWEP summer fellowship from W.E. Upjohn Institute for Employment Research. All remaining errors are our own.

†Ge: Department of Economics, Virginia Tech, Blacksburg, VA 24061 (email: ges@vt.edu).

‡Yang: Department of Economics, University at Albany, Albany, NY 12222 (email: fyang@albany.edu).

1 Introduction

One striking phenomenon in the U.S. labor market is the reversal of the gender gap in college attainment. In 1980, 57 percent of young men aged 25 to 34, as compared with 46 percent of young women, had some college education by age 34. By 1996, however, female college attainment had reached 64 percent, 5 percentage points higher than that of males in the same cohort. In fact, females overtook males in college attainment in 1987 and have led ever since.

There is a large body of empirical research that emphasizes the role of the earnings premium as a key explanatory variable for the determination of education outcomes (see, for example, Becker 1967; Mincer 1974; and Willis and Rosen 1979). In addition, an extensive literature shows that family background is an important determinant of the schooling decision (see, among others, Kane 1994; Cameron and Heckman 1998, 2001; Eckstein and Wolpin 1999; and Ge 2008). Recently there have been several papers that argue, empirically and theoretically, that expected marriage is important in determining schooling decision (e.g., Chiappori, Iyigun, and Weiss 2006; and Ge 2008).

Based on this literature, we construct a formal economic model that includes the potential costs and benefits from the labor market and marriage market which determine individual college decisions. In our model, individuals with differing learning abilities first decide whether or not to enter college. Then they might get married and have children. Parents are altruistic and value their children's learning ability, which increases with the parents' education. Forward-looking individuals take into account the impact of their own schooling on their children's learning ability. Other factors that affect an individual's decision on whether to pursue higher education include the expected direct labor market returns to college over one's lifetime, the expected marriage market returns to college, and the costs in effort to attend college. These costs and benefits can differ by gender.

We estimate the parameters of the model by matching data on aggregate college attainment and college attainment conditional on parents' education by gender from the Panel Study of Income Dynamics (PSID). We present evidence on how well the model fits the data. We then use the parameter estimates to simulate counterfactual experiments, which break down the sources of changes in college attainment into the effects of changes in relative earnings, changes in parental education, and changes in the marriage market.

We calculate marriage distributions by education, lifetime earnings by education and marital status, and parents' education distributions as exogenous inputs of the model. In the marriage market, there has been a substantial decline of marriage probabilities for both genders, regardless of college attainment status. In addition, the probability of marrying a college spouse, regardless of one's own education level, has increased over time for males and decreased for females. Lifetime earnings by cohort are decreasing slowly for males of all marital statuses, especially for married males. Lifetime earnings for married females are increasing gradually, while those for single females are decreasing. We observe that the number of college-educated parents increases over time. To formally endogenize those changes is beyond the scope of our paper. We instead focus on the mechanism in which, under perfect foresight, these changes affect education decisions.

What accounts for the increase in college attainment over the past few decades? We find that the increasing gap in earnings between college and high school graduates has important effects on the increase in college attainment for both genders. When earnings are fixed at 1980 levels, attainment rates in 1996 drop by 15.5 and 17.4 percentage points for males and females, respectively. We also emphasize the importance of intergenerational persistence in schooling on the increase in college attainment for both genders. If the parents' schooling distribution is fixed at its 1980 level, college attainments in 1996 drop by 7.8 and 7.0 percentage points for males and females. The model endogenously generates the pattern that a college-educated parent is substantially more likely to have a college-educated daughter or son than a noncollege graduate, even after controlling for the education of the other parent. This link between parents' and children's schooling provides an intergenerational propagation mechanism: as the number of college-educated parents increases, their children become more likely to attend college. Thus, the gradual transformation of parental education acts as a mechanism to propagate changes in college attainment.

What accounts for females in the last generation overtaking males in college attainment? We find that decreasing marriage probabilities are crucial in explaining the relative increase in female college attainment. The decline in marriage probabilities decrease college attainment for both genders because the return to college from improving children's learning ability for married couples dominates their lower earnings return in the labor market. Under our parameters, the decline in the marriage probability decreases returns to college for females less than those for males. Therefore, *ceteris paribus*, college attainment rate of females declines less than that of males. Without the observed changes in marriage probabilities, females' college attainment would have been always lower than that of males. Two factors are relevant here. First, among married persons, the returns to college education are higher for males than those for females. Second, among singles, the return to college education is higher for females than for males. As marriage probabilities decrease, the returns to college for single females become high enough to compensate for the low returns to college for married females, and female college enrollment exceeds that of males.

This paper contributes to an active and growing literature on gender differences in educational attainment. Several papers have studied college enrollment and graduation by gender for one cohort. Averett and Burton (1996) focus on the gender differences in college enrollment for young individuals in 1979. Rios-Rull and Sanchez-Marcos (2002) construct a model to explain why males had higher college attainment than females in the 1970s. Jacob (2002) finds that higher noncognitive skills and college premiums among women account for most of the gender gap in higher education enrollment in 1988. Those papers focus only on one cohort and thus cannot examine the trends.

Among works that study the reversal of the gender gap in higher education enrollment over time, Anderson (2002) suggests that increasing discount rates over time have a role in explaining the gender gap in college enrollment. Charles and Luoh (2003) emphasize the effect of the uncertainty of future wages on relative schooling by gender. Those papers do not consider the effects of marriage and children on college entry decisions. Chiappori, Iyigun, and Weiss (2006) show, in a theoretical framework, that women can acquire more schooling than men if the gender wage gap narrows with the level of education. One crucial assumption of their model is that the intramarital share of the marriage surplus one can extract increases with his or her education. Our results do not rely on this assumption. Goldin, Katz, and Kuziemko (2006) show that improvement in test score and high school performance, driven by the increase of expected labor market return to education, can explain most of the relative increases in women's college completion rate. They do not quantify different returns to education. To our knowledge, our paper is the first that incorporates several factors in a structural model to quantitatively account for the reversal of the gender gap in college attainment.[1]

The paper is organized as follows. In Section 2, we present some empirical results from the PSID documenting college attainment rates in 1980–1996. In Section 3, we present our model. Section 4 provides parameters estimated from the data that are used in the model. Section 5 presents the quantitative results of the benchmark model and investigates the quantitative importance of changes in relative earnings, changes in parental education, and changes in the marriage market. Section 6 conducts sensitivity

[1] Sanchez-Marcos (forthcoming) quantifies the reduction in gender gap in college attainment in a structure model. She does not study the overtaking of female college attainment afterwards.

analysis. Brief concluding remarks are provided in Section 7.

2 Data on College Attainment

We use the PSID to calculate college attainment rates. The PSID is a longitudinal survey of U.S. families and the individuals who make up those families. We select individuals in the core sample whose ages were between 25–34 in that year and who had valid information on parents' education.[2] We use completed schooling among mature adults as the measure of an individual's schooling.[3] An individual who has more than 12 years of education completed by age 34 is defined as having a college education. The college attainment rate is calculated as the fraction of individuals that have college education among each specific group.

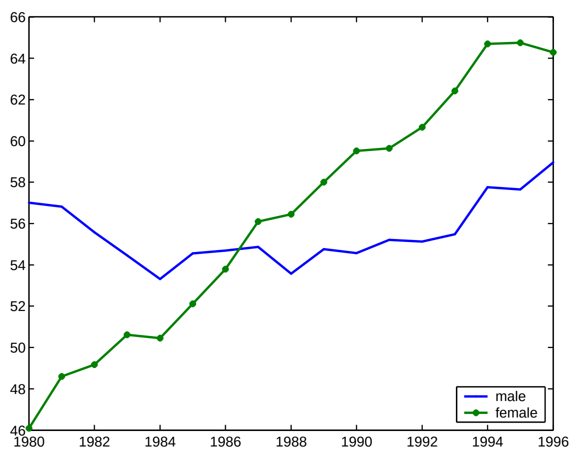

Figure 1: College attainment rates by age 34 among those aged 25–34.
Source: Authors' calculations from the PSID data files.

Figure 1 illustrates the changes in relative college attainment by males and females over the sample period considered here, 1980 to 1996.[4] Among those whose ages were

[2]We thus use the average college attainment of 10 birth cohorts. The sample size in the PSID is too small for us to analyze each birth cohort.

[3]See Charles and Luoh (2003) for a discussion of the advantage of using school attainment among mature adults over enrollment.

[4]We choose this beginning period to avoid the high male-to-female ratio in the early 70s after the Vietnam War. We choose this ending period because of the availability of data. The latest year of data

between 25 and 34 in 1980, 57 percent of young men had some college education, which was 11 percentage points higher than young women. By 1996, male college attainment rate had increased slightly by 2 percentage points, while female college attainment had increased by 18 percentage points. In fact females have led males in college attainment since 1987.[5, 6]

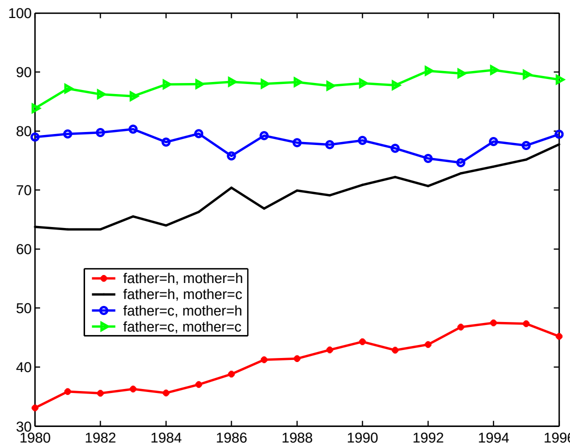

Figure 2: Female college attainment rates conditional on parental education.
Source: Authors' calculations from the PSID data files. h denotes high school and below, and c denotes some college and above.

We also calculate college attainment rates conditional on parents' education. A detailed description of the data processing procedure is provided in Appendix 8.1. Figure 2 shows female college attainment conditional on parental education. We observe that a college-educated parent is substantially more likely to have a college-educated daughter than a noncollege-educated parent, even after controlling for the education of the other parent.[7] For example, among those who were in the age range of 25 to 34 in 1980, 84 percent of females whose parents both had a college education had attended some college,

available to us is 2005 PSID. Since we use education completed by age 34, individuals at the age of 34 in 2005 were 25 in 1996. For the years 1997 and later, of course, education by age 34 is not available for individuals at age 25.

[5]Other studies (see, for example, Charles and Luoh 2003; Goldin, Katz, and Kuziemko 2006), which use different measures of education or different data sets, find similar patterns.

[6]The sample size in PSID is too small if we divide the sample by race/ethnicity. The process of convergence and ultimate ascendancy by women in completed schooling among successive generations of men and women is evident, however, when we divide sample by race/ethnicity using the CPS.

[7]Similar patterns hold for sons, and the results are available from the authors upon request.

which was 5 percentage points higher than those whose father had a college education but whose mother did not, and 20 percentage points higher than those whose mother had a college education but whose father did not. Therefore, the marginal effect of a father's education on his children's education is larger than that of a mother's. We also observe that the conditional attainment rates increase at a much slower pace than the aggregate attainment rates. This indicates that a large fraction of the observed increase in female attainment can be accounted for by the increase in their parents' attainment.

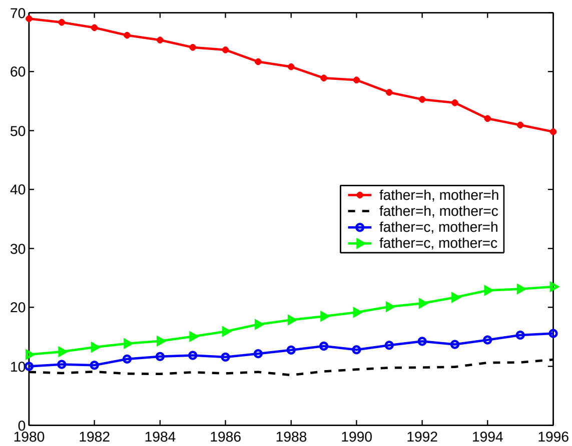

Figure 3: Parents' education distribution.
Source: Authors' calculations from the PSID data files.

The schooling distribution of our PSID sample's parents is shown in Figure 3. We observe that the number of college-educated parents increases over time. In 1980, 12 percent of individuals ages 25–34 had parents that both had college educations, and 69 percent had parents that both had only high school educations or below. By 1996, the fractions have changed to 23 percent and 50 percent.

3 The Model

The economy is populated by individuals that live for two periods. We assume that going to college entails an idiosyncratic nonpecuniary effort cost $D \in [0, \infty)$.[8] In each period, the adult population is characterized by a distribution of effort costs. At the beginning of the first period, individuals with different costs make schooling decisions. In the second period, they might get married and have children. Parents are altruistic and care about their children's learning ability. We assume that the higher a parent's education, the higher is his or her children's learning ability. Factors that affect an individual's decision on whether to attend college include the direct labor market returns to college and the marriage market returns to college, the impact of one's own schooling on his or her children's learning ability, as well as the effort cost. These costs and benefits can differ by gender. We now describe the model in more detail.

3.1 Marriage and the Labor Market (Second Period)

In the second period, individuals of schooling type s_f and s_m might marry at an exogenously given rate, and they work.[9] We denote the education of a male as $s_m = \{0, 1, 2\}$ and the education of a female as $s_f = \{0, 1, 2\}$, where 0 stands for that person being absent, 1 stands for being a high school graduate, and 2 stands for having a college education. Let Y_{g,s_m,s_f} denote the earnings, net of monetary costs of attending college, of an individual of gender $g = \{f, m\}$, male's education $= s_m$, and female's education $= s_f$. For example, $Y_{f,0,2}$ denotes the earnings of a single female with a college education,

[8] We can interpret the effort cost as net of the psychic benefit of attending college. Heterogeneity in effort cost in our model is equivalent to heterogeneity in the consumption value of school in the literature (Keane and Wolpin 1997, 2001; Eckstein and Wolpin 1999; and Ge 2008). These papers consider the lifecycle decisions of one cohort. They normally allow individual heterogeneity in other dimensions, for example, different wage offers. However, Ge (2008) shows that heterogeneity in the consumption value of school is the most important determinant of women's college enrollment decision.

[9] For simplicity's sake, we do not model marriage as a match outcome. Fernandez, Guner, and Knowles (2005) study the interactions between household matching, inequality, and per capita income.

and $Y_{f,2,2}$ denotes the earnings of a female with a college education whose spouse has a college education.

Each individual values his or her own consumption and children's learning ability, if that person has children. We abstract from out-of-wedlock birth and assume that a single individual does not have children.[10] [11] If a person does not marry, that person enjoys the full benefits of his or her own earnings. The person also gains or loses utility from the status of being single. The lifetime utility functions for a single female and a single male of schooling level s are, respectively,

$$U^f(s_m = 0, s_f = s) = \log(Y_{f,0,s}) + \delta, \quad (1)$$

$$U^m(s_m = s, s_f = 0) = \log(Y_{m,s,0}) + \delta, \quad (2)$$

where δ is the additional value from being single, which can be negative or positive.

Let's take the case of married couple with children. We assume that fertility is exogenous and the total number of children a couple has is independent of each spouse's education.[12] The cost of having children as the opportunity cost of time will be incorporated into our estimates of the earnings process. We abstract from the financial cost of raising children.

The learning ability of the couple's children, a', is a function of the couple's human capital, s_m and s_f. The production function of children's learning ability is Cobb-Douglas

$$a'_{s_m,s_f} = s_m^{1-\theta_s} s_f^{\theta_s}. \quad (3)$$

[10]Allowing single individuals to have children would significantly complicate our analysis, as it is hard to measure directly the effect of the absent parent on the children's education outcome.

[11]We abstract from divorce since, as is shown in Stevenson and Wolfers (2007), divorce rate is quite stable since 1980s.

[12]Fernandez and Rogerson (1998), Rios-Rull and Sanchez-Marcos (2002), and Greenwood, Guner, and Knowles (2003) show that fertility declines with income and education. Adopting the assumption that fertility declines with education should only change the results marginally since in our model an individual's decision is not affected by the number of children.

This functional form captures the fact that when parents are more educated, their children tend to have high learning ability. This could occur because more educated parents provide a better environment for children to flourish, or because parental learning ability is passed on genetically (Plug and Vijverberg 2003). Children of different genders from the same family have the same learning ability.

Each spouse gets a share of the total family income, with the weight of each spouse depending on his or her individual relative earnings through a parameter $\lambda \in [0,1]$. The income share of the wife is $(1-\lambda)0.5 + \lambda Y_{f,s_m,s_f}/(Y_{m,s_m,s_f} + Y_{f,s_m,s_f})$, and the share of the husband is $(1-\lambda)0.5 + \lambda Y_{m,s_m,s_f}/(\ Y_{m,s_m,s_f} + Y_{f,s_m,s_f})$. Notice that $\lambda = 0$ is the case of full income-pooling, while $\lambda \in (0,1]$ implies that each spouse's share of household earnings increases with his or her own earnings.

The utilities of men and women of marriage type (s_m, s_f) are given, respectively, as

$$U^f(s_m, s_f) = \log[0.5((1+\lambda)Y_{f,s_m,s_f} + (1-\lambda)Y_{m,s_m,s_f})] + \lambda_a \log[a'_{s_m,s_f}], \tag{4}$$

$$U^m(s_m, s_f) = \log[0.5((1+\lambda)Y_{m,s_m,s_f} + (1-\lambda)Y_{f,s_m,s_f})] + \lambda_a \log[a'_{s_m,s_f}], \tag{5}$$

where λ_a measures the weight on the utility from children's learning ability.[13]

3.2 The College Decision (First Period)

At the beginning of the first period, an individual decides whether to go to college or not. A female individual chooses whether to attend college, $s_f = 1$ (high school) and $s_f = 2$ (college), given her conditional marriage probabilities $P_f(s_m|s_f)$ and her

[13] An alternative specification of altruism is to use a dynastic model. This would complicate our analysis significantly because the environment is not stationary. Thus we use children's learning ability to approximate their expected lifetime utility. Learning ability is an important component of an individual's endowments when the college entry decision is made. As is shown in Keane and Wolpin (1997), variance in expected lifetime utility between endowment types could account for 90 percent of the total variance.

individual cost of schooling D, by solving

$$(6)\ \ \max_{s_f}\Big\{\sum_{s_m=0}^{2} U^f(s_m, s_f=1)P_f\left(s_m|s_f=1\right),\sum_{s_m=0}^{2} U^f(s_m, s_f=2)P_f\left(s_m|s_f=2\right)-D\Big\}.$$

Note that $P_f\left(0|s_f\right)$ is the probability of being single. A male's problem is defined analogously.

An individual is indifferent as to whether he or she goes to college or not if the expected utility gain from going to college is equal to the effort cost D. We define the threshold levels as

$$(7)\ D_f^* \equiv \sum_{s_m=0}^{2} U^f(s_m, s_f=2)P_f\left(s_m|s_f=2\right)-\sum_{s_m=0}^{2} U^f(s_m, s_f=1)P_f\left(s_m|s_f=1\right),$$

$$(8) D_m^* \equiv \sum_{s_f=0}^{2} U^m(s_m=2, s_f)P_m\left(s_f|s_m=2\right)-\sum_{s_f=0}^{2} U^m(s_m=1, s_f)P_m\left(s_f|s_m=1\right).$$

Therefore a female with an idiosyncratic effort cost D chooses to go to college, $s_f=2$, if and only if $D<D_f^*$, and a male chooses $s_m=2$ if and only if $D<D_m^*$.

3.3 Distribution

Each individual receives a draw of effort cost, D, in the first period. We assume that the individual's learning ability, a, affects the distribution of effort cost from which he or she draws. More specifically, we assume that the effort cost D is log-normally distributed with mean $\mu(a)$ and variance σ^2, where $\mu(a)$ is decreasing in the learning ability level a. Recall from Equation (3) that a is determined by parent's type, $a_{s_{m_{-1}},s_{f_{-1}}}$, where $s_{j_{-1}}$ is parent j's schooling. In each period, there are 4 different values of a. Let $\psi_g^c(s_{m_{-1}}=i, s_{f_{-1}}=j)$ denote the college attainment rates of individuals of gender g, conditional on parents' education, which are calculated using cumulative distribution function of D at D_g^* as follows:

$$\psi_g^c(s_{m_{-1}} = i, s_{f_{-1}} = j) = F[D_g^* | a_{i,j}]. \tag{9}$$

Notice that the fraction of individuals that go to college will depend on the parents' type, because the parents' type determines the average effort cost these individuals bear.

Let the total fraction of individuals of gender g attending college be Φ_g^c. Denote $p_{-1}\left(s_{m_{-1}} = i, s_{f_{-1}} = j\right)$ as the fraction of fathers and mothers with education level i and j, respectively. Thus the aggregate college attainment, Φ_g^c, is the average of the conditional attainment rates weighted by parents' education distribution:

$$\Phi_g^c = \sum_{i,j=1}^{2} \psi_g^c(s_{m_{-1}} = i, s_{f_{-1}} = j) * p_{-1}\left(s_{m_{-1}} = i, s_{f_{-1}} = j\right). \tag{10}$$

4 Data Inputs

We calculate parents' education distributions, marriage distributions, and lifetime earnings as inputs of the model. We compute the distribution of parents' education from the PSID. The results were presented in Section 2. Since the Current Population Surveys (CPS) cover longer periods and have a larger sample than the PSID, we use the CPS to estimate earnings and marriage distributions.[14] This section describes the estimation procedure and results of those parameters in detail.

4.1 Marriage Distributions

We estimate the probability that each individual will be married from the March supplement of the CPS 1964–2007. We define an individual as having a college education if he or she completes more than 12 years of schooling, and define an individual as

[14]The PSID and the CPS show similar patterns of college attainment.

having a high school education if he or she completes 12 years of schooling or below.[15] To be consistent with our model, we exclude individuals whose marital status is that of widowed, divorced, or separated. Individuals in our sample are either never married or currently married. We define an individual as single if he or she has never married.

It is undoubtable that most individuals will eventually get married. However, a person married at a later age stays in single status longer and thus enjoys less the benefit of marriage. Therefore, the relevant measurement of marriage probability is the expected lifetime marriage probability for an individual at the beginning of the life cycle. Unfortunately, this is hard to measure in CPS. We thus use the probability of being married at a middle age as a proxy of the average lifetime marriage probability.[16] For each birth cohort, we first construct a pseudo-panel of people between the ages of 18 and 65. In each pseudo-panel we construct, we calculate the life-cycle profile of marriage probability, which is defined as the fraction of individuals that are married at each age. Usually not the entire life-cycle profile is observed.[17] We use a fourth-order polynomial in age to estimate the life-cycle profile. Then we pick the predicted probability of being married at age 35 (38) as a proxy of the average lifetime marriage probability for a typical female (male).[18] For each year between 1980 and 1996, we compute the average marriage probability of these cohorts who were between the ages of 25 and 34 in that year.

Figure 4 shows marriage probability by gender and education. For a male, having attained the level of college implies a higher probability of being married than if one

[15]The CPS changed schooling classification in 1992. Prior to 1991, we have information on the number of grades attended and completed up to 18 years. After 1992, information on an individual's highest degree received was provided. We classify those who have some college but no degree, and those who have bachelor's degrees and above as being college educated.

[16]To avoid the potential bias caused by the fact that more-educated persons marry later, we use marriage probabilities at a middle age instead of an earlier age.

[17]For example, for a cohort born in 1970, the available CPS data only provide us with a marriage probability profile from ages 18 to 37.

[18]By using a later age for males than for females, we thus capture the pattern that males usually marry later than females.

had not. For a female, having attended college implies a lower probability of marriage before 1990 and a higher probability after 1990. Figure 4 also shows that there has been a substantial decline in marriage probabilities for both genders. From 1980 to 1996 the marriage probability decreased by 18 percentage points for high school males, 13 percentage points for college males, 15 percentage points for high school females, and 10 percentage points for college females.[19]

We assume the decrease in marriage probabilities is exogenous.[20] In this paper we focus on how future expected marriage probability affects education decisions, when individuals take into account that going to college will change their future perspective on marriage. We realize that more-educated persons marry later; thus we choose to measure marriage probability at late ages. We choose not to postpone the age for marriage probability proxy further since doing so forces us to rely more on extrapolation of those probabilities for many later cohorts.[21] We conduct sensitivity analysis in Section 6 using marriage probabilities that decline less dramatically.

We then calculate, conditional on being married, the probability of marrying each type of spouse. We use household and spousal identification information to match couples. For the years from 1980 to 1996, we compute the marriage distribution of those married individuals who were between the ages of 25 and 34 in each year.

Figure 5 confirms the well-known phenomenon that people do not marry randomly

[19]The marriage market is clear at each point in time by construction. The marriage probability for males is lower than that for females for each cohort mainly because males marry three years later than females. At each point in time, males marry females three cohorts younger. Shifting males' profiles by three years to the right reduces most of the differences in marriage probabilities by gender.

[20]Stevenson and Wolfers (2007) review the potential reasons to explain the changes in marriage probabilities. Greater access to birth control and abortion might reduce marriage (Akerlof, Yellen, and Katz 1996; Goldin and Katz 2002). Labor-saving technology might decrease the return to be gotten from specialization within a household. Increasing wage inequality might increase the time needed to search within the marriage market (Gould and Paserman 2003).

[21]Nevertheless, our estimated marriage probabilities for earlier cohorts are good approximations of lifetime marriage probabilities. For example, Stevenson and Wolfers (2007) show that among those who were born in 1950-1955, 10.5% were never married by age 45. Our estimates of 10-15% are only slightly higher.

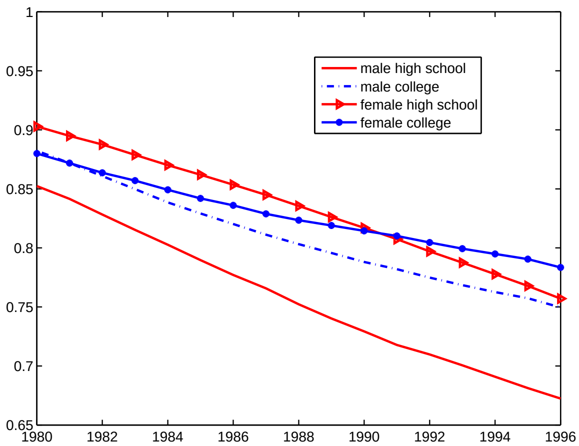

Figure 4: Marriage probabilities at age 35 (38) for female (male) by education

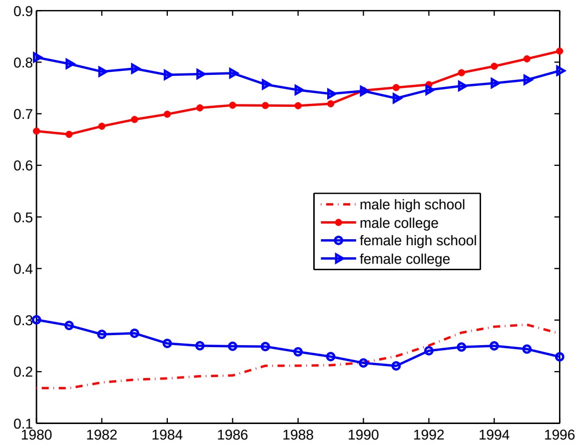

Figure 5: Probability of marrying a college-educated spouse conditional on own gender and education

and there exists assortative matching (Becker 1973; Mare 1991; and Pencavel 1998).[22] A college-educated person is more likely to marry a college-educated spouse and benefit from the spouse's earnings. Meanwhile, Figure 5 shows that as the female college attainment rate increases, the probability of marrying a college female for a male, with or without college education, increases over time.

4.2 Earnings

We need to estimate expected lifetime earnings at each marriage status for an individual at the beginning of the life cycle. We do not observe wages for those who do not work, since there are none. If labor force participation is correlated with unobservable determinants of wages, a simple OLS regression is biased. To control for the selection bias, we estimate the wage by a two-stage procedure: First we estimate equations of observed labor market participation as functions of explanatory variables along with random disturbance terms representing unobservable factors. Then we specify and estimate equations of the logarithm of wage, controlling for participation selection.

[22] Benham (1974), Boulier and Rosenzweig (1984), Behrman, Rosenzweig, and Taubman (1994), and Weiss (1997) point out that one's own schooling can improve spousal schooling acquired in the marriage market, but it is difficult to conclude whether this effect is due to human capital accumulation within the household or assortative mating.

4.2.1 Estimation Procedure

We estimate a regression function for each subsample of working individuals by gender as

$$\log w_i \quad = \quad X_i\beta + \alpha V_i + \eta_i, \tag{11}$$

where $\log w_i$ is the logarithm of real hourly wage and X is a vector of characteristics such as schooling, work experience, etc. The variable V, the inverse Mills ratio, represents the selection effect of participation.

We apply the Heckman (1979) and Lee (1978) two-stage estimation methods to this model to obtain consistent estimates. First, we estimate equations linking observed labor market participation to a set of explanatory variables and a random disturbance terms representing unobservable factors.[23] Second, we use these estimates to construct the inverse Mills ratios for the wage equation. Then, we run an OLS regression of log wage equations on X, using the estimated inverse Mills ratios as additional regressors, as is specified in Equation (11). Finally, we predict hourly wage for each individual using the fitted equation:

$$\log \widehat{w}_i \quad = \quad X_i\widehat{\beta}, \tag{12}$$

where $\widehat{\beta}$ is the consistent estimation of β. The estimation results, along with a full description of our methodology, are provided in Appendix 8.2.

[23]The standard procedure for ensuring identification is to have this set of variables not be identical to X. In our specification, the number of children is assumed to affect the participation decision, but not wages directly.

4.2.2 Lifetime Earnings

The earnings concept that is consistent with our model is the expected lifetime earnings. To calculate the lifetime earnings, first we use the following procedure to estimate the average life-cycle profiles of earnings from the CPS. For a typical individual who is in birth cohort t, with gender $g = \{f, m\}$, education of the husband $s_m = \{0, 1, 2\}$, and education of the wife $s_f = \{0, 1, 2\}$, we denote earnings at age age as $y^t_{g,s_m,s_f}(age)$. For each birth cohort t, we first construct a pseudo-panel between ages 18 and 65. Then $y^t_{g,s_m,s_f}(age)$ is calculated as the product of mean predicted hourly wages (as in Equation [12]) and mean annual hours worked by that particular gender, education, marital status, age, and cohort. We then use a polynomial in age, age, to estimate the life-cycle earnings profile for each birth cohort by gender and education of both spouses:

(13)

$$y^t_{g,s_m,s_f}(age) = \beta_{0,g,s_m,s_f} + \beta_{1,g,s_m,s_f} \cdot age + \beta_{2,g,s_m,s_f} \cdot age^2 + \beta_{3,g,s_m,s_f} I(cohort = t) + \varepsilon^t_{g,s_m,s_f}(age),$$

where $I(cohort = t)$ is a dummy for birth cohort t.

We calculate total lifetime earnings by gender, marital status, and cohort using the estimated life-cycle earnings profiles described in Equation (13). For a male who is in birth cohort t, with education of the husband $s_m = \{0, 1, 2\}$ and education of the wife $s_f = \{0, 1, 2\}$, we calculate total discounted life-cycle earnings at the beginning of his adult life, Y^t_{m,s_m,s_f}, as

$$Y^t_{m,s_m,s_f} = \sum_{age=18}^{65} (\frac{1}{1+r})^{age-18} y^t_{m,s_m,s_f}(age), \text{ if } s_m = h;$$

$$Y^t_{m,s_m,s_f} = \sum_{age=22}^{65} (\frac{1}{1+r})^{age-18} y^t_{m,s_m,s_f}(age) - \sum_{age=18}^{21} (\frac{1}{1+r})^{age-18} cs^t(age), \text{ if } s_m = c,$$

where r is the annual real interest rate, $cs^t(age)$ is the annual cost of college for cohort t at age age, which, including tuition, room and board, and $y^t_{m,s_m,s_f}(age)$, is the annual real earnings at age $age = \{18, 19, ..., 65\}$ as given by Equation (13).[24] A female's lifetime earnings are calculated analogously. An interest rate of $r = 4\%$ is used. We set the annual cost of college based on the estimates from the National Center for Education Statistics (NCES, Digest of Education Statistics, 2004, Table 313).

For the years from 1980 to 1996, we compute the average discounted life-cycle earnings of those cohorts aged 25–34 in each year, by gender and marriage type. Our measurement of lifetime earnings thus incorporates both the changes in labor supply and changes in the wage. We do not disentangle those two forces in the data because in our model, those two forces affect education decision through the same channel by changing earnings.

Figure 6 shows lifetime earnings for males at each marriage status. We observe that high school graduates on average have much less earnings than college graduates, regardless of marriage status. We also notice that married males on average earn more than never-married ones do.[25] The marriage premium is the highest for those whose spouses have college degrees. Over time, lifetime earnings by cohort are decreasing slowly for males at all marital statuses, especially for married males.[26]

Figure 7 shows lifetime earnings for females at each marriage status. We observe that high school graduates on average have much less earnings than college graduates, regardless of marriage status. However, we do not find that married females–unlike males–earn more than singles. The marriage premium for females is negligible among high school graduates and is in fact negative among college graduates.[27] Over time, lifetime earnings for married females are increasing gradually, partially because of the

[24]We assume college students cannot work, thus do not have earnings between the ages of 18 and 21.

[25]Korenman and Neumark (1991), among others, attribute most of the male marriage wage premium to productivity increased in marriage.

[26]Our finding is consistent with Kambourov and Manovskii (2005), who show that the life-cycle profiles of males' earnings for younger cohorts are lower than those for older cohorts.

[27]Papers measuring marriage premium using wages generally find negligible premium for females. We find negative marriage premium among female college graduates because we use lifetime income and single college females work more than the married counterpart.

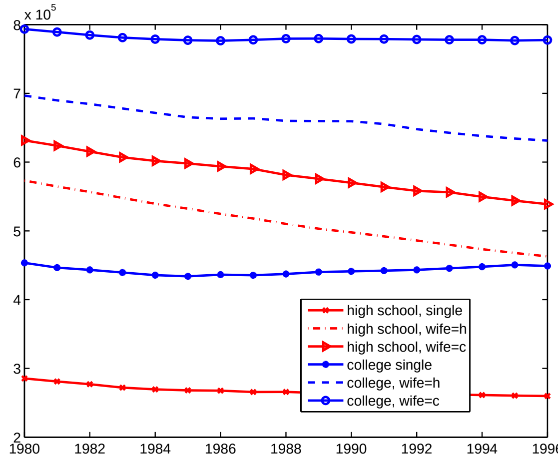

Figure 6: Male's lifetime earnings
Source: Authors' estimation from the CPS data files.

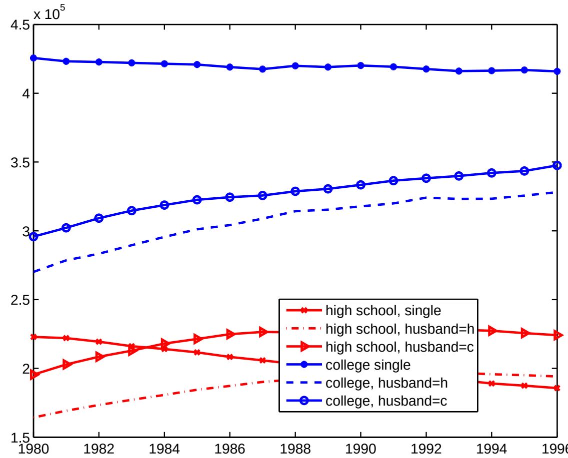

Figure 7: Female's lifetime earnings
Source: Authors' estimation from the CPS data files.

increasing female labor supply, partially because of the increasing of wages.[28] On the contrary, lifetime earnings for single females are decreasing gradually, since the increase of wages is offset by the decrease of labor supply.[29] Compared with lifetime earnings for males shown in Figure 6, females earn much less than males of the same marital status. The differences in earnings between married males and married females are decreasing gradually over time.

5 Findings

Can the model replicate the change in college attainment that occurred between 1980 and 1996? To do this, we use the data reported in Section 4 and estimate the other model's parameters which are constant over time by matching aggregate and conditional

[28]Existing theories that explain the increase in female labor participation include the following: technological innovation (Greenwood, Seshadri, and Yorukoglu 2005; Goldin and Katz 2002; and Albanesi and Olivetti 2006), falling child care costs (Attanasio, Low, and Sanchez-Marcos, forthcoming), an increase in the number of jobs that are less physically demanding (Goldin 1990), cultural acceptance of maternal employment (Fernandez, Fogli, and Olivetti 2004; Fernandez 2007; Fogli and Veldkamp 2008), and increases in women's wages (Jones, Manuelli, and McGrattan 2003). Existing theories that explain the decrease of the gender wage gap include gender differences in qualifications and discrimination (Blau and Kahn 2000), and self-selection (Mulligan and Rubinstein 2007).

[29]Our results confirm McGrattan and Rogerson (2004), who use census figures and find a decline of hours worked by single females.

college attainment rates obtained in the data. Then we run counterfactual simulations to study the effects of different mechanisms on college attainment by comparing college attainments from each simulation with those in the benchmark.

5.1 Benchmark

We use calculated earnings, marriage distributions, and parent education distributions as inputs of the model and estimate the rest of the parameters to match observed aggregate and conditional attainment rates by gender. The estimated parameters are presented in Table 1. We find that $\lambda = 0$, indicating that each spouse consumes half of the household earnings. The preference parameter δ is positive, indicating that the status of being single brings additional utility. The ability production parameter θ_s is less than 0.5, indicating that fathers' education affects children's learning ability more than mothers' education, consistent with Figure 2, where the marginal effect of fathers' education on children's education is larger than that of mothers'.[30] The fact that θ_s is less than 0.5 implies the following order of learning ability levels by parents' education: $a_{1,1} < a_{1,2} < a_{2,1} < a_{2,2}$. This in turn implies that the effort cost distribution parameters by parents' education $\mu_{i,j} = \mu(a_{i,j})$ have a corresponding rank order. In particular, since $\mu(a)$ is decreasing in a, we have $\mu_{1,1} > \mu_{1,2} > \mu_{2,1} > \mu_{2,2}$.

Parameters	Estimates
Preference	$\lambda_a = 2.023, \lambda = 0, \delta = 1.005$
Ability production	$\theta_s = 0.425$
Effort cost distributions	
Means	$\mu_{1,1} = 0.257$, $\mu_{2,1} = 0.077$
	$\mu_{1,2} = 0.137$, $\mu_{2,2} = -0.006$
Standard Deviation	$\sigma = 0.191$

Table 1: Parameters used in the benchmark model

Figure 8 compares college attainment rates from the model with those in the data.

[30]This is also consistent with the empirical results of Behrman and Rosenzweig (2002).

The model is able to generate the pattern that college attainments for females were lower in 1980 and higher in 1996 than those for males, as is observed in the data. In the model, female college attainment began to exceed that of males in 1986, one year earlier than observed in the data.

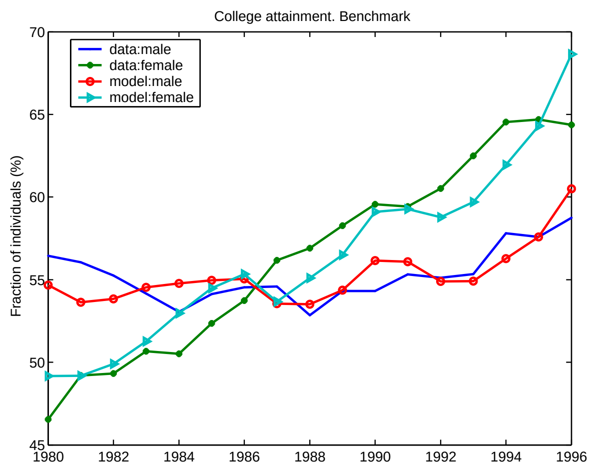

Figure 8: College attainment rates

Figure 9 compares females' college attainment rates conditional on parent's education from the model with those in the data. The model is able to generate the pattern that a college-educated parent is substantially more likely to have a college-educated daughter than a parent who is a noncollege graduate, even after controlling for the education of the other parent.[31] In our model, parents' type determines the average effort cost these individuals bear. Thus the order of μ's, $\mu_{1,1} > \mu_{1,2} > \mu_{2,1} > \mu_{2,2}$, is critical to generate the order of school attainment by parent's type.

5.2 Counterfactual Simulations

In the benchmark economy, changes in college attainment over time are caused by the exogenous changes in parental education, lifetime earnings by education, and marriage distributions. To study the effects of different mechanisms on college attainment, we run

[31]Similar patterns hold for males, and the results are available from the authors upon request.

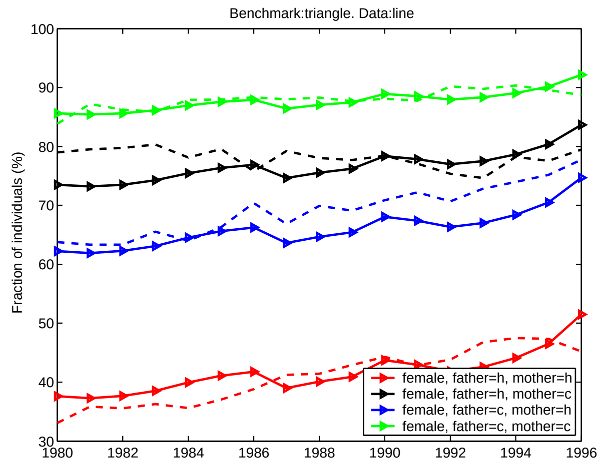

Figure 9: Female's college attainment rates (dashed line: data; solid line: model)

counterfactual simulations. For each simulation, we keep the values of the variables that we want to focus on fixed in the 1980 level, and we keep the values of other variables the same as in the benchmark model. Therefore the comparison between each simulation and the benchmark model results will quantify the direct effects of those variables.

5.2.1 Parents' background

First, we investigate the intergenerational schooling effects. The results are shown in Figure 10. When the parents' schooling distribution is fixed at the 1980 level, college attainment drops by 7.8 and 7.0 percentage points in 1996 for males and females. We notice the gender reversal of college attainment occurs in the same year as in the benchmark. Therefore, parental education is an important source of the increase in college attainment but cannot in itself account for the reversal of the gender gap.

The benchmark model captures the intergenerational persistence in schooling: When parents are more educated, their children tend to have high learning ability and are more likely to go to college. Thus the gradual transformation of parental schooling composition, as is shown in Figure 3, acts as a mechanism to propagate change in college attainment: as the number of college-educated parents increases, so does the proportion of children with high learning ability (a low value of the effort cost D), which

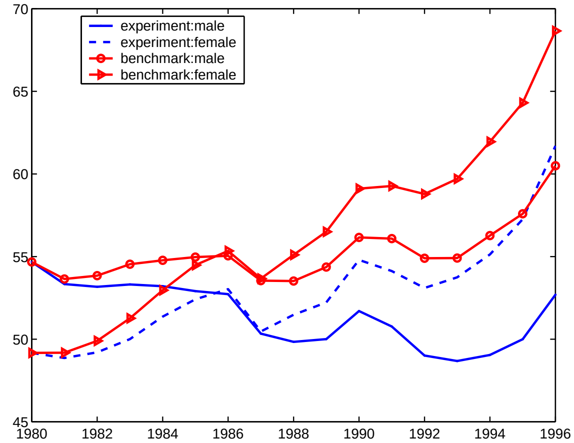

Figure 10: No change in parents' distribution since 1980

then helps to increase the attainment rate of the children's generation. This propagation mechanism seems to affect females and males in similar magnitude, so that it had little effect on the timing of gender reversal of college attainment.

These results for intergenerational schooling effects are broadly consistent with previous research. Many studies report a significant positive relationship between parents' education and the schooling of their children for one cohort (Behrman 1997; Behrman and Rosenzweig 2002). Based on data from the National Longitudinal Survey of Youth 1979 (NLSY79), Ge (2008) estimates a sequential college choice model and shows that improvements in parental education can account for a large part of the college attendance difference between NLSY79 young women and those born almost 20 years later. To our knowledge, our paper is the first attempt to investigate the importance of intergenerational schooling effects in accounting for the trends of college attainment for both genders.

5.2.2 Earnings

To understand the effect of earnings on education, we calculate the labor market return to education. For singles, we compute the ratios of life-cycle earnings between college and high school for males and females. For married couples, the relevant concept of earnings is household lifetime earnings, since $\lambda = 0$ implies full income pooling between spouses. For a married female (male), we compare the earnings of a household in which the wife (husband) has a college education, but the husband (wife) does not, with the earnings of a household in which both spouses are high school graduates.[32]

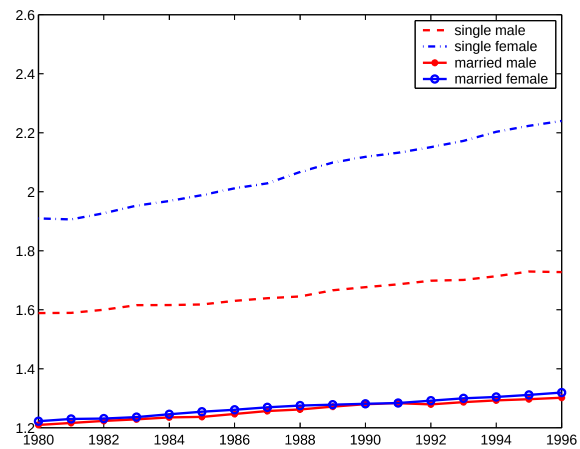

Figure 11: Earnings Return to College by Gender and Marital Status

Figure 11 presents the earnings return to college by gender and marital status. Several patterns are observed over the period 1980–1996. First, the earnings return to college increases for both genders and for all marital statuses.[33] Second, the earnings return to college is higher for single females than for single males. In 1980, single college-educated males had 59 percent more earnings than single high-school-educated males, while single college-educated females had 91 percent more earnings than single high-school-educated females. The earnings return to college for single females has increased more than that for single males between 1980 and 1996. Third, the earnings return to college is similar for married females as for married males. Compared to a typical household in 1980 in which neither spouse attained college, a household in which the man went to college had

[32]We also compared earnings in households where both spouses are college graduates with earnings in households in which the wife (husband) has a college education but the husband (wife) does not; the returns are only slightly higher and the overtime trends are almost identical.

[33]Using cross-sectional earnings or wages, many authors have documented recent increases in the earnings return to college. (See, for example, Juhn, Murphy, and Pierce 1993; Katz and Murphy 1992; Card and DiNardo 2002; and Eckstein and Nagypál 2004) Our measure using lifetime earnings gives similar results.

21 percent more earnings, while a household in which the woman went college had 22 percent more earnings.

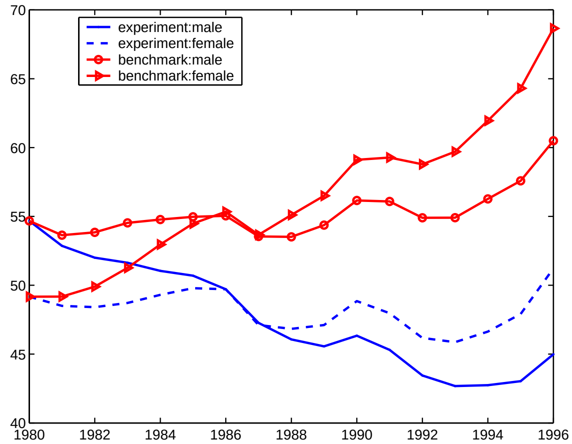

Figure 12: No change in earnings since 1980

We now analyze the case in which there is no change in earnings since 1980. The results are shown in Figure 12. Male and female attainment rates drop by 15.5 and 17.4 percentage points by 1996. This indicates that the increasing returns to college in the labor market for those cohorts, as shown in Figure 11, have an important impact on college attainment for those cohorts.

The change of earnings has a larger effect on college attainment for females than that for males. This is due to the fact that, over time, the earnings return to college for single females has been increasing at a faster rate than that for single males. The gender reversal of college attainment occurs in the same year, however, as in the benchmark model. Thus the change in earnings over time cannot account for the reversal of the gender gap in college attainment.[34]

5.2.3 Marriage market

The next two simulations try to isolate the effects of changes in the marriage market on college attainments. First, we quantify the effects of declines in marriage probabilities,

[34] We also simulate a version that fixes monetary costs of attending college at 1980 level and the resulting attainments for both genders are very similar to the benchmark results.

keeping conditional marriage probabilities as in the data. Then, we show the effects of changes in conditional marriage probabilities, keeping marriage probabilities as in the data.

Marriage probabilities Figure 13 shows that, without change in the probabilities of marriage, both males and females would reach higher college attainment in 1996 and females' college attainment would always be lower than that of males.

The decline in marriage probabilities decreases college attainment for both genders. This can be explained by the differences in the total returns to education by marital status. As is shown in Figure 11, the earnings return to college in the labor market is higher for singles than for married couples. However, married individuals receive an additional benefit from college through increasing their children's learning ability. Under our parameters, the return from children for married couples dominates their lower return in the labor market; thus the returns to college increase with marriage probability. As the marriage probability declines, returns to college decrease and so does college attainment.

The comparison also indicates that, as marriage probabilities decline, female college attainment decreases less than that of males. This is because single females receive a larger return to college in the labor market than single males. Moreover, under our parametrization, θ_s is less than 0.5; therefore married females receive a smaller return to college through their children than married males. As a result, the decline in the marriage probability decreases the returns to college for females less than those for males. Therefore, college attainment for females declines less than that for males. This implies that changes in marriage probabilities are crucial in accounting for the reversal of the gender gap in college attainment. In fact, Figure 13 shows that without the decline in marriage probabilities, female college attainment never exceeds male attainment.

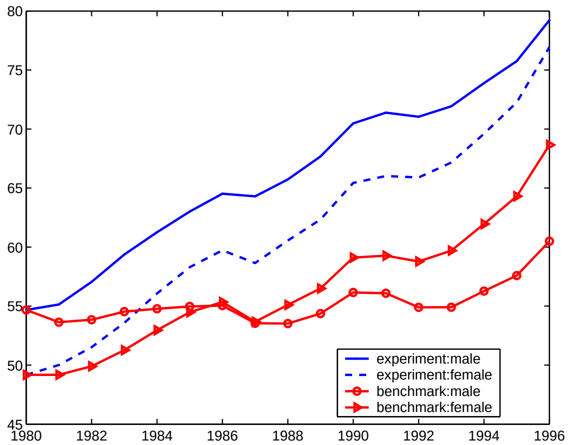

Figure 13: No change in single rates since 1980

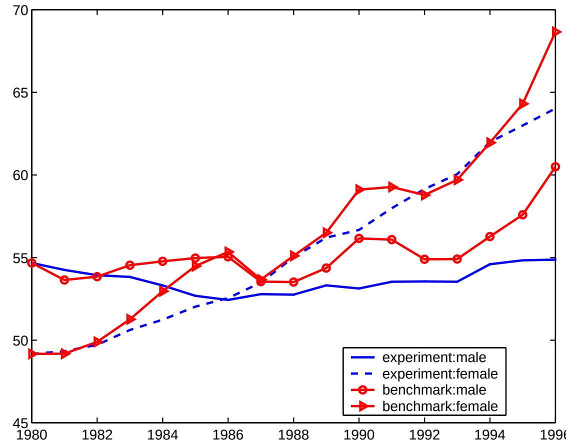

Figure 14: No change in conditional marriage probability since 1980

Conditional marriage probabilities In the fourth simulation we fix the conditional marriage probabilities at the level they were in 1980 and keep the marriage probabilities in the data. The results are shown in Figure 14. The college attainment in 1996 would be 4.6 and 3.3 percentage points lower for males and females, respectively. Therefore the change of conditional marriage probabilities plays a quantitatively minor role in accounting for the increase in college attainment for both genders.

The gender reversal of college attainment occurs at the same time as in the benchmark model. The change of the marriage probability has a larger effect on college attainment for males than for females. This is in part due to the fact that, over time, the probability of marrying a college spouse for males has increased quite substantially, while the probability for females has barely changed. In our model, spousal education increases household income and children's human capital. In the benchmark, over time, males benefit more from marrying college spouses than females do; thus college attainment for males increases more than that for females.

6 Sensitivity

6.1 Transmission of ability

In the benchmark, we assume that children's learning ability is affected only by parents' education. Now we relax this assumption by also allowing for genetic transmission of innate ability. We introduce innate ability π, and acquired learning ability a. The acquired learning ability of a couple's children, a', is a function of the children's innate ability, π', and the couple's human capital, s_m and s_f. The production function of children's learning ability is the Cobb-Douglas functional form

$$a'_{s_m,s_f}(\pi) = (\pi')^{\theta_\pi}(s_m^{1-\theta_s} s_f^{\theta_s})^{1-\theta_\pi}. \tag{14}$$

We assume that there are two types of innate ability, and that there is a positive correlation between parents' and children's innate ability.

The expected utilities of men and women at a marriage type (s_m, s_f) are given, respectively, by

$$U^f(s_m, s_f, \pi_m, \pi_f) = \log \frac{(1+\lambda)Y_{f,s_m,s_f} + (1-\lambda)Y_{m,s_m,s_f}}{2} + \lambda_a E_{\pi'|\pi_m,\pi_f} \log a', \tag{15}$$

$$U^m(s_m, s_f, \pi_m, \pi_f) = \log \frac{(1+\lambda)Y_{m,s_m,s_f} + (1-\lambda)Y_{f,s_m,s_f}}{2} + \lambda_a E_{\pi'|\pi_m,\pi_f} \log a'. \tag{16}$$

At the beginning of the first period, a female individual with innate ability π_f chooses whether to attend college, $s_f = 1$ (high school) and $s_f = 2$ (college), given her conditional marriage probabilities $P_f(s_m, \pi_m | s_f, \pi_f)$ and her individual cost of schooling D, by solving

$$\max_{s_f} \left\{ \begin{array}{l} \sum_{s_m=0}^{2} \sum_{\pi_m} U^f(s_m, s_f{=}\,1, \pi_m, \pi_f) P_f(s_m, \pi_m | s_f = 1, \pi_f), \\ \sum_{s_m=0}^{2} \sum_{\pi_m} U^f(s_m, s_f{=}\,2, \pi_m, \pi_f) P_f(s_m, \pi_m | s_f = 2, \pi_f) - D \end{array} \right\}.$$

A female individual is indifferent about whether to go to college or not if the expected utility gain from going to college is equal to the effort cost D. We define the threshold levels as

$$
\begin{aligned}
(17) \qquad D_f^*(\pi_f) = & \sum_{s_m=0}^{2} \sum_{\pi_m} U^f(s_m, s_f = 2, \pi_m, \pi_f) P_f\left(s_m, \pi_m | s_f = 2, \pi_f\right) \\
& - \sum_{s_m=0}^{2} \sum_{\pi_m} U^f(s_m, s_f = 1, \pi_m, \pi_f) P_f\left(s_m, \pi_m | s_f = 1, \pi_f\right),
\end{aligned}
$$

Therefore a female with an idiosyncratic effort cost D and innate ability π_f chooses to go to college, $s_f = 2$, if and only if $D < D_f^*(\pi_f)$. A male's problem is defined analogously.

We assume that the effort cost D is log-normally distributed with mean $\mu(a)$ and variance σ^2, where $\mu(a)$ is decreasing in the learning ability level a. In each period, there are 8 different values of a. Let $\psi_g^c(s_{m_{-1}} = i, s_{f_{-1}} = j, \pi)$ denote the college attainment rates of individuals of gender g, conditional on innate ability and parents' education, which are calculated using the cumulative distribution function of D at $D_g^*(\pi)$:

$$
(18) \qquad \psi_g^c(s_{m_{-1}} = i, s_{f_{-1}} = j, \pi) = F[D_g^*(\pi) | a_{i,j}(\pi)].
$$

We denote $p_{-1}\left(\pi | s_{m_{-1}} = i, s_{f_{-1}} = j\right)$ as the fraction of individuals that have innate ability π among those whose fathers and mothers have education levels i and j, respectively. Let $\psi_g^c(s_{m_{-1}} = i, s_{f_{-1}} = j)$ denote the college attainment rate of individuals of gender g, conditional on parents' education, which is calculated as

$$
(19) \quad \Psi_g^c(s_{m_{-1}} = i, s_{f_{-1}} = j) = \sum_{\pi} \psi_g^c(s_{m_{-1}} = i, s_{f_{-1}} = j, \pi) * p_{-1}\left(\pi | s_{m_{-1}} = i, s_{f_{-1}} = j\right).
$$

Let the total fraction of individuals of gender g attending college be Φ_g^c. Denote $\widetilde{p}_{-1}\left(s_{m_{-1}} = i, s_{f_{-1}} = j\right)$ the fraction of individuals whose fathers and mothers have education level i and j, respectively. Thus the aggregate college attainment, Φ_g^c, is the average of the conditional attainment rates weighted by parents' education distribution:

$$\Phi_g^c = \sum_{i,j=1}^{2} \Psi_g^c(s_{m_{-1}} = i, s_{f_{-1}} = j) * \widetilde{p}_{-1}\left(s_{m_{-1}} = i, s_{f_{-1}} = j\right). \tag{20}$$

We calculate $p_{-1}\left(\pi|s_{m_{-1}}, s_{f_{-1}}\right)$ using NLSY79. The NLSY79 is a nationally representative sample of young men and women who were 14–22 years old when they were first surveyed in 1979. They were 25–33 in 1990. We use AFQT (Armed Force Qualification Test) percentile score to measure innate ability. An individual is defined as having high innate ability if his or her AFQT score is above the age-adjusted median score. The calculated distribution of innate ability by parents' education seems to be stationary over birth cohorts. Thus we assume that $p_{-1}\left(\pi|s_{m_{-1}}, s_{f_{-1}}\right)$ is stationary from 1980 to 1996.

We also assume that innate ability is an unobservable individual characteristic. This justifies the assumption that $P_f(s_m, \pi_m|s_f, \pi_f)$ is independent of π_f, and $P_m(s_f, \pi_f|s_m, \pi_m)$ is independent of π_m. Thus $P_f(s_m, \pi_m|s_f, \pi_f) = P_f(s_m, \pi_m|s_f) = P_f(s_m|s_f) * P_m(\pi_m|s_m)$. We calculate $P_f(\pi_f|s_f), P_m(\pi_m|s_m)$ from NLSY79. We assume that those distributions are stationary from 1980 to 1996.

Table 2 compares the results of counterfactual simulations between the benchmark and the model with innate ability. The results indicate that in the benchmark model, which is without the intergenerational correlation of innate ability, the quantitative importance of earnings, parents' background, and marriage market on college attainment are very similar to a model with innate ability.

6.2 Marriage probability

In the benchmark model, the decline in marriage probabilities is crucial in accounting for the reversal of the gender gap in college attainment. If part of the decline in marriage

	male	female	taken-over time
1996 attainment accounted for by (in percentage points)			
1 earnings			
benchmark	15.5	17.4	1988
adding innate ability	14.4	16.2	1986
less change in marriage probabilities	15.7	17.4	1986
2 parents' distribution			
benchmark	7.81	6.96	1986
adding innate ability	7.96	7.20	1985
less change in marriage probabilities	7.86	7.1	1985
3 marriage probability			
benchmark	-18.7	-8.30	no
adding innate ability	-16.1	-7.36	no
less change in marriage probabilities	-17.6	-8.92	no
4 conditional marriage probability			
benchmark	5.62	4.63	1986
adding innate ability	3.86	3.76	1986
less change in marriage probabilities	5.41	5.02	1986

Table 2: Comparison of benchmark model, model with innate ability, and model with less change in marriage probabilities

probabilities is due to the delay in marriage, the change in marriage probabilities would have been less dramatic. To see how sensitive the results to marriage probabilities are, we reestimate the model assuming that marriage probabilities decrease by 80 percent of the size as in the benchmark every year. Table 2 shows results of decomposition. The results indicate that, using marriage probabilities that decline less dramatically, the quantitative importance of earnings, parents' background, and marriage market on college attainment is very similar to the benchmark model.

7 Conclusion

We develop a dynamic model of college entry decision that incorporates intergenerational persistence on learning ability, marriage, and differential earnings by gender and marital status. Using this model, we study the effects of changes in relative earnings, changes in parental education, and changes in the marriage market on changes in college attainment by gender. We find that the rises in parental education and relative earnings between college and high school persons increase college attainment for both genders. The declining marriage probabilities decrease college attainment for females less than that for males, and thus are crucial in explaining the reversal of the gender gap in college attainment.

There are several directions in which this work can be extended. First, we abstract from divorce and out-of-wedlock birth and assume a single individual does not have children. The divorce rate has been stabilized since the early 1980s but out-of-wedlock birth has been increasing. Those changes in family structure might affect female's and male's college attainment decisions differently. Second, we assume marriage probabilities and earnings are exogenous. An extension that we wish to explore is the relationship among college attainment, marriage, and labor supply for both genders. Even though labor earnings are sacrificed, a parent who stays at home and takes care of children would contribute to the household by increasing the learning ability of children. We plan to study these issues in future work.

8 Appendix

8.1 PSID sample

The PSID is a longitudinal survey of U.S. families and the individuals who make up those families. Approximately 4,800 U.S. families were sampled in 1968, and these families were reinterviewed annually until 1997. From 1997 onwards PSID was changed to a biennial data collection and two major changes were made: a reduction of the core sample, and the addition of a new sample of post-1968 immigrant families and their adult children.

We first find parents' education for the selected sample by linking parents and children from Individual Files (1968–2005). The PSID facilitates the intergenerational linkage by providing the parent's ID in the Individual Files. If a linkage can not be found in Individual Files, we use 2003 Parent Identification Files to link an individual with his or her parents. If the above procedure fails to provide parents' education information, we find parents' education by using parents' and parents-in-law's education as reported by the head in Family Files. In 1974, questions were asked about how much education had been completed by the household head's parents and by the spouse's parents. In the later waves, these parental education questions were asked for new heads and spouses. By merging Individual Files with Family Files, we are able to find parents' education for those who were heads or spouses or siblings of the heads.

8.2 Estimation of wage

The model is estimated on the March CPS from 1964 to 2007. We restrict the sample to individuals who are between the ages of 18 and 65 who are not in the armed forces and not self-employed. To be consistent with the decision model, we restrict our attention to individuals who are either married or single (never married). Hourly wage is deflated to 2006 dollars using the CPI. Definitions of variables are given in Appendix section 8.2.2. We run separate probit wage selection and log wage regression for each gender in each year. The reduced-form probit selection results and estimated coefficients of the wage equations in 2007 are provided in Appendix sections 8.2.3 and 8.2.4.

8.2.1 Estimation procedure of wages

Consider the following wage function on a sample of working men and women:

$$\log w_i = X_i\beta + \mu_i,$$

where $\log w_i$ is the logarithm of hourly wage, and X is a vector of characteristics such as schooling, work experience, etc. It is argued, however, that the sample of employed workers is not a random sample, and that this selectivity might bias the coefficients. Formally, we can write down a participation equation

$$\begin{aligned} E_i &= 1 \text{ if } Z_i\gamma + \varepsilon_i \geq 0, \\ E_i &= 0 \text{ if } Z_i\gamma + \varepsilon_i < 0, \end{aligned}$$

where Z includes variables that predict whether or not a person works. Therefore the probability of an individual working is

$$\Pr(E_i = 1) = \Pr(\varepsilon_i \geq -Z_i\gamma) = \Phi(\frac{Z_i\gamma}{\sigma}), \tag{21}$$

where σ_ε^2 is the variance of ε_i, and $\Phi(\cdot)$ is cumulative distribution function of the standard normal.

The selectivity problem is apparent by taking expectations of the wage function over the sample of employed workers:

$$E(\log w_i | E_i = 1, X_i) = X_i\beta + E\left(\mu_i | \varepsilon_i \geq -Z_i\gamma\right).$$

Supposing μ_i and ε_i are jointly normally distributed, let $\sigma_{\mu,\varepsilon}$ be the covariance between μ_i and ε_i. We can now write

$$E\left(\mu_i|\varepsilon_i \geq -Z_i\gamma\right) = \frac{\sigma_{\mu,\varepsilon}}{\sigma_\varepsilon}\frac{\phi(Z_i\gamma/\sigma)}{\Phi(Z_i\gamma/\sigma)},$$

where $\phi(\cdot)$ is the standard normal density. When $\sigma_{\mu,\varepsilon}$ is not zero, selectivity bias occurs. To estimate the potential wage consistently, we need to add the selection term (the inverse Mills ratio)

$$\frac{\phi(Z_i\gamma/\sigma)}{\Phi(Z_i\gamma/\sigma)} \equiv V_i \tag{22}$$

in the OLS regression as

$$\log w_i = X_i\beta + \alpha V_i + \eta_i.$$

8.2.2 Definitions of variables in X and Z

Age	Respondent's age
Age^2	Square of variable Age
HI	Dummy variable: 1 if respondent is a high school dropout
HG	Dummy variable: 1 if respondent is a high school graduate
SC	Dummy variable: 1 if respondent has some college education
CG	Dummy variable: 1 if respondent is a college graduate
Exp	Respondent's years of work experience
Exp^2	Square of variable Exp
Black	Dummy variable: 1 if respondent is black
Married	Dummy variable: 1 if respondent is married
Nchild	Number of own children in household
Nchlt5	Number of own children under age 5 in household
Northeast	Dummy variable: 1 if household is located in Northeast region
Midwest	Dummy variable: 1 if household is located in Midwest region
South	Dummy variable: 1 if household is located in South region
West	Dummy variable: 1 if household is located in West region
Metro	Dummy variable: 1 if household is located in a metropolitan
Manager	Dummy variable: 1 if respondent is a manager or professional
Whitecollar	Dummy variable: 1 if respondent has white-collar occupation other than those in management
Bluecollar	Dummy variable: 1 if respondent has blue-collar occupation
V	See Equation (22)

8.2.3 Estimation results: probit selection

The reduced-form probit selection rule in equation (21) is estimated in each year for men and women, respectively. We estimate these probits year by year because there are some evidences that how individuals select themselves into the workforce have shifted

	Males		Females	
Variable	Coefficient	t	Coefficient	t
Constant	-2.5929	-41.75	-2.6571	-43.10
HG	0.3134	15.67	0.5029	24.77
SC	0.3882	18.68	0.6520	32.05
CG	0.7044	31.09	0.8212	38.86
Age	0.1627	46.82	0.1448	41.89
Age^2	-0.0022	-52.07	-0.0018	-44.29
Black	-0.3328	-16.14	-0.0018	-0.10
Marry	0.4641	22.41	-0.0499	-2.91
Nchild	0.0396	4.82	-0.0800	-12.86
Nchlt5	0.0315	1.79	-0.2708	-22.21
No. of obs.	48,145		51,315	
-2 ln(likelihood ratio)	8285.05		5252.96	
χ^2 degree of freedom	9		9	

Table 3: Participation Selection Rules: Probit Analysis (CPS 2007)

over time (Mulligan and Rubinstein 2007). Table 3 presents estimated coefficients and asymptotic t-statistics of the reduced form participation probit for 2007.[35] Our findings are generally in accord with previous research. Specifically, we find that educational attainment has a positive and statistically significant impact on the probability of participation for both men and women. The probability of working increases in age at a decreasing rate for both men and women. Black men are less likely to participate than nonblacks. Men who are married or have children are more likely to participate than other men, even though the effect of number of children is not statistically significant. Married women and women with children are less likely to participate.

8.2.4 Estimation results: wage equations

Estimated coefficients and asymptotic t-statistics of the wage equations in 2007 corrected for selections are found in Table 4. Estimated coefficients on education, experience, occupation dummies, race, and region dummies are similar to estimates from the typical wage equations found in the literature. College education attainments are generally more important for women's wage than for men. Experience has more of a positive impact on men's wage than on women's wage.

Selectivity biases are particularly interesting. One would expect that individuals with higher wage potential should be more likely to participate in the labor force. The estimation results confirm that individuals who expect to earn more are more likely to participate in the labor force. The coefficients of V (defined in equation (22) in Appendix 8.2.1) are positive and statistically significant for both men and women. Therefore,

[35]Estimates for other years are available from the authors.

	Males		Females	
Variable	Coefficient	t	Coefficient	t
Constant	1.5958	29.79	1.4525	31.64
HG	0.3278	22.01	0.2822	13.43
SC	0.4650	28.34	0.4704	20.27
CG	0.7743	36.96	0.8072	31.61
exp	0.0462	19.04	0.0336	19.29
exp^2	-0.0009	-13.81	-0.0006	-15.43
manager	0.3618	36.55	0.4974	38.16
white-collar	0.0099	1.09	0.1966	19.19
Midwest	-0.0713	-6.83	-0.0746	-6.61
South	-0.0829	-8.39	-0.0766	-7.09
West	-0.0445	-4.38	-0.0453	-4.01
metro	0.1150	12.74	0.1476	15.27
black	-0.1424	-9.38	-0.0298	-2.41
married	0.2748	16.84	0.0425	4.01
V	0.3131	5.08	0.2283	6.51
R^2	0.3026		0.2362	

Table 4: Estimates of wage equation (11): CPS 2007

observed wage patterns of men and women are higher than the population mean pattern would have been.

References

[1] Akerlof, George, Janet Yellen, and Michael Katz. 1996. "An Analysis of Out-of-Wedlock Childbearing in the United States." *Quarterly Journal of Economics* 111(2): 277–317.

[2] Albanesi, Stefania, and Claudia Olivetti. 2006. "Home Production, Market Production and the Gender Wage Gap: Incentives and Expectations." Working Paper 2005-004. Boston: Boston University.

[3] Anderson, Patricia M. 2002. "Where the Boys No Longer Are: Recent Trends in U.S. College Enrollment Patterns." Working Paper. Hanover, N.H.: Dartmouth College.

[4] Attanasio, Orazio, Hamish Low, and Virginia Sanchez-Marcos. Forthcoming. "Explaining Changes in Female Labour Supply in a Life-Cycle Model." *American Economic Review.*

[5] Averett, Susan L., and Mark L. Burton. 1996. "College Attendance and the College Wage Premium: Differences by Gender." *Economics of Education Review* 15(1): 37–49.

[6] Becker, Gary S. 1967. *Human Capital and the Personal Distribution of Income.* Ann Arbor: University of Michigan Press.

[7] Becker, Gary S. 1973. "A Theory of Marriage: Part I." *Journal of Political Economy* 81(4): 813–846.

[8] Behrman, Jere R. 1997. "Women's Schooling and Child Education: A Survey." Working Paper. Philadelphia: University of Pennsylvania.

[9] Behrman, Jere R., and Mark R. Rosenzweig. 2002. "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?" *American Economic Review* 92(1): 323–334.

[10] Behrman, Jere R., Mark R. Rosenzweig, and Paul Taubman. 1994. "Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment." *Journal of Political Economy* 102(6): 1,131–1,174.

[11] Benham, Lee. 1974. "Benefits of Women's Education within Marriage." *Journal of Political Economy* 82(2): S57–S71.

[12] Blau, Francine D., and Lawrence M. Kahn. 2000. "Gender Differences in Pay." *Journal of Economic Perspectives* 14(4): 75–99.

[13] Boulier, Bryan, and Mark R. Rosenzweig. 1984. "Schooling, Search, and Spouse Selection: Testing Economic Theories of Marriage and Household Behavior." *Journal of Political Economy* 92(4): 712–732.

[14] Cameron, Stephen V., and James J. Heckman. 1998. "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males." *Journal of Political Economy* 106(2): 262–333.

[15] Cameron, Stephen V., and James J. Heckman. 2001. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *Journal of Political Economy* 109(3): 455–499.

[16] Card, David, and John E. DiNardo. 2002. "Skill Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles." NBER Working Paper 8769. Cambridge, M.A.: National Bureau of Economic Research.

[17] Charles, Kerwin K., and Ming-Ching Luoh. 2003. "Gender Differences in Completed Schooling." *The Review of Economics and Statistics* 85(3): 559–577.

[18] Chiappori, Pierre A., Murat Iyigun, and Yoram Weiss. 2006. "Investment in Schooling and the Marriage Market." IZA Discussion Paper 2454. Bonn, Germany: Institute for the Study of Labor.

[19] Eckstein, Zvi, and Éva Nagypál. 2004. "The Evolution of U.S. Earnings Inequality: 1961–2002." *Federal Reserve Bank of Minneapolis Quarterly Review* 28(2): 10–29.

[20] Eckstein, Zvi, and Kenneth I. Wolpin. 1999. "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities." *Econometrica* 67(6): 1295–1339.

[21] Fernández, Raquel. 2007. "Culture as Learning: The Evolution of Female Labor Force Participation over a Century." Working Paper. New York: New York University.

[22] Fernández, Raquel, Alessandra Fogli, and Claudia Olivetti. 2004. "Mothers and Sons: Preferences Formation and Female Labor Force Dynamics." *Quarterly Journal of Economics* 119(4): 1249–1299.

[23] Fernández, Raquel, Nezih Guner, and John Knowles. 2005. "Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality." *Quarterly Journal of Economics* 120(1): 273–344.

[24] Fernández, Raquel, and Richard Rogerson. 1998. "Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform." *American Economic Review* 88(4): 813–833.

[25] Fogli, Alessandra, and Laura Veldkamp. 2008. "Nature or Nurture? Learning and Female Labor Force Dynamics." Working Paper. New York: New York University.

[26] Ge, Suqin. 2008. "Women's College Choice: How Much Does Marriage Matter?" Working Paper. Blacksburg, V.A.: Virginia Tech.

[27] Gould, Eric, and Daniele Paserman. 2003. "Waiting for Mr. Right: Rising Inequality and Declining Marriage Rates." *Journal of Urban Economics* 53(2): 257–81.

[28] Goldin, Claudia. 1990. *Understanding the Gender Gap*. Oxford: Oxford University Press.

[29] Goldin, Claudia, and Lawrence Katz. 2002. "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions." *Journal of Political Economy* 110(4): 730–770.

[30] Goldin, Claudia, Lawrence F. Katz, and Ilyana Kuziemko. 2006. "The Homecoming of American College Women: The Reversal of the College Gender Gap." *Journal of Economic Perspectives* 20(4): 133–156.

[31] Greenwood, Jeremy, Nezih Guner, and John A. Knowles. 2003. "More on Marriage, Fertility, and the Distribution of Income." *International Economic Review* 44(3): 827–862.

[32] Greenwood, Jeremy, Ananth Seshadri, and Mehmet Yorukoglu. 2005. "Engines of Liberation." *Review of Economic Studies* 72(1): 109–133.

[33] Heckman, James. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1): 153–161.

[34] Jacob, Brian 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.

[35] Jones, Larry, Rodolfo Manuelli, and Ellen McGrattan. 2003. "Why Are Married Women Working So Much?" Research Department Staff Report 317. Minneapolis, M.N.: Federal Reserve Bank of Minneapolis.

[36] Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce. 1993. "Wage Inequality and the Rise in Returns to Skills." *Journal of Political Economy* 101(3): 410–442.

[37] Kambourov, Gueorgui, and Iourii Manovskii. 2005. "Accounting for the Changing Life-Cycle Profile of Earnings." Working Paper. Philadelphia: University of Pennsylvania.

[38] Kane, Michael. 1994. "College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education." *Journal of Political Economy* 102(5): 878–911.

[39] Katz, Lawrence, and Kevin M. Murphy. 1992. "Changes in Relative Wages, 1963–1987: Supply and Demand Factors." *Quarterly Journal of Economics* 107(1): 35–78.

[40] Keane, Michael P., and Kenneth I. Wolpin. 1997. "The Career Decisions of Young Men." *Journal of Political Economy* 105(3): 473–522.

[41] Keane, Michael P., and Kenneth I. Wolpin. 2001. "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." *International Economic Review* 42(4): 1051–1103.

[42] Korenman, Sanders, and David Neumark. 1991. "Does Marriage Really Make Men More Productive?" *Journal of Human Resources*, 26 (2): 282-307.

[43] Lee, Lung-Fei. 1978. "Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables." *International Economic Review* 19(2): 415–433.

[44] Mare, Robert D. 1991. "Five Decades of Educational Assortative Mating." *American Sociological Review* 56(1): 15–32.

[45] McGrattan, Ellen R., and Richard Rogerson. 2004. "Changes in Hours Worked, 1950–2000." *Federal Reserve Bank of Minneapolis Quarterly Review* 28(1): 14–33.

[46] Mincer, Jacob. 1974. *Schooling Experience and Earnings.* New York: Columbia University Press.

[47] Mulligan, Casey, B., and Yona Rubinstein. 2007. "Selection, Investment, and Womens Relative Wages." Working Paper. Chicago: University of Chicago.

[48] Pencavel, John. 1998. "Assortative Mating by Schooling and the Work Behavior of Wives and Husbands." *American Economic Review* 88(2): 326–329.

[49] Plug, Eric, and Wim Vijverberg. 2003. "Schooling, Family Background, and Adoption: Is It Nature or Is It Nurture?" *Journal of Political Economy* 111(3): 611–641.

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

[51] Stevenson, Betsey, and Justin Wolfers. 2007. "Marriage and Divorce: Changes and their Driving Forces." *Journal of Economic Perspectives* 21(2): 27–52.

[52] Sanchez-Marcos, Virginia. Forthcoming. "What accounts for the increase in college attainment of women?" *Applied Economic Letter*.

[53] Weiss, Yoram. 1997. "The Formation and Dissolution of Families: Why Marry? Who Marries Whom? and What Happens upon Marriage and Divorce?" In *Handbook of Population and Family Economics*, Volumn 1A, Rosenzweig and Oded Stark, eds. Amsterdam: Elsevier Science, pp. 81–124.

[54] Willis, Robert J., and Sherwin Rosen. 1979. "Education and Self-Selection." *Journal of Political Economy* 87(5): S7–S36.