Marriage, Intergenerational Schooling Effect, and Gender Gap in College Attainment

Suqin Ge  Fang Yang*
Virginia Tech       SUNY-Albany

First draft: August 2007
This version: December 2007

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 decision 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 increases of parental education and relative earnings between college and high school persons have important effects on the increase in college attainment for both genders, while the decrease of marriage rates is crucial in explaining the reversal of gender gap in college attainment.

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

JEL Classification: J24, J16, I20

*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). We thank Michele Boldrin, Mariacristina De Nardi, Zvi Eckstein, and seminar participants at Virginia Tech for helpful comments and suggestions. All remaining errors are our own.
I Introduction

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

We construct a formal economic model including the potential costs and benefits which determine individual college decisions. In our model, individuals with different ability first make schooling decisions. Then they may form marriages and have children. Parents are altruistic and value children’s ability. We assume that the higher a parent’s education, the more capable he/she is at increasing children’s ability. Forward-looking individuals take into account the impact of their own schooling on their parenting skills. Other factors that affect an individual’s education choices include the direct labor market returns to college, marriage market returns to college, and effort costs for attending 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 1980-1996 Panel Study of Income Dynamics (PSID). We present evidence on the fit of the model to the data. We then use the parameter estimates to simulate counterfactual experiments to decompose 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.

What accounts for the increase of college attainment? We find that the increasing relative earnings between college and high school persons have important effects on the increase in college attainment for both genders. We also emphasize the inter-generational persistence in schooling: a college-educated parent is substantially more
likely to have a college-educated daughter or son than a non-college graduate, even after controlling for the education of the other parent. This link between parents’ and children’s schooling provides an inter-generational propagation mechanism: as the number of college-educated parents increases, their children become more likely to attend college. Thus, the gradual transformation of the parental education itself acts as a propagation mechanism in changes in college attainment.

What accounts for females overtaking males in college attainment? We find that the decreasing marriage rates are crucial in explaining the relative increase in female college attainment. Two factors are relevant here. First, when married, the returns to college education are higher for males than for females. Second, when single, the return to college education is higher for females than for males. As marriage rates decrease, the return to college for single females becomes 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 the college enrollment and graduation by gender in the 1970s (see, for example, Susan L. Averett and Mark L. Burton (1996), David Card and Thomas Lemieux (2001), Jose-Victor Rios-Rull and Virginia Sanchez-Marcos (2002)). Among works that study the reversal of gender gap in college enrollment, proposed explanations include higher non-cognitive skills and college premiums among women (Brian A. Jacob (2002)), the increase of discount rates over time (Patricia M. Anderson (2002)), uncertainty of future wages (Kerwin K. Charles and Ming-Ching Luoh (2003)), the negative correlation between gender wage gap and female’s education (Pierre A. Chiappori, Murat Iyigun, and Yoram Weiss (2006)).

The paper is organized as follows. In Section II, we present some empirical results
from the PSID documenting college attainment rates in 1980-1996. In Section III, we present our model. Section IV 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. Brief concluding remarks are provided in Section V.

II Data

We use the PSID to calculate college attainment rates in 1980-1996. The PSID is a longitudinal survey of U.S. families and the individuals who make up those families. Approximately 4,800 U.S families are sampled in 1968 and these families have been re-interviewed annually until 1997. From 1997 onwards PSID was changed to biennial data collection and two major changes were made to the PSID: a reduction of the core sample, and the introduction of a refresher sample of post 1968 immigrant families and their adult children.

We select individuals in the core sample whose age is between 25-34 in each year and with valid information on parents’ education. We use completed schooling among mature adults as a measure of the schooling. Individual whose education completed by age 34 is more than 12 is defined as having college education. We first find parents’ education for the selected sample by linking parents and children from Individual Files (1968-2005). PSID facilitates the intergenerational linkage by providing parent’s ID in the individual files. If a linkage can not be found in Individual Files, we use 2003 Parent Identification File to link an individual with his/her parents. If the above procedure fails to provide parent’s education information, we find parents’ education by using parents’ and parents-in-laws’ education 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 wife’s parents. In the later waves, these parental education questions were asked for new head and wife. By merging Individual Files with Family Files, we are able to find parents’s education for those who were Head or wife or siblings of the Head.

Figure 1: College attainment by age 34. Source: Authors’ calculations from 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.\(^1\) 57% of young men aged 25 to 34 in 1980, as compared with 46% of young women, had some college education. In 1996, however, female college attainment reached 64%, 5 percentage points higher than that of males. In fact females have taken over males in college attainment since 1987.

We also calculate college attainment rates conditional on parent’s education. As Figure 2 shows, a college-educated parent is substantially more likely to have a college-educated daughter than a non-college graduate, even after controlling for the education of the other parent.\(^2\)

\(^1\)Other studies (see, for example, Kerwin K. Charles and Ming-Ching Luoh (2003), Claudia Goldin, Lawrence F. Katz and Ilyana Kuziemko (2006)), which use different measures of education or different data sets, find similar patterns.

\(^2\)Similar patterns hold for sons.
III The Model

The economy is populated by overlapping generations that live for 2 periods. We assume that going to college entails an idiosyncratic non-pecuniary effort cost $D \in [0, \infty)$. At the beginning of the first period, individuals with different costs make schooling decisions. In the second period, they form marriages and have children. Parents are altruistic and value their children’s ability. We assume that the higher is a parent’s education, the more capable he/she is at increasing his/her children’s ability. Forward-looking individuals take into account the impact of their own schooling on their parenting skills. Other factors that affect an individual’s education choices include the direct labor market returns to college and marriage market returns to college. These costs and benefits can differ by gender. We now describe the model in more detail.

Marriage and the Labor Market (second period)

In the second period, individuals of schooling type $s_f$ and $s_m$ marry at an exogenously given rate, and they work. Let $Y_{g,s_m,s_f}$ denote the earnings of an individual of gender, $g = \{f, m\}$, the education of husband is denoted by $s_m = \{0, h, c\}$, and the

---

3We abstract from tuition costs which should have similar marginal effects on the education of each gender and are unlikely to explain the gender difference in college attainment.
education of wife is denoted by \( s_f = \{0, h, c\} \), where 0 stands for single. If a person does not marry, he/she enjoys his/her own consumption\(^4\). The lifetime utility function for a single female and a single male of schooling \( s \) is, respectively,

\[
(1) \quad U_f(s_m = 0, s_f = s) = \log(Y_{f,0,s}), \\
(2) \quad U_m(s_m = s, s_f = 0) = \log(Y_{m,s,0}).
\]

A married couple has two children: one boy and one girl. Each values his/her own consumption and children’s learning ability. Each spouse gets a share of the total family income, with the weight of each spouse depending on his/her individual relative earnings through a parameter \( \lambda \in [0,1] \). The 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 weight is increasing in his/her share of household earnings. 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 ability is Cobb-Douglas: \( a' = s_m^{1-\theta_s}s_f^{\theta_s} \). A boy and a girl from the same family have the same learning ability. The utilities of men and women at a marriage type \((s_m, s_f)\) are given, respectively, by

\[
(3) \quad U_f(s_m, s_f) = \log[0.5((1 + \lambda)Y_{f,s_m,s_f} + (1 - \lambda)Y_{m,s_m,s_f})] + \lambda_a \log[s_m^{1-\theta_s}s_f^{\theta_s}], \\
(4) \quad U_m(s_m, s_f) = \log[0.5((1 + \lambda)Y_{m,s_m,s_f} + (1 - \lambda)Y_{f,s_m,s_f})] + \lambda_a \log[s_m^{1-\theta_s}s_f^{\theta_s}].
\]

**The College Decision (first period)**

The decision to go to college depends on the cost and the expected returns to

\(^4\)We abstract from divorce and out-of-wedlock birth and assume a single individual does not have children. We plan to extend our analysis in future research.
college. A female individual chooses whether to attend college, $s_f = 1$ (high school) and $s_f = 2$ (college), given her conditional marriage probabilities $P(s_m|s_f)$ and her individual cost of schooling $D$, by solving

$$\max_{s_f \in \{1,2\}} \sum_{s_m=0}^{2} U^f(s_m, s_f)P(s_m|s_f) - D.$$  

Note that $P(0|s_f)$ is the probability of being single. A male’s problem is defined analogously.

An individual is indifferent between going 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

$$(6) \quad D_f^* \equiv \sum_{s_m=0}^{2} U^f(s_f = 2, s_m)P(s_m|s_f = 2) - \sum_{s_m=0}^{2} U^f(s_f = 1, s_m)P(s_m|s_f = 1),$$

$$(7) \quad D_m^* \equiv \sum_{s_f=0}^{2} U^m(s_m, s_f = 2)P(s_f|s_m = 2) - \sum_{s_f=0}^{2} U^m(s_m, s_f = 1)P(s_f|s_m = 1).$$

Thus a female with the idiosyncratic schooling cost $D$ chooses $s_f = 1$ if and only if $D > D_f^*$ and a male chooses $s_m = 1$ if and only if $D > D_m^*$.

**Distribution**

Each individual receives a draw of effort cost in the first period. We assume that the children’s learning ability affects the distribution of effort cost from which each child draws $D$. More specifically, we assume that the effort cost $D$ is log-normally distributed with mean $\mu(a)$ and variance $\sigma^2$, where $\mu(a)$ is decreasing in the ability level $a$. Recall from above 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 $a$. The conditional enrollment rates $\psi^c_g(s_{m-1} = i, s_{f-1} = j)$ are calculated using cumulative distribution
function of $D$ at $D^*_g$: 

$$(8) \quad \psi^c_g(s^{-1}_m = i, s^{-1}_f = j) = F[D^*_g|a_{i,j}].$$

Notice that although an individual born to any family type would make decisions based only on his/her gender and the value of effort cost $D$, the fraction of individuals born to each family type 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 and high school be $\Phi^c_g$ and $\Phi^h_g$ respectively. Let $p^{-1}(s^{-1}_m = i, s^{-1}_f = j)$ be the fraction of fathers and mothers with education level $i$ and $j$, respectively. Thus the aggregate college attainment, $\Phi^c_g$, is the average of the conditional attainment rates weighted by parents education distribution.

$$(9) \quad \Phi^c_g = \sum_{i,j=1}^2 \psi^c_g(s^{-1}_m = i, s^{-1}_f = j) \ast p^{-1}(s^{-1}_m = i, s^{-1}_f = j).$$

### IV Findings

Can the model replicate the change of college attainment that occurred between 1980 and 1996? To do this, we calculate some of the parameters of the model directly from the data and estimate other parameters by matching aggregate and conditional moments on college attainment by gender in the data.

It is well known that people do not marry randomly and there exists assortative matching. A college educated person is more likely to marry a college educated spouse and benefit from the spouse’s earnings. We calculate the marriage distribution that each individual faces from the March supplement of Current Population Surveys (CPS). To be consistent with college attainment rates calculated from PSID, we restrict our
sample to those whose age is between 25-34 in each year between 1980 and 1996. Individuals in our sample are either never married or currently married. We calculate both the probabilities of being married and the probability of marrying each type of spouse conditional on being married. In the marriage market, there has been a decline of marriage rates for both genders. As Table 1 shows, from 1980 to 1996 the marriage rate has decreased by 21 percentage points for high school males, 14 percentage points for college males, 17 percentage points for high school females, 11 percentage points for college females. Meanwhile, Table 2 shows that as female college attainment increases, the probability of marrying a college female for a male increases over time and the opposite happens for a female.

<table>
<thead>
<tr>
<th></th>
<th>High School</th>
<th>College</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>1980</td>
<td>0.77</td>
<td>0.86</td>
<td>0.69</td>
<td>0.77</td>
</tr>
<tr>
<td>1985</td>
<td>0.69</td>
<td>0.80</td>
<td>0.64</td>
<td>0.72</td>
</tr>
<tr>
<td>1990</td>
<td>0.63</td>
<td>0.76</td>
<td>0.58</td>
<td>0.69</td>
</tr>
<tr>
<td>1996</td>
<td>0.56</td>
<td>0.69</td>
<td>0.55</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Table 1: Marriage Rates for Individuals Age 25-34 by Gender and Education

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>high school</td>
<td>college</td>
<td>high school</td>
<td>college</td>
</tr>
<tr>
<td>1980</td>
<td>16.8</td>
<td>66.6</td>
<td>30.1</td>
<td>80.9</td>
</tr>
<tr>
<td>1985</td>
<td>19.1</td>
<td>71.1</td>
<td>25.0</td>
<td>77.7</td>
</tr>
<tr>
<td>1990</td>
<td>21.8</td>
<td>74.5</td>
<td>21.7</td>
<td>74.4</td>
</tr>
<tr>
<td>1996</td>
<td>27.4</td>
<td>82.1</td>
<td>22.9</td>
<td>78.3</td>
</tr>
</tbody>
</table>

Table 2: Probability of Marrying a College Spouse Conditional on Own Gender and Education

We calculate the life-cycle earnings by gender, education, and marital status using earnings data from March CPS in 1964-2006. We define an individual as having college education if he/she completes more than 12 years of schooling and we define an individual as single if he/she has never married. For an individual who is in birth cohort $t$, with gender $g = \{f, m\}$, education $s = \{h, c\}$, and marital status $z = \{single, (s_m, s_f)\}$, we
define total discounted life-cycle earnings at the beginning of his/her adult life, \( Y^t_{g,s,z} \), as:

\[
Y^t_{g,s,z} = \sum_{a=25}^{65} \left( \frac{1}{1+r} \right)^{a-25} y^t_{g,s,z}(a),
\]

where \( r \) is the annual real interest rate and \( y^t_{g,s,z}(a) \) is the annual real earnings at age \( a = \{25, 26, ..., 65\} \).

For each birth cohort \( t \), we first construct a pseudo-panel between age 25 and 65 and deflate observed age-specific earnings by CPI. In the pseudo-panels we construct, usually not the entire life-cycle earnings profiles are observed. We use the following polynomial in age, \( a \), to estimate the life-cycle earnings profile for each birth cohort by gender, \( g = \{f, m\} \), education, \( s = \{h, c\} \), and marital status, \( z = \{\text{single}, (s_m, s_f)\} \):

\[
y^t_{g,s,z}(a) = \beta_{0,g,s,z} + \beta_{1,g,s,z} \cdot a + \beta_{2,g,s,z} \cdot a^2 + I(\text{cohort} = t) + \varepsilon_{g,s,z},
\]

where \( I(\text{cohort} = t) \) is a dummy for birth cohort. When we calculate total discounted life-cycle earnings in equation (10), \( Y^t_{g,s,z} \), we use an interest rate of \( r = 4\% \). Furthermore, predicted age-specific earnings from equation (11) are used if these earnings are not observed in our sample.

Between 1980 and 1996, we compute the average life-cycle earnings of those cohorts who aged 25-34 in each year. Table 3 presents the earnings return to college by gender and marital status. For singles, we compute the ratio between life-cycle earnings between college and high school males and females. For married couples, we compare earnings of households in which either the husband or the wife has college education with earnings of households where both spouses are high school graduates.

Several patterns are observed from Table 3 over the period 1980-1996. First, the earnings return to college is increasing for both genders. Second, the earnings return
Table 3: Earnings Return to College by Gender and Marital Status

to college is higher for single females than for single males. In 1980, single college-
educated males had 98% more earnings than single high school males, while single
college-educated females had 129% more earnings than single high school females.
Third, the earnings return to college is higher for married males than for married
females. For a typical household in 1980, where neither spouse attained college, having
the man to college gives the household 36% more earnings, while having the female to
college gives only 31% more earnings.

We calculate the schooling distribution of our PSID sample’s parents and show its
change over time in Figure 3. In 1980, 12% of individuals at age 25-34 had college-
educated parents, and 69% had high school parents. In 1996, the fraction changes to
23% and 50% respectively.

![Figure 3: Parents’ education. Source: Authors’ calculations from PSID data files.](image)

We use calculated earnings, marriage distributions, and parent education distribu-

tions as inputs of the model and estimate the rest of the parameters to match observed aggregate and conditional attainment rates by gender as in Figure 1 to 2. The estimated parameters are presented in Table 4. We find that \( \theta_s \) is less than 0.5, indicating that fathers’ education affects children’s ability more than mothers’ education, consistent with the observations from Figure 2. Effort cost distribution parameter \( \mu_{1,1} \) corresponds to \( \mu(a_{1,1}) \), etc. The rank order of \( \mu \)’s is consistent with that of \( a \)’s.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preference</td>
<td>( \lambda_a = 3.41, \lambda = 0 )</td>
</tr>
<tr>
<td>Ability production</td>
<td>( \theta_s = 0.41 )</td>
</tr>
<tr>
<td>Effort cost distributions</td>
<td></td>
</tr>
<tr>
<td>Means</td>
<td>( \mu_{1,1} = 0.61, \mu_{2,1} = 0.38 )</td>
</tr>
<tr>
<td></td>
<td>( \mu_{1,2} = 0.50, \mu_{2,2} = 0.31 )</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>( \sigma = 0.26 )</td>
</tr>
</tbody>
</table>

Table 4: Parameters used in the benchmark model

Figure 4 compares the college attainment from the model with that in the data. The model is able to generate the pattern that female college attainment was lower in 1980 and higher in 1996 than males, as is observed in the data. In the model, female college attainment began to exceed that of males in 1988, one year later than observed in the data.

![Figure 4: College attainment](image)

We run counterfactual simulations to study the effects of different mechanisms on
college attainment by comparing college attainment from each simulation with the benchmark given estimated parameters. The first simulation analyzes the case when there is no change in earnings since 1980. The results are shown in Figure 5. Both male and female attainment rates would drop by more than 10 percentage points in 1996, indicating that the increasing returns to college in the labor market have important impact on college attainment. In the second simulation, we investigate the intergenerational schooling effects. The results are shown in Figure 6. When parents’ schooling distribution is fixed at 1980 level, college attainment would drop by 6.8 and 6.2 percentage points in 1996 for males and females respectively. Therefore the gradual transformation of family schooling composition acts as a propagation mechanism in changing college attainment: as the number of college-educated parents increases, so does the proportion of children with high ability, which then helps to increase the attainment rate of the children generation. Both earnings and parental education are important sources of the increase in college attainment, but neither is about to account for the reversal of the gender gap.

The next two simulations try to isolate the effects of changes in the marriage market. Without change in the rates of marriage, as shown in Figure 7, both males and females
would reach higher college attainment in 1996 and female’s college attainment is always lower than that of males. The earnings returns to college in the labor market is higher for singles than for married couples. However, married individuals receive 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 rate. As marriage rate declines, returns to college decrease and so does college attainment. Moreover, as marriage rates decline female college attainment decreases less than that of males, because single females receive larger return to college in the labor market. This shows that changes in marriage rates are crucial in accounting for the reversal in gender gap in college attainment. In the fourth simulation we fix the conditional marriage probabilities as in 1980 and keep the marriage rates in the data. The results are shown in Figure 8. The 1996 college attainment would be 2.7 and 4.0 percentage points lower for males and females. Therefore the change of conditional marriage probabilities plays a relative minor role in accounting for the change in college attainment.

![Figure 7: No change in single rates since 1980](image1)

![Figure 8: No change in conditional marriage probability since 1980](image2)
V Conclusions

We develop a model to 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 increases of parental education and relative earnings between college and high school persons have important effects on the increase in college attainment for both genders, while the decrease of marriage rates is crucial in explaining the reversal of 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 earlier 1980 but out-of-wedlock birth has been increasing. The change in family structure might have different impact on female and male’s college attainment decision. Secondly, we assume earnings are exogenous. An extension that we wish to explore is the relationship among college attainment, marriage and labor supply for both gender. Even though labor earnings are sacrificed, a parent who stays at home and takes care of children would contribute to household utility by increasing the ability of children. We plan to study these issues in future work.

References


