What Explains the Growing Gender Education Gap?

The Effects of Parental Background, the Labor Market and the Marriage Market on College Attainment

by

Zvi Eckstein*, Michael Keane** and Osnat Lifshitz*

August 15, 2023

Abstract: In the 1960 cohort, American men and women graduated from college at the same rate, and this was true for Whites, Blacks and Hispanics. But in more recent cohorts, women graduate at much higher rates than men. To understand the emerging gender education gap, we formulate and estimate a model of individual and family decision-making where education, labor supply, marriage and fertility are all endogenous. Assuming preferences that are *common* across ethnic groups and *fixed* over cohorts, our model explains differences in all endogenous variables by gender/ethnicity for the ‘60-‘80 cohorts based on three exogenous factors: family background, labor market and marriage market constraints. Changes in parental background are a key factor driving the growing gender education gap: Women with college educated mothers get greater utility from college, and are much more likely to graduate themselves. The marriage market also contributes: Women’s chance of getting marriage offers at older ages has increased, enabling them to defer marriage. The labor market is the largest factor: Improvement in women’s *labor* market return to college in recent cohorts accounts for 50% of the increase in their graduation rate. But the labor market returns to college are still greater for men. Women go to college more because their *overall* return is greater, after factoring in marriage market returns and their greater utility from college attendance. We predict the recent large increases in women’s graduation rates will cause their children’s graduation rates to increase further. But growth in the aggregate graduation rate will slow substantially. [Abstract = 250 words]

Keywords: Returns to college, parental background, college graduation, education, gender wage gap, assortative mating, labor supply, marriage, fertility

Affiliations: Eckstein, Reichman University (zeckstein@runi.ac.il), Keane, Johns Hopkins University and UNSW (m.keane@unsw.edu.au), Lifshitz, Reichman University (osnat.lifshitz@runi.ac.il).

Acknowledgements: This research was supported by Australian Research Council (ARC) grants DP210103319 and CE170100005.

I. Introduction

College graduation rates in the US exhibit several key patterns that we seek to explain. Figure 1 plots college graduation rates of men and women using CPS data on 5-year birth cohorts from 1960 to 1990. In general, graduation rates have grown substantially. But they have grown much faster for women. In the 1960 birth cohort, women in all three ethnic groups graduated at rates very similar to men. But more recent cohorts of American women graduate from college at substantially higher rates than men, reversing a gender education gap in favor of men that existed for generations.[1] Differences across ethnic groups are also notable: Whites graduate at a much higher rate than Blacks, who in turn graduate at a higher rate than Hispanics. But, as a result of the general upward trend in the graduation rates across cohorts, the graduation rate for Blacks in the 1990 birth cohort is close to that for whites in the 1960 cohort.

In order to explain these patterns we formulate and estimate a model of individual and family decision making that can explain differences in college graduation rates by gender, and across ethnic groups and cohorts. Our model relies on 3 exogenous factors to explain education differences: Parental background (i.e., parent's education, marital status and immigration status), labor market constraints (wage offer functions and job offer functions) and marriage market constraints. We use the model to decompose differences in college graduation rates by gender/ethnicity/cohort into parts due to each of these 3 factors.[2] We impose discipline on the model by assuming common preferences across ethnic groups and cohorts. In our model education, labor supply, marriage and fertility are all endogenous, and we require the model to explain not only education but employment, marriage and fertility as well.

Parental background plays an important role in our model, as it affects skill endowments at age 16, tastes for school, and tastes for marriage. We model family background as consisting of three characteristics: mothers' education, parent's marital status and parent's immigration status. Parental education and immigration status play two key roles: They are correlated with endowments of labor market skill, and also shift tastes for education.[3] Similarly,

---

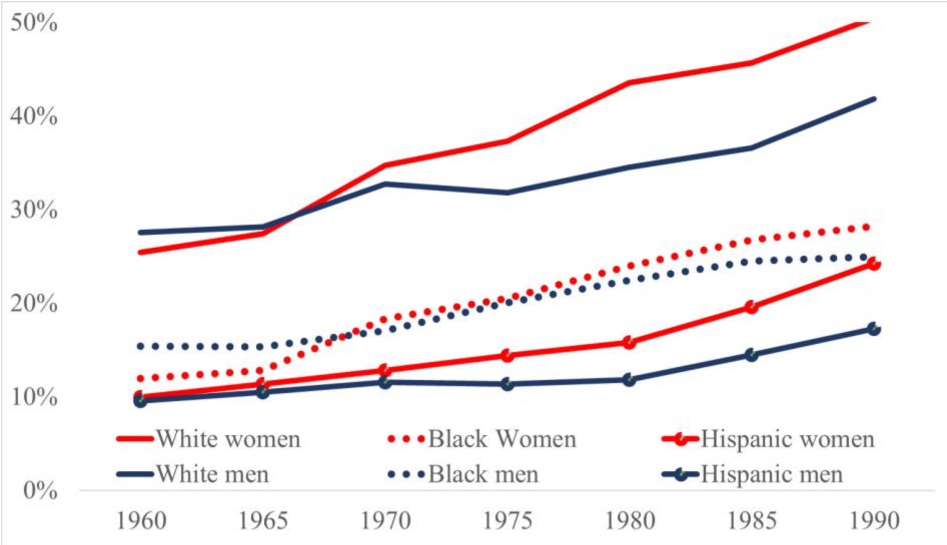
[1] In the '40 to '50 birth cohorts white men graduated college at an 8 percentage point higher rate than women. This gender gap closed suddenly in the '55 and '60 cohorts (i.e., people who started college in about 1973 to 1978). This was driven by a substantial drop in college attendance of men, and a very slight increase for women.

[2] In Eckstein, Keane and Lifshitz (2019) we explored factors that drove changes in white men's and women's college attendance in earlier birth cohorts from 1935-75. During that period white women's college graduation rate increased from 6% to 36%. We find this enormous increase can be explained by three factors (one third each): (i) increasing returns to college in the marriage market for women, (ii) changes in the wage structure that favored women, and (iii) increasing mother's education, which increased daughters' tastes for college.

[3] Obviously other mechanisms may also be at work. Families with college educated mother's may invest more in daughters human capital, and/or have higher aspirations for daughters. Having a college educated mother may raise a daughter's aspirations, or result in a daughter having better information about career opportunities after college. Our model captures these various mechanisms under the umbrella of increased taste for college.

parents' marital status is correlated with labor market skill, and it shifts tastes for marriage. As we show in Table 1, the differences in parental background across cohorts and ethnic groups are substantial, so they may be an important factor driving differences in education:

Figure 1: The Growth in the Gender Education Gap



Note: We plot the college graduation rate for 5-year birth cohorts from 1960 to 90.

Table 1 shows that the fraction of children whose mothers were college graduates is much greater for Whites than for Blacks or Hispanics, and it has increased substantially over time for all three groups. For example, for whites, it increased from 14% in the 1960 birth cohort to 38% in the 1990 cohort. Of special importance for Hispanics is that the percent with U.S. born parents increased substantially over cohorts – See Table 1, last column. This may have improved education prior to age 14, leading to better initial skill endowments. It may also imply changing tastes for education on the part of both parents and children.

Table 1: Mother's Education, Marital Status and Immigration Status by Cohort

	White		Black		Hispanic		
Cohort	% of CG+PC mothers	% of single mothers	% of CG+PC mothers	% of single mothers	% of CG+PC mothers	% of single mothers	% of US born mothers
1960	14.0%	6.4%	6.0%	37.0%	7.0%	12.7%	42.0%
1970	24.0%	11.8%	12.0%	48.9%	8.0%	18.6%	44.0%
1980	26.0%	18.7%	13.0%	61.5%	11.0%	25.9%	52.0%
1990	38.0%	26.6%	19.0%	67.6%	14.0%	31.1%	75.0%
2000	45.0%	21.7%	25.0%	69.8%	17.0%	33.2%	
2010	51.0%	21.9%	28.0%	74.5%	23.0%	40.0%	

Note: The table reports the % of college graduate mothers (two decades earlier at age 30), and the % of single mothers (two decades earlier at ages 23-30), both taken from CPS data. The % of US born Hispanics is taken from the American Community Survey.

The fraction of children born into single mother households also differs substantially across ethnic groups and cohorts. For whites, it increased from 6% in the 1960 cohort to 27% in the 1990 cohort. The rate for Hispanics is about 5 points higher. For Blacks the prevalence of single mother households is much higher: it increased from 37% in the 1960 cohort to 68% in the 1990 cohort. Table 2 shows that having a college graduate mother increases probability of college attendance while having a single mother reduces it, but the association is stronger for Whites and Hispanics than for Blacks.

Table 2: Association of College Graduation with Mother Education and Marital Status

	White		Black		Hispanic	
	Men	Women	Men	Women	Men	Women
Single mom	-0.565***	-0.656***	-0.160	-0.366	-0.648**	-0.492*
	(0.17)	(0.16)	(0.26)	(0.22)	(0.24)	(0.23)
College mom	1.141***	0.922***	0.340	0.423*	0.925**	1.539***
	(0.18)	(0.19)	(0.25)	(0.20)	(0.30)	(0.32)

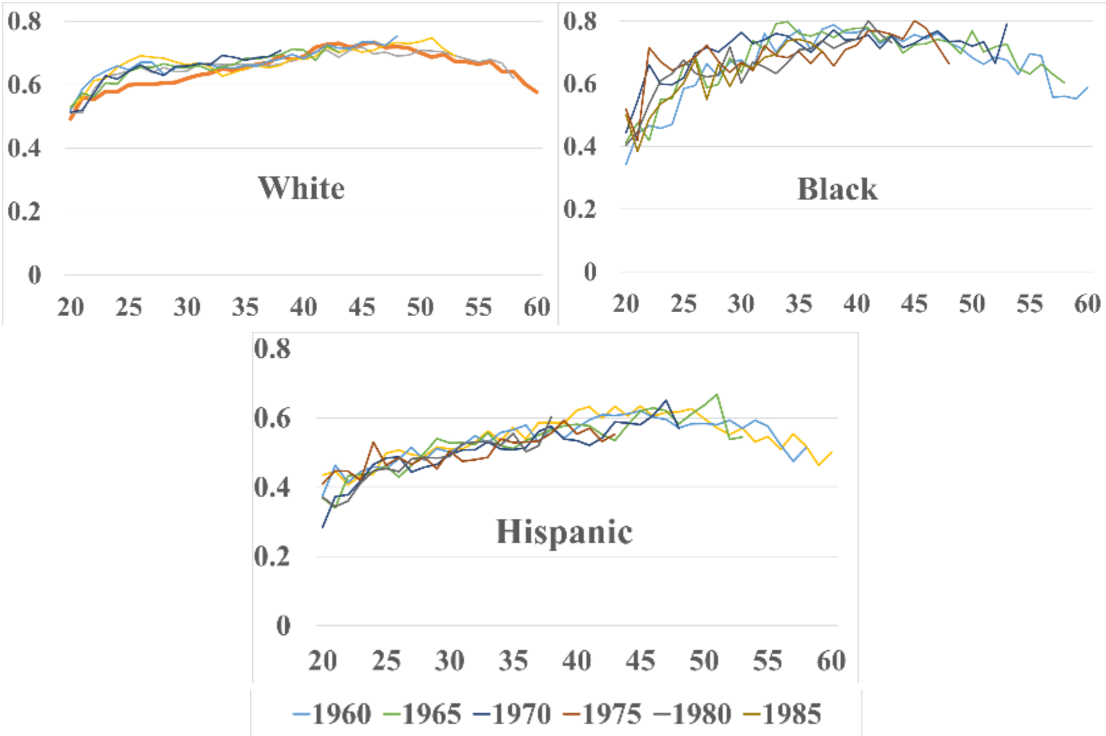
Note: NLSY data, college graduation by age 30, as a function of mother's education and marital status. We control for income at childhood, age of mother, year of birth, geographical variables.

A potential explanation for the rapid increase in women's education is that women are investing more in human capital because they are forward-looking and expect to work more. But this is inconsistent with the data. Figure 2 shows employment rates of married women have been remarkably stable since the 1960 cohort. As we showed in EKL, Figure C1, employment rates of married women increased dramatically from the '25 through '60 birth cohorts, but then stabilized. (Employment rates of unmarried women have always been high). Thus, increased employment plays a key role in explaining the growth in women's education prior to the '60 cohort, but not since. A key challenge is to explain substantial increases in women's education while holding their employment fixed, using fixed preferences for leisure across cohorts.

A brief summary of our results is as follows: Assuming preferences that are *common* across ethnic groups and *fixed* over cohorts, our model is successful in explain the differences in all endogenous variables (education, labor supply, marriage and fertility) by gender and ethnicity for cohorts from 1960-80 based on just three exogenous factors: family background, labor market and marriage market constraints. We find changes in parental background are a key factor driving the growing gender education gap: Women with college educated mothers get greater utility from college, and are much more likely to graduate. The marriage market also contributes: Women's chance of getting marriage offers at older ages has improved, making it easier to defer marriage until after college. The largest factor is the labor market:

The improvement in womens' labor market return to college in recent cohorts accounts for 50% of the increase in their graduation rate. Nevertheless, the labor market returns to college are still greater for men than women. The reason women go to college more than men is not that their labor market return is greater: It is that their *overall* rate of return is greater after factoring in marriage market returns and their greater utility from college attendance.

Figure 2: Married Women's Employment over the Life-Cycle, by Birth Cohort



Note: CPS data on 5-year birth cohorts from 1960 to 1985.

Using our model to forecast behavior of future cohorts, we predict the graduation rate of white women has peaked with the '90 to '00 cohorts, and will plateau at about 53% going forward. The graduation rate of white men will (very) slowly catch up, causing the gender education gap to narrow gradually. Education gaps between whites and both Blacks and Hispanics will remain substantial unless offer wage functions converge. Otherwise, substantial subsidies would be required to equalize educational opportunity between ethnic groups.

The outline of the paper is as follows: In Section II we present and overview of our approach and discuss the literature. Section III presents our model. In Section IV we explain how we fit the model to CPS data from the '60, '70 and '80 birth cohorts. We also show how, in an out-of-sample validation, our model provides a good fit to behavior of the 1990 cohort. Section V discusses the parameter estimates. Section VI presents our decompositions of education differences across groups into parts due to the various exogenous factors. Section

VII presents our forecasts of future education levels. Section VIII presents education subsidy experiments aimed at equalizing opportunity across groups, and Section IX concludes.

II. Overview of our Approach

In this paper we present a structural life-cycle model of education, marriage, fertility and labor supply decisions of men and women that succeeds in capturing all the key patterns in Figure 1, as well as fitting a broad range of other demographic and labor market outcomes. We fit the model to the behavior of the 1960-1980 birth cohorts. We then use the model to decompose the sources of the changes in educational attainment across cohorts, broken down by gender and race/ethnicity.

Our model builds on Eckstein, Keane and Lifshitz (2019) – henceforth EKL. They present a life-cycle model of education, labor supply, marriage and fertility decisions. That model, in turn was based on the separate life-cycle models for men and women developed in Keane and Wolpin (1997, 2010). EKL's innovation was to *jointly* model life-cycle behavior of men and women in a unified framework that accounts for the possibilities of marriage/divorce. When in the married state couples make labor supply decisions jointly, while in the single state agents make decisions as individuals. All agents in the model retain their individual identity (and utility functions) through marriage and seek to maximize own lifetime utility. EKL show their models provides an excellent fit to education, labor supply, marriage and fertility decisions, as well as observed wages, for five cohorts of white men and women in the US.

Here we extend the EKL model to account for behavior of Blacks and Hispanics in the four cohorts born from 1960-90. This would be a trivial exercise if it were simply a matter of fitting the same model to Blacks and Hispanics. However, as we have discussed, the behavior of the 1960-80 cohorts differs tremendously between Whites, Blacks and Hispanics, particularly with regard to education. Our challenge is to find a *parsimonious* generalization of the EKL model that can explain these substantial behavioral differences, including both differences *across* ethnic groups as well as changes in behavior of each group *over time*.

A key aspect of the EKL model is that it imposes invariant preferences across cohorts. Thus, all changes in behavior over time are attributed to changes in the environment that the cohorts faced.[4] Here we show – rather remarkably – that we can also explain differences in behavior across Whites, Blacks and Hispanics using preferences that are both common across the three groups and invariant over time.

[4] We found the most important differences that explained different outcomes for the 1935-75 birth cohorts were differences in (i) offer wages, (ii) parental background – which altered skill endowments, (iii) differences in the distribution of potential marriage partners, and (iv) improvements in contraception technology.

The only parameters that we allow to differ by ethnic group are the wage offer functions, job offer probabilities, marriage offer distribution – i.e., the probabilities of meeting potential partners with particular education levels – and the stochastic process for health. Importantly, these parameters can only vary across groups in a rather restricted way, as the model must generate simulated wages, employment rates and marriage rates (by education) that are consistent with observables. For example, in the data we see that Hispanic men have very low rates of college attendance. The model cannot automatically explain this simply by assuming they have low returns to education, as that explanation is only tenable if it is consistent with the patterns of wages by education level actually observed in the wage data.

We extend EKL in three key ways: First, we let the distribution of the labor market skill endowment depend not only on parents' education but also on their marital status. This accommodates in a simple way the possibility that investments in child development may be greater in couples than in single parent families.

Second, we let the taste for marriage depend on parents' marital status. This generates inter-generational persistence in tastes for marriage. Frequency of marriage differs substantially across parents of Whites/Blacks/Hispanics, which in turn generates differences in skill endowments and tastes for marriage among their children. This is important for our model to explain the behavioral differences across the three groups.

Third, we include welfare participation as a choice. We model welfare as an AFDC/TANF type system that provides benefits to single mothers only, and where participation generates "welfare stigma" as in Moffitt (1984). The welfare reform of 1996 made welfare receipt more difficult by imposing work requirements and time limits on recipients. We model the reform as introducing a 5-year time limit on benefit receipt. And we model work requirements parsimoniously as increasing the fixed cost of participation.

The impact of welfare programs on behavior depends crucially on both skill endowments and marriage market opportunities. As these differ across Whites, Blacks and Hispanics, the impact of welfare on the behavior of the three race/ethnic groups differs in important ways as well.[5] We find that changes in welfare rules play a modest role in explaining increases in education for Blacks and Hispanic women from the 1960 to 1980 cohorts. The latter cohort is more affected by the welfare reforms of the mid to late 1990s.

[5] In EKL (2019) we found that accounting for welfare participation was not especially important for explaining the behavior of white cohorts over time.

III. A Life-Cycle Model of Education, Labor Supply, Marriage/Divorce and Fertility

Agents enter the model at age 17 as single individuals in school. Both men and women make annual private decisions about school continuation and work. In addition, women make annual decisions about fertility, and single mothers decide whether to participate in a welfare program (if eligible). We assume only single people can attend school.[6] Retirement is enforced at age T=65, at which point agents receive a terminal value function. The men and women in the model also interact in a marriage market, so they can choose to form (and later dissolve) couples. Once a couple is formed, decisions about labor supply and fertility are made jointly.

To make marriage decisions, individuals compare the values of the married and single states. Thus, we first describe the problem of married couples, followed by that of single individuals. We are then in a position to explain how we model the marriage market.

III.A. The Decisions of a Married Couple

We assume a collective model of household decision making. Let t denote the annual time period, and let $j = f, m$ denote gender. Individuals have time endowments of one unit per period. This is split between market work (h), and leisure (l), hence, $h_t^j + l_t^j = 1$. Agents can choose to work full-time, part-time, or not at all. Thus, the labor supply choice each period is $h_t^j \in \{1, 0.5, 0\}$, corresponding to a leisure choice of $l_t^j \in \{0, 0.5, 1\}$. We assume that full-time and part-time work correspond to 2000 and 1000 annual hours, respectively.

Conditional on marriage, couples have three choice variables: Leisure of the husband and wife, $\{l_t^m, l_t^f\}$, and pregnancy, $p_t \in \{0, 1\}$. Pregnancy leads deterministically to arrival of a child at t+1. Let X_t^j denote work experience, and N_t denote number of children under 18. These state variables have laws of motion $X_{t+1}^j = X_t^j + h_t^j$ for $j = m, f$ and $N_{t+1} = N_t + p_t - p_{t-18}$. Couples also make annual decisions about divorce/marriage continuation. We ignore this for now and focus on the joint decisions of couples' *conditional* on marriage.

III.A.1. *Preferences and Constraints*

Married couples have total gross income GY_t^M given by:

$$(1) \qquad GY_t^M = (w_t^m h_t^m + w_t^f h_t^f)$$

Here w_t^j and h_t^j for $j = f, m$ are annual full-time wage rates.[7] We will use the M superscript throughout to indicate values for married individuals.

---

[6] School attendance by married people is rare. We rule out school attendance after age 30 for the same reason.

[7] In EKL (2019) we also included a term $b_m I[h_t^m = 0] + b_f I[h_t^f = 0]$, where the b_j are unemployment benefits plus values of home production. But both there and here we find the b_j are not separately identified from the value of leisure parameters β_{jt} defined in equation (4b) below.

Net income is Y_t^M given by the equation:

(2) $$Y_t^M = GY_t^M - \tau_t^M\big((w_t^m h_t^m + w_t^f h_t^f), N_t\big),$$

where $\tau_t^M(\cdot,\cdot)$ is the tax function for married couples based on the time *t* tax rules. We model the US federal tax system in detail, including deductions, exemptions, EITC, and the joint taxation of couples (see Appendix B). We assume perfect foresight regarding tax rules.

The household budget constraint takes the form:

(3) $$C_t^M = (1 - \kappa(N_t))Y_t^M$$

Here $\kappa(N_t)$ is the fraction of Y_t^M spent on children, based on a square root equivalence scale.[8] We assume a static budget constraint as it is computationally infeasible to add saving in addition to our other state variables. However, the terminal value function (at age 65) proxies for how labor supply affects Social Security and retirement assets, so these key aspects of savings do enter our model in a reduced form way.

The period utility of a married person of age *t* and gender *j* in state Ω_{jt} is given by:[9]

(4a) $$U_t^{jM}(\Omega_{jt}) = \frac{1}{\alpha}(\psi C_t^M)^\alpha + L_{jt}(l_t^j) + \theta_t^M + \pi_t^M p_t + A_j^M Q(l_t^f, l_t^m, Y_t^M, N_t) \qquad j = m, f$$

(4b) $$L_{jt}(l_t^j) = \frac{\beta_{jt}}{\gamma}(l_t^j)^\gamma + \mu_{jt} l_t^j \qquad \gamma < 1, \alpha < 1$$

The first term in (4a) is a CRRA in consumption with curvature parameter α. We assume household consumption C_t^M is a "public" good. That is, the full amount C_t^M enters the utility of both the husband and wife. The parameter $\psi \in (½, 1)$ captures household economies of scale in consumption. The square root equivalence scale gives $\psi = 1/\sqrt{2} = 0.707$, so a couple needs 41% more expenditure than a single person to obtain an equivalent consumption level.

The second term in (4a) captures the value of leisure and home production. The third term (θ_t^M) is the utility from marriage itself (i.e., the match quality), the fourth term captures the utility (or dis-utility) from pregnancy ($p_t = 1$), and the fifth term captures utility from the quality and quantity of children. We now discuss the 2nd through 5th terms in more detail:

[8] For a household with two adults, the square root scale implies that $\kappa(N) = 1 - \sqrt{2/(2+N)}$.

[9] The state vector Ω_{jt} contains four variables that are relevant for $U_t^{jM}(\Omega_{jt})$. These are N_t and μ_{jt}, as well as parent's education and marital status, and health. The state vector Ω_{jt} contains several additional variables, whose role will only become clear after the full model is laid out. Thus, we defer giving the complete list of elements of Ω_{jt} until we finish expositing the full model and turn to discussing the DP problem solution (Section IV).

III.A.1.1. ***Tastes for Leisure and Value of Home Production***

The second term in (4a) consists of two parts, written out explicitly in (4b):

The first part is a CRRA in leisure with curvature parameter γ. The parameter β_{jt}, which must be positive, shifts tastes for leisure. For women we allow β_{jt} to depend on p_t, while for both men and women we allow β_{jt} to depend on education and on health status.

The second part captures stochastic variation in the marginal utility of leisure. In (4b) this is denoted by $\mu_{jt}l_t^j$ where μ_{jt} is a random variable. We assume that:

(5) $$\ln(\mu_{jt}) = \tau_{0j} + \tau_{1j}\ln(\mu_{j,t-1}) + \tau_{2j}p_{t-1} + \varepsilon_{jt}^l \qquad where \qquad \varepsilon_{jt}^l \sim iidN(0, \sigma_\varepsilon^l)$$

where $0 < \tau_{1j} < 1$. Thus, shocks to tastes for leisure (i.e., home time) follow a stationary AR(1) process. Importantly, the arrival of a new child at time t (i.e., $p_{t-1} = 1$) leads to a shift in tastes for home time (τ_{2j}). We expect that, particularly for women, the marginal utility of home time will jump up when a newborn arrives (i.e., $\tau_{2f} > 0$), capturing an increase in time required for home production and the desire to spend time with the child. Afterward, provided no new children arrive, tastes for home time gradually revert to normal, as $\tau_{1f} < 1$. This lets us generate the decline in women's employment after childbirth, as well as their subsequent gradual return to the labor force. The stochastic terms ε_{jt}^l generate heterogeneity in these response patterns.

III.A.1.2. ***Match Quality, Utility of Pregnancy, Child Production***

The last three terms in equation (4a) capture the utility from marriage, the dis-utility from pregnancy ($p_t = 1$), and the utility from children. First, consider the utility from marriage (θ_t^M) – i.e., the match quality. We write:

(6) $$\theta_{tj}^M = d_1 + d_2 \cdot I[E^m - E^f > 0] + d_3 \cdot I[E^f - E^m > 0] + d_4(H_t^m - H_t^f)^2 + \varepsilon_t^M$$

where $\varepsilon_t^M \sim iidN(0, \sigma_\varepsilon^M)$ and E^j denotes education, rank ordered as high school dropout (HSD), high school (HSG), some college (SC), college (CG) and post-college (PC), and $H_t^j \in \{1, 2\}$ denotes health (i.e., good or poor). The 2nd and 3rd terms capture assortative mating on education. $I[E^m - E^f > 0]$ indicates the man has greater education than the woman, and $I[E^f - E^m > 0]$ indicates the reverse. If $d_3 < 0$ people are averse to matches where the woman has more education. The 4th term captures assortative mating on health. If $d_4<0$ people prefer matches where partners have similar health. Finally, ε_t^M is a transitory shock to match quality.

Next consider the utility from pregnancy, π_t. We specify that:

(7) $$\pi_t = \pi_1 M_t + \pi_2 H_{ft} + \pi_3 N_t + \pi_4 p_{t-1} + \varepsilon_t^p$$

where $\varepsilon_t^p \sim iidN(0,\sigma_\varepsilon^p)$. Here π_t is a function of marital status, where M_t is a 1/0 indicator for marriage, the woman's health, the number of already present children and lagged pregnancy. The presence of health H_{ft} helps generate that fertility declines with age. Note that equation (7) contains nothing individual specific. We assume pregnancy decisions are made jointly by the couple, and each party gets the same utility from the decision.[10]

Finally, consider the function $Q(\cdot)$ that determines the utility a couple receives from the quality and quantity of children. We assume it is a CES function of the inputs, as follows:

$$(8)\quad Q\left(l_t^f, l_t^m, Y_t^M, N_t\right) = \left(a_f\left(l_t^f\right)^\rho + a_m(l_t^m)^\rho + a_g(\kappa(N_t)Y_t^M)^\rho + (1 - a_f - a_m - a_g)N_t^\rho\right)^{1/\rho}$$

The first three inputs, which are home time of parents and spending per child, $\kappa(N_t)Y_t^M$, all increase child quality. The parameter A_j^M in the utility function (4a) is a scale parameter that multiplies $Q(\cdot)$. This parameter is allowed to differ in the single state (see below).

III.A.2. *Value Function of a Married Couple*

We are now able to write the choice-specific value functions for married *individuals*. These depend on both a person's own state and that of their partner:

$$(9)\quad V_t^{jM}\left(l_t^m, l_t^f, p_t | \Omega_{mt}, \Omega_{ft}\right) = \frac{1}{\alpha}(\psi C_t^M)^\alpha + L\left(l_t^j\right) + \theta_t^M + \pi_t p_t + A_j^M Q\left(l_t^f, l_t^f, Y_t^M, N_t\right)$$
$$+\delta E_{MAX}\left(M_{t+1}V_{t+1}^{jM}\left(\Omega_{m,t+1}, \Omega_{f,t+1}\right) + (1 - M_{t+1})V_t^j\left(\Omega_{j,t+1}\right)\right) \qquad j = f, m$$

The current payoff simply reproduces (4). The future component in (9) consists of two parts, corresponding to whether the marriage continues at t+1 or not. The term $V_{t+1}^{jM}\left(\Omega_{m,t+1}, \Omega_{f,t+1}\right)$ is the value of next period's state for partner j if the marriage continues. The newly defined term $V_t^j\left(\Omega_{jt+1}\right)$ is the value of next period's state for partner j if he/she becomes single (i.e., a divorce occurs). We discuss divorce and the value functions for single persons below.

The t+1 state depends on the current state $\{\Omega_{mt}, \Omega_{ft}\}$ and current choices $\left\{l_t^m, l_t^f, p_t\right\}$ via the laws of motion of the state variables. δ is the discount rate and $E_{MAX}(\cdot)$ is the expectation taken over elements of the t+1 state that are unknown at t. These include M_{t+1}, $\left\{\varepsilon_{jt+1}^l\right\}$ for $j = m$, f, ε_{t+1}^M and ε_t^p, as well as realizations of wage shocks and job offers. We defer a detailed discussion of these until Section III.C which describes the labor market.

III.A.3. *Household Decision Making for Married Couples*

In our collective model the household value function is given by:

[10] Of course, one could imagine individuals in a couple getting different utilities from a pregnancy decision, but we cannot infer such differences from the data so we ignore them.

$$(10)\quad V_t^M\left(l_t^m, l_t^f, p_t | \Omega_{mt}, \Omega_{ft}\right) = \lambda V_t^{fM}\left(l_t^m, l_t^f, p_t | \Omega_{mt}, \Omega_{ft}\right) + (1-\lambda) V_t^{mM}(l_t^m, l_t^f, p_t | \Omega_{mt}, \Omega_{ft})$$

Here λ and (1-λ) are Pareto weights. We set $\lambda=0.5$ for simplicity.[11] The V_t^{jM} for *j=f, m* are the choice-specific value functions of the *individual* married partners. The Ω_{jt} for *j=f, m* are the state vectors of these individuals. Couples seek a choice vector $\{l_t^m, l_t^f, p_t\}$ to maximize (10), subject to the constraint that both parties prefer marriage over the outside option of divorce.[12]

Let $V_t^m(\Omega_{mt})$ and $V_t^f(\Omega_{ft})$ denote the maximized value functions of single males and females in period *t*. Utility is not transferable, so a divorce occurs if the value of the outside (single) option exceeds the value of marriage for *either* party. Let $\mathcal{F}$ denote the feasible set of choice options. A choice vector $\{l_t^m, l_t^f, p_t\} \in \mathcal{F}$ if:

$$(11)\qquad V_t^{jM}\left(l_t^m, l_t^f, p_t | \Omega_{mt}, \Omega_{ft}\right) \geq V_t^j\left(\Omega_{jt}\right) - \Delta_{jt} \quad for \quad j = f, m$$

where Δ_{jt} is the cost of divorce.[13] If no choice vector $\{l_t^m, l_t^f, p_t\}$ satisfies (11) then $\mathcal{F} = \emptyset$.

We can now formally define the solution to the maximization problem. Denote the vector of household choices that maximize equation (10) as $\{l_t^{m*}, l_t^{f*}, p_t^*\}$. That is,

$$\{l_t^{m*}, l_t^{f*}, p_t^*\} = \begin{cases} arg \max\limits_{\{l_t^m, l_t^f, p_t\} \in \mathcal{F}} V_t^M\left(l_t^m, l_t^f, p_t | \Omega_{mt}, \Omega_{ft}\right) & if \quad \mathcal{F} \neq \emptyset \\ \emptyset & if \quad \mathcal{F} = \emptyset \end{cases}$$

The form of (10) insures that $\{l_t^{m*}, l_t^{f*}, p_t^*\}$ is a Pareto efficient allocation. If one or more parties prefer to remain single for all possible $\{l_t^m, l_t^f, p_t\}$ then $\mathcal{F} = \emptyset$ and a divorce occurs.

The maximized value function of a married *individual* in state Ω_{jt} is given by:

$$(12)\quad V_t^{jM}\left(\Omega_{mt}, \Omega_{ft}\right) \equiv \begin{cases} V_t^{jM}\left(l_t^{m*}, l_t^{f*}, p_t^* | \Omega_{mt}, \Omega_{ft}\right) & for \quad j = f, m \quad if \quad \mathcal{F} \neq \emptyset \\ -\infty & for \quad j = f, m \quad if \quad \mathcal{F} = \emptyset \end{cases}$$

The maximized value function depends on both the own state Ω_{jt} and that of the partner.[14]

[11] This simple specification is similar to Voena (2015), who considers a household planning problem with a unilateral divorce regime. We discuss our decision to use this simple specification in EKL (2019) Section V.

[12] If we take the *unconstrained* maximum of (10), we might obtain a solution for $\{l_t^m, l_t^f, p_t\}$ where only one party prefers to stay married. In a transferable utility framework, a marriage may persist in such a case, using transfers between partners. We do not adopt this approach, as such transfers would be difficult to enforce.

[13] The cost of divorce depends on the number of children, $\Delta_{jt} = \alpha_4^j + \alpha_5^j \mathrm{N}_t$ where parameters α_4^j and α_5^j may change over time due to changing divorce laws.

[14] Note that if $\mathcal{F} = \emptyset$ then no action exists such that person *j* can be married at time *t*, so a divorce occurs. Then $V_t^{jM} = -\infty$, so behavior is governed solely by the single value function $V_t^j\left(\Omega_{jt}\right)$.

III.B. The Decisions of Single Households

In this Section we describe the optimization problems of single (i.e., unmarried) women and men. Both single men and women choose their levels of labor supply $h_t^j \in \{1, 0.5, 0\}$ and whether to attend school, $s_t \in \{0,1\}$. In addition, eligible single women with children may also decide to participate in a welfare program, $g_t \in \{0,1\}$. The gross income of a single man is simply $GY_t^m = w_t^m h_t^m$, while gross income of a single woman is:

$$GY_t^j = w_t^j h_t^j + CS \cdot a_t \cdot I[N_t > 0] + wb(N_t, w_t^j h_t^j, G_t) \cdot g_t \quad (13)$$

Here a_t is an indicator that a single woman with children receives child support *CS*. This occurs with probability $P(a_t = 1)$. We let $CS = \exp(\varepsilon_t^p)$ where $\varepsilon_t^a \sim N(\mu_a, \sigma_\varepsilon^a)$.

The term $wb(N_t, w_t^j h_t^j, G_t)$ denotes the level of welfare benefits to which a single mother with N_t children and income $w_t^f h_t^f$ is eligible. The state variable G_t is number of years the woman has received benefits in the post-1996 period and prior to age *t*.

The $wb_t(\cdot)$ function is designed to capture the array of social benefits targeted at single mothers in the US. These include AFDC/TANF benefits, public housing, childcare subsidies, etc. Rather than model the rules of these programs in detail, we treat $wb_t(\cdot)$ as an exogenous stochastic process that we fit from data prior to estimation, using the simple function form:

$$wb_t(N_t, w_t^f h_t^f, G_t) = \begin{cases} \beta_{0t} + \beta_{1t} N_t + \beta_{2t} w_t^f h_t^f & if \quad G_t < 5 \\ 0 & if \quad G_t > 5 \end{cases} \quad (14)$$

Importantly, the benefit rule $wb_t(\cdot)$ provides a natural exclusion restriction. It affects behavior of single women *directly* through the budget constraint, but it only affects behavior of married women, and all men, *indirectly* through the marriage market and household bargaining.

The net income of a single person is given by:

$$Y_t^j = GY_t^j - \tau_t^S(w_t^j h_t^j, N_t) \quad j = f, m \quad (15)$$

where $\tau_t^S(w_t^j h_t^j, N_t)$ is the time *t* tax function for single individuals calculated using the tax rules described in the Appendix B. Thus, the budget constraint for a single person is simply:

$$C_t^j = (1 - \kappa(N_t)) Y_t^j \quad (16)$$

Note that both single men and women may have children ($N_t > 0$). These may be children from a previous marriage or, in the case of single women, children born outside of marriage.

In our model, utility functions exist at the *individual* level, and are not fundamentally altered by marriage. Within marriage, collective household decisions are made by constrained maximization of a weighted average of the individual partners' utility functions, as in (10). Consistent with this, we specify the utility functions of singles to be as similar as possible to those of married individuals. Consider the per-period utility function of a single female:

$$(17) \quad U_t^f(\Omega_{ft}) = \left(\frac{1}{\alpha}(C_t)^\alpha + L_j(l_t) - \Psi g_t\right)(1 - s_t) + \vartheta_{ft} s_t + \pi_t p_t + A_f^s Q(l_t, 0, Y_t, N_t)$$

where Ψ is a disutility welfare participation and ϑ_{ft} is utility from attending school.

The Ψ may be thought of as capturing social "stigma," along with the time and effort costs associated with the various work/training/search and reporting requirements imposed on welfare participants. The welfare reform of 1996 can be thought of as making these requirements more stringent, so we let Ψ increase in 1996.

Recall that only single people can attend school (s_t = 1). In (17), students receive ϑ_{jt} for j=m,f as a current payoff to school attendance, rather than the function of (C_t, L_t) that workers receive. This asymmetry is motivated by the fact that we cannot measure "leisure" time for students in a way comparable to that for workers,[15] and a desire to avoid modelling how consumption is financed by students.[16] Hence, consistent with prior work like Keane and Wolpin (1997), we simply define a "utility while in school" variable, given by:

$$(18) \quad \vartheta_{jt} = \vartheta_{0j} + TC \cdot I(E_t > HSG) + \vartheta_{1j} PE + \vartheta_{2j} \mu_j^W \quad \text{for } j=m,f$$

Here ϑ_{jt} is a function of tuition cost TC, which is only relevant for higher education,[17] the unobserved skill endowment μ_j^W, and parents' education, denoted by PE. Parental education may affect tastes for school directly, or it may affect ability at school, and hence the effort required for success. More educated parents also make larger financial transfers to support students in school – see Keane and Wolpin (2001).

Education evolves as follows: All people start out in school at age 16, at education level "HSD." Two years of school are required to become a high school graduate (HSG). For 1 to 3 additional years of school the person is at some college (SC) level, while 4 more years is needed to become a college graduate (CG). Any additional years lead to the PC level.

If a single woman is not in school (s_t = 0), her utility function is fundamentally identical to that of a married woman, as one can see by comparing (4) and (17). The only differences are that in (17) consumption is individual specific (i.e., ψ = 1), utility from marriage is (of course)

[15] While we can see hours of market work, we cannot measure hours spent on school work.

[16] Consumption while in school may be financed by a combination of parental transfers, financial aid, part-time work, etc. (see Keane and Wolpin (2001). It is beyond the scope of our paper to model all these possibilities.

[17] Note that ϑ_{ft} captures utility of school net of costs. Without data on costs these can't be identified separately.

dropped, utility from children is allowed to differ from the married state ($A_f^s \neq A_j^M$), the home-time of the husband is set to zero in the Q function, and welfare participation is an option.

We can now write the choice-specific value functions for single females:

(19a) $$V_t^f\left(l_t, p_t, s_t, g_t | \Omega_{ft}\right) = \left(\frac{1}{\alpha}(C_t)^\alpha + L_f(l_t) - \Psi g_t\right)(1 - s_t) + \vartheta_{ft} s_t + \pi_t p_t + A_f^s Q(l_t, 0, Y_t, N_t) + \delta E_{MAX} V(\Omega_{f,t+1})$$

(19b) $$E_{MAX} V\left(\Omega_{f,t+1}\right) = E_{MAX}\left(M_{t+1} V_{t+1}^{fM}\left(\Omega_{m,t+1}, \Omega_{f,t+1}\right) + (1 - M_{t+1}) V_t^f\left(\Omega_{f,t+1}\right)\right)$$

where $E_{MAX} V\left(\Omega_{f,t+1}\right)$ takes into account that the person may get married at *t*+1. The choice-specific value functions $V_t^m(l_t, s_t | \Omega_{mt})$ for single men are analogous, except they do not include the welfare participation (g_t) and pregnancy (p_t) options.

Now we consider the optimization problem of singles. In Section III.E we discuss the marriage market, but we must first consider decision making *conditional* on being single – i.e., the state where no marriage offer is available or where it has already been declined.

Let $V_t^m(\Omega_{mt})$ and $V_t^f(\Omega_{ft})$ denote the maximized value functions of single males and females in period *t*. Let $\mathcal{S}_t^m$ and $\mathcal{S}_t^f$ denote the feasible set of choice options for a single male and female in period *t*, respectively. As we will see in Section III.C, workers receive job offers probabilistically, so $\mathcal{S}_t^m$ and $\mathcal{S}_t^f$ may not include all possible levels of work hours and leisure. To proceed, for women and men we have, respectively:

(20) $$V_t^f\left(\Omega_{ft}\right) = \max_{\{l_t, p_t, s_t, g_t\} \in \mathcal{S}_t^f} V_t^f\left(l_t, p_t, s_t, g_t | \Omega_{ft}\right)$$

(21) $$V_t^m(\Omega_{mt}) = \max_{\{l_t, s_t\} \in \mathcal{S}_t^m} V_t^f(l_t, s_t | \Omega_{mt})$$

These value functions appear in (11) and (27) that govern divorce and marriage decisions.

III.C. The Labor Market – Wages Offers and Job Offers

The wage offer functions have a standard Ben-Porath (1967), Mincer (1974) form:

(22) $$ln w_{et}^j = \omega_{1e}^j + \omega_{2e}^j X_t - \omega_{3e}^j X_t^2 + \varepsilon_{jt}^W \qquad for \qquad j = f, m$$

where X_t is work experience (years) and $e \in \{HSD, HSG, SC, CG, PC\}$ is education level. We let the wage function parameters $\{\omega_{ke}^j\}_{k=1,3}$ vary freely by education, gender, race/ethnicity and cohort. Thus, at a given education level, both starting wages and returns to experience may differ between males and females and between race/ethnic groups, capturing two potential dimensions of discrimination.[18] Our specification allows returns to experience to differ by education, as recent studies (e.g., Imai and Keane, 2004, Blundell et al., 2016) find greater experience returns for more educated workers. We let parameters vary by cohort to allow for changes in the wage structure over time, particularly changes in returns to education/experience and relative wages between men and women and between race/ethnic groups.

The error term ε_{jt}^{W} in equation (22) has a permanent/transitory structure:

$$\text{(23)} \qquad \varepsilon_{jt}^{W} = \mu_j^{W}(PE, PM) + \tilde{\varepsilon}_{jt}^{W} \qquad \text{where} \qquad \tilde{\varepsilon}_{jt}^{W} \sim iidN(0, \sigma_{\varepsilon}^{W})$$

The time-invariant error component μ_j^{W} is the person's skill endowment, as in Keane and Wolpin (1997). We allow for three initial skill endowment levels (low, medium, high). The probability a person is each type depends on parents' education (*PE*) and marital status (*PM*). This accounts for intergeneration transition of ability, and different levels of investment in child development among parents of different education levels or marital status.

An important discipline we impose on the model is that, like preferences, we do not allow the distribution of latent ability *conditional* on parents' education and marital status to differ by race/ethnicity or cohort or gender. The probability a person is each of the three skill types is determined by a multinomial logit (MNL) with the latent indices:

$$\text{(24)} \qquad \begin{aligned} \nu^{High} &= \eta_0^{H} + \eta_1^{H} \cdot PE + \eta_2^{H} \cdot PM + \epsilon^{H} \\ \nu^{Med} &= \eta_0^{M} + \eta_1^{M} \cdot PE + \eta_2^{H} \cdot PM + \epsilon^{M} \end{aligned}$$

where we normalize $\nu^{Low} = \epsilon^{L}$ for identification and where the ϵ vector is *iid* extreme value.

Now consider job offers: In each period, people who were unemployed at the start of the period ($h_{t-1} = 0$) may receive full- and/or part-time job offers probabilistically. Thus, their possible choice sets for hours are D_t = {0}, {0, 0.5}, {0,1} or {0, 0.5, 1}. The probabilities of getting a full-time offer and a part-time offer are each determined by a logit model of the form:

$$\text{(25)} \qquad P_j(k \in D_t) = \frac{exp(\phi_{j0k} + \phi_{j1k} e_t^{r} + \phi_{j2k} X_t + \phi_{j3k} H_t)}{1 + exp(\phi_{j0k} + \phi_{j1k} e_t^{r} + \phi_{j2k} X_t + \phi_{j3k} H_t)} \qquad \text{for} \quad k=1, 2$$

where k=1,2 denote full- and part-time, respectively. Here $e_t^{r} = 1, \ldots, 5$ corresponds to the five education levels in ascending order, X_t is work experience and H_t is health.[19] An employed

[18] Different intercepts or slopes between males and females may also capture that males and females of a given age and education are not perfect substitutes in production, causing rental rates on male/female labor to differ.
[19] By introducing job offer probabilities that depend on work experience we can capture the idea that women who leave the labor force (perhaps after marriage or childbirth) may have a difficult time obtaining job offers later, as their work experience lags behind that of women who continue working. Hence, the impact of not working on experience (and wages) can accumulate over time.

individual (h_{t-1}>0) has the option to keep his/her previous job, unless an exogenous separation occurs. The separation probability obeys a similar logit function that also depends on e_t^r, X_t and H_t.[20] We let the parameters of the logit models for job offers and separation probabilities differ freely by cohort and race/ethnicity (just like the wage function parameters).

III.D. Health Status

There are substantial disparities across race/ethnic groups in health and mortality in the US, so it is important for our model to account for these. We assume that health evolves over the life-cycle according to a two-state Markov chain, where $H_{jt} \in \{1,2\}$ indicate good and fair/poor, respectively. The transition probabilities differ by age, cohort and race/ethnicity. We assume health is an *exogenous* process, so it can be estimated outside the model.

Health plays several important roles in our model. For example, we require people to retire by age 65, but declining health may induce them to retire earlier, as health affects both tastes for work (eqn. 4b) and job offer probabilities (25). Health is also a dimension on which people sort in the marriage market (eqn. 6), and it shifts tastes for pregnancy (eqn. 7). Furthermore, as we assume health is not affected by employment, marriage or fertility decisions, it generates exogenous variation in these decisions (*given* our model).

III.E. The Marriage Market

The final component of the model is the marriage market. Single people may receive marriage offers, and they choose to become married if they draw a good enough match. To make this decision, they must compare the value of remaining single to the value of entering the married state. This section describes how the matching process works.

III.E.1. ***Marriage Offers***

At the start of a period a single individual may receive a marriage offer. Denote the probability of receiving an offer as $p_j^H(\Omega_{jt})$ for $j = f, m$. We assume the probability is given by a binomial logit model that depends on age and age-squared, whether a person is below 18, and whether a person is in school. The age effects differ by gender.

A marriage offer is characterized by a vector of attributes of a potential spouse, denoted by $\mathcal{M}_{jt}$. We assume marriage offers always come from a potential spouse of the same age (t). This is necessitated by technical issues that arise in solving the dynamic programming problem (see EKL (2019) for details). We do not think this assumption will have too great an effect on the results, because the large majority of married couples are in fact close in age.[21] It is

[20] A person may be involuntary unemployed due to exogenous separation or due to drawing an empty choice set (i.e., $D = \{0\}$). A person may be voluntarily unemployed due to quitting their prior job or rejecting all offers.
[21] For the cohorts of 1960-1980, the age gap between partners is below 5 years for 79% of all couples. It is below 7 years for 88% of couples, and below 10 years for 94% of couples.

convenient to describe the construction of marriage offers in three steps:

First, we draw the education of the potential spouse. We assume potential spouses have three possible education levels: high-school and below (HS, *ed = 0*), some college (SC, *ed = 1*) or college or above (C, *ed = 2*). The probability of receiving an offer from a potential spouse of the HS, SC or C type depends on a person's own education.[22]

Specifically, if the individual gets a marriage offer, we draw the potential partner's education using a multinomial logit (MNL) with the following latent indices:

$$
\begin{aligned}
v_{jt}^{C} &= \eta_{0j}^{C} + \eta_{1j}^{C} \cdot I[ed^m - ed^f = 2] + \eta_{2j}^{C} \cdot I[ed^m - ed^f = 1] + \epsilon_{jt}^{C} \\
v_{jt}^{SC} &= \eta_{0j}^{SC} + \eta_{1j}^{SC} \cdot I[ed^m - ed^f = 1] + \epsilon_{jt}^{SC}
\end{aligned} \tag{25}
$$

High school is the base case with $v_{jt}^{HS} = 0$. The parameters η govern the probability that a person (of given education) receives offers from potential partners with different education levels. The η reflect both the supply of potential partners and tastes for partners of different types.[23]

Rather than solve explicitly for marriage market equilibrium, we estimate parameters η that, when combined with the rest of our model, generate the observed distribution of match outcomes between types of partners.[24] Crucially, we let the η parameters differ by race/ethnicity and cohort. This captures different supplies of potential partners within each race/ethnic group and over time, as well as different and changing tastes for partners of different education levels.

Our approach allows us to side-step making explicit assumptions about intermarriage. We do this by searching for parameters η that enable us to match the frequencies with which both men and women in each race/ethnic group marry partners with each level of education, *irrespective* of the race/ethnic identify of the partners. For example, our estimation does not constrain the number of Hispanic women who marry college men to equal the number of married college-educated Hispanic men (or vice versa).[25]

[22] To simplify the MNL we combine the HSD and HSG levels into "HS," and the CG and PC levels into "C." Then, if a person draws "HS" we assign education level HSD or HSG to the potential partner according to the actual fraction in the data (by cohort and age). We do the same to convert "C" draws to CG and PC offers.

[23] For example, suppose there is a strong preference for partners with similar education. Then the chance a HSD woman receives an offer form a college educated man may be negligible regardless of the supply of college educated men.

[24] Our estimated η are therefore reduced form parameters that implicitly combine (i) structural parameters of preferences for different types of partners with (ii) endogenously determined supplies of partners. This approach has two key advantages: (1) It greatly simplifies estimation of the model relative to a case where we solve explicitly for the marriage market equilibrium, and (2) it allows us to avoid making detailed assumptions about how the marriage market works. The downside of course, is that we must assume the η are invariant to any policy experiments we may choose to consider.

[25] The rate of college education among Hispanic men is very low, so if marriage markets were segmented by race/ethnicity we would expect Hispanic women to receive a low rate of offers form college men. But our

Of course, substantial segregation of marriage markets along race/ethnic lines does exist, so we expect the η parameters to imply that black and Hispanic women have a much lower rate of receiving offers from college-educated men than white women (and to find similar differences for men). We find that such differences in marriage market prospects are important for explaining differences in behavior between race/ethnic groups.

Once the education is drawn, we calculate the potential work experience as the age of the individual minus his years of schooling minus 6.

Second, we draw the remaining elements of $\mathcal{M}_{jt}$. The five *observed* elements are drawn from the population distribution of all potential partners within a person's own age cell. These elements of $\mathcal{M}_{jt}$ are partner's health, number of children, PE, PM and lagged work. Their distributions are not conditional on un-observables, so we can obtain them from the raw data. We account for intermarriage by weighting the white, black, Hispanic matrixes according to the proportion of intermarriages in each group. However, inter-marriages are rare so the new matrixes are not very different from the unweighted.

Third, the four *unobserved* elements are drawn from their population distributions as specified in the model. These are the potential partner's tastes for leisure μ_{jt}, labor market ability μ_j^W, transitory wage shock $\tilde{\varepsilon}_{jt}^W$, and the taste for marriage, ε_t^M. The stochastic terms μ_{jt}, μ_j^W, $\tilde{\varepsilon}_{jt}^W$ and ε_t^M, are observed by both parties as part of the marriage offer. Both parties also understand which terms are permanent and which terms are only transitory.

Putting this all together, the marriage offer for a single female consists of the vector:

$$(26) \quad \mathcal{M}_{ft} = (E^m, X^m, H^m, N^m, PE^m, PM^m, h_{t-1}^m, \mu_m^l, \mu_m^W, \tilde{\varepsilon}_{mt}^W, \varepsilon_t^M)$$

Marriage offers for males ($\mathcal{M}_{mt}$) have an analogous form.

III.E.2. *Marriage Decisions*

Given a marriage offer $\mathcal{M}_{jt}$, a single person can construct the vector $(\Omega_{ft}, \Omega_{mt})$ that characterizes the state of the couple if they marry. That is, $(\Omega_{jt}, \mathcal{M}_{jt}) \rightarrow (\Omega_{ft}, \Omega_{mt})$ for $j=f,m$. The potential partner also knows $(\Omega_{ft}, \Omega_{mt})$. Both parties calculate the value of marriage, denoted by $V_t^{jM}(\Omega_{mt}, \Omega_{ft})$ for $j = f, m$ in equation (12). A marriage is formed if and only if:

$$(27) \quad V_t^{fM}(\Omega_{mt}, \Omega_{ft}) - \Delta(\mathrm{PM}_f) > V_t^f(\Omega_{ft}) \text{ and } V_t^{mM}(\Omega_{mt}, \Omega_{ft}) - \Delta(\mathrm{PM}_m) > V_t^m(\Omega_{mt})$$

Here $\Delta(\mathrm{PM}_j)$ is a fixed cost of marriage that we allow to depend on the marital status of the

estimated η parameters allow Hispanic women to receive offers from college-educated men at a higher rate – allowing implicitly for the possibility that they also receive offers from college educated white or black men.

parents of each partner. This allows for the possibility that there is intergenerational transmission in tastes for marriage.

If the pair decides to marry they proceed to make collective decisions about work and fertility as described in Section III.A. If the pair decides to remain single they make decisions about work, school and (for women) fertility as described in Section III.B.

III.F. Terminal Period and Retirement

The terminal period in the model is fixed at age 65, at which point everyone must retire. Of course, people can choose to stop working earlier if desired. By setting the terminal period at 65 we avoid the complications of modelling Social Security and the accumulation of retirement savings.[26] To reduce computational burden, we assume the terminal value function $V_{T+1}^{j}(\Omega_{jT})$ at *T*=65 is a simple function of state variables. Thus, the terminal value function accounts for retirement savings in a reduced form way.

III. G. Summary

This completes the exposition of the model. Note that the choice set of a married couple is $\{l_t^m, l_t^f, p_t\}$, as well as whether to stay married. The choice set of both single men and women includes work hours, school attendance, and whether to marry. Single women also make choices about pregnancy and, if eligible, welfare participation. We have stochastic terms in tastes for leisure, pregnancy, school, marriage and welfare. As there is a stochastic term associated with every choice, the model will produce a non-degenerate likelihood.

Finally, it is useful to discuss the mechanisms that drive marriage in the model. First, there is the public good nature of couples' consumption. Each partner consumes 71% of total household consumption. Thus, marriage may increase consumption of both parties. However, if a person has much higher earning capacity than their potential spouse, his/her consumption may fall with marriage. Thus, a person with higher earning capacity will tend to have a higher reservation earning capacity for their spouse (*ceteris paribus*). This occurs for two reasons: (i) the higher a person's income, the higher the income of his/her spouse must be to prevent consumption from falling after marriage, and (ii) a person with higher earning capacity will have a higher probability that his/her offers are accepted, enabling them to be more selective. These mechanisms help to generate assortative mating in the model.

Second, people get utility from marriage itself (see eqn. (6). But the magnitude of this

[26] For our five cohorts, the "normal" age for claiming Social Security (SS) benefits was gradually increased from 65 to 67. But workers can also opt to receive "early" SS benefits at age 62, subject to a penalty in the form of a reduced benefit level. To avoid having to model this decision, we chose to fit our model only to data up through age 61. By setting the terminal value function at age 65, we implicitly assume away the early SS option.

utility differs across potential spouses. This gives the individual an incentive to search over marriage offers (i.e., an option value for waiting). In (6) we specified that people prefer spouses with similar education. This helps to generate assortative mating on education. Interestingly, there is a trade-off between θ_t^M and w_t^j (as noted earlier). This means a person is willing to accept a larger education difference if it is compensated by a higher wage.

IV. Solution, Estimation and Fit

We back-solve the model from age 65 to 17, assuming a terminal value function at age 65, as in EKL.[27] We stress that we solve the dynamic programming (DP) problems of *individual* males and females. The individual solves his/her problem understanding the probabilities of marriage and divorce, as well as how decisions are made by couples. The state space Ω_{jt} of our DP problem is discrete. The state variables are marital status, number of children, taste for leisure (which is shifted by arrival of children), education, experience, age, the lagged choice, latent skill type and parental background. The state of a married person also includes the state variables of the spouse, and the stochastically evolving match quality (which determines the flow utility from marriage).

Starting at age 17, a single person makes choices taking into account how these affect his/her marriage market opportunities. This requires predicting the distribution of potential spouse *conditional* on own age/education in *future* periods. We assume people have perfect foresight about these distributions. This is imposed implicitly in estimation by: (i) using the same offer distribution that we fit within the estimation as the distribution that people use to forecast offers, and (ii) requiring that the model based on this assumed distribution provide a good fit to realized assortative mating patterns. This circumvents the need to solve for the spouse offer distribution as an endogenous object that emerges from the marriage market equilibrium, which would be infeasible in a dynamic model with many state variables.

We estimate our model using repeated cross-section data from the CPS, as in EKL. Specifically, we use annual data from the March CPS for the period 1962 to 2021. The sample is divided into three separate ethnic groups: Whites, Blacks and Hispanics. We consider only civilian non-institutionalized adults over the age of 16. We divide the sample into three cohorts born within two years of the reference years 1960, 1970 and 1980. For out-of-sample validation and forecasting we also use data for cohorts born within two years of 1990, 2000 and 2010.

The estimation method is the Method of Simulated Moments (MSM), as proposed by McFadden (1989) and Pakes and Pollard (1989). We use simulated life-cycle histories from

[27] We assume everyone lives to at least age 65, and abstract from marriages that end via the death of a spouse.

the model to generate statistics that we match to the data to form moments (currently 5000 men and women for each cohort). The moments we attempt to match include moments by age groups on completed schooling, employment rates, annual wages, marriage rates, assortative mating and fertility, for men and women of each ethnic group in the 1960, 70 and 80 cohorts. There are a total of 2772 moments that we list in detail in Appendix F.[28] The identification of the model is discussed in detail in EKL.

We estimate the model jointly for all the three cohorts and three ethnic groups, because we assume that 87 model parameters are common across cohorts and ethnic groups. These common parameters include 47 preference parameters, as well as 24 terminal value function parameters, 8 random shock variances, 6 parameters that describe how ability depends on parental background, and 2 parameters that govern the distribution of alimony. The 47 preference parameters include tastes for leisure (12), school (7), kids (8), pregnancy (5), marriage (8), welfare (2), divorce (4) and the CRRA in consumption.

We observe cohorts for different age ranges; the last observation is age 61 for the 1960, 51 for the 1970 cohort and 41 for the 1980 cohort. In simulating the model, we use the actual rates of mother's college graduates that were born in the US and single mothers for each cohort and ethnic group as indicated by Table 1. It should be noted that we assume that mothers of whites and blacks are born in the US and that Hispanic mothers that were not born in the US are equivalent to mothers without college degree. We use these to derive the labor market skill endowments and taste for college as explained above.

We allow wage offer functions, job offer functions and marriage offer functions to differ by ethnic group, cohort and gender, which generates a large number of parameters. For each cohort/ethnic group/gender (3*3*2) group we have 15 parameters for wage offers and 12 parameters for job offers, giving a total of 27 labor market parameters per group (giving 486 labor market parameters in total over all 18 groups). For each cohort/ethnic group we have 15 parameters for the marriage market (giving 135 in total for all 9 groups). So, the model has 708 parameters in total for the 18 cohort/ethnic group/gender. This is an average of about 39 1/3 parameters per group. Thus, the total number of parameters is not very large considering the large number of statistics being fit (for each of 18 different cohort/ethic/gender groups).

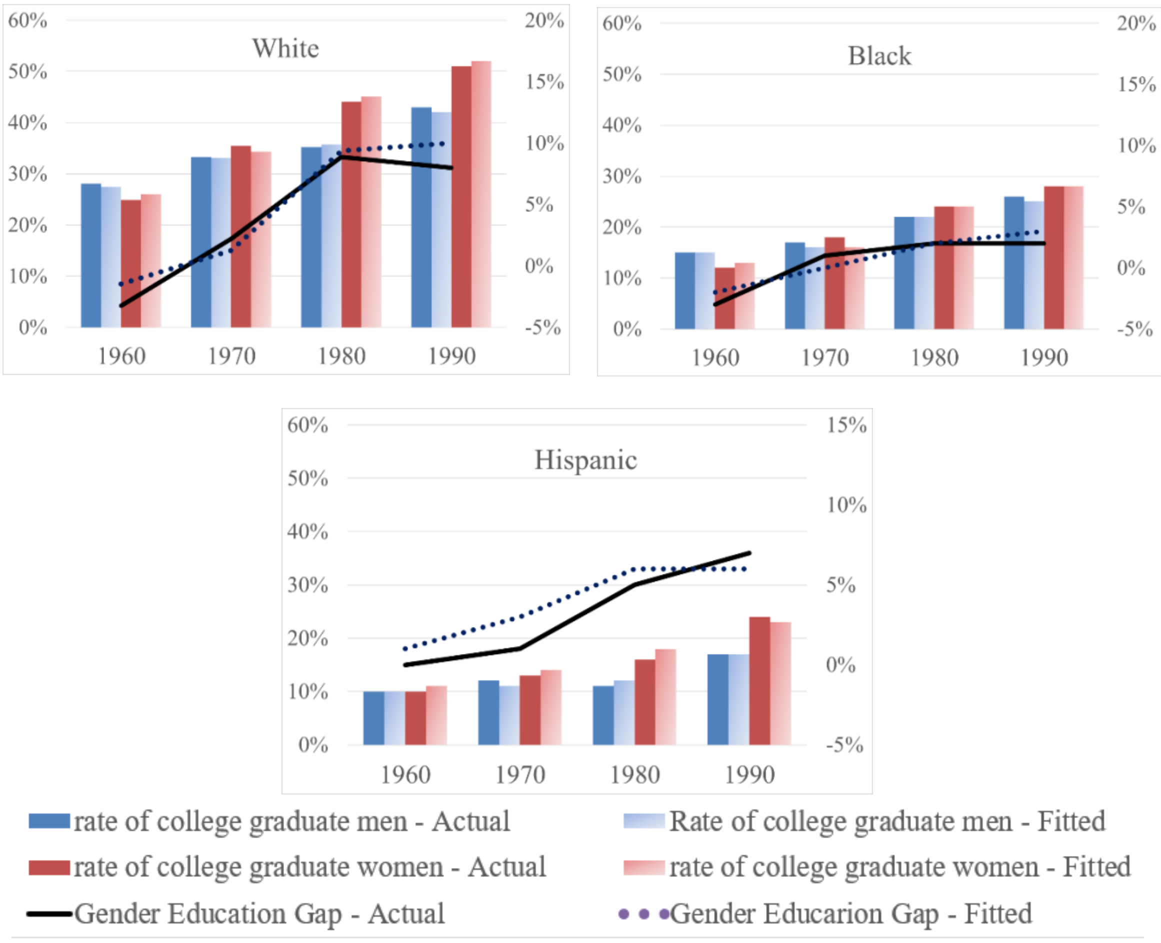
Our assumption that preferences parameters are common across all cohorts and ethnic groups is an important discipline imposed on the model. An important result is that we can achieve a very good fit to a large range of behaviours and outcomes (education, labor supply,

[28] This includes 388 moments for each of the three ethnic groups in the 1960 cohort, 308 moments for each group in the 1970 cohort, and 228 moments for each group in the 1980 cohort.

wages, marriage, fertility, welfare participation, assortative mating) for all cohorts and ethnic groups assuming common preferences – see fit statistics in Appendix H.

The main focus of the paper is on college graduation, so we display the fit to college graduation rates in Figure 3. Our model provides an excellent fit to the college graduation rates for each gender/ethnic group/cohort. It succeeds in predicting the college graduation rate gaps between groups, the growing graduation rates across cohorts, and the substantial increase in college attendance of women relative to men in 1970 and 1980 cohorts.

Figure 3: Fit to College Graduation Rate by Gender and Cohort and Ethnic Group



As we are able to fit the education differences across gender/ethnic/cohort groups almost exactly using fixed preferences, our model decomposes these differences completely into parts due to each of the exogenous factors in the model (i.e., differences in parental background, labor market constraints, marriage market constraints, etc.).

Our model also gives an excellent fit to many other statistics (employment rates, assortative mating, fertility and wages). Figure 4 highlights one particularly important statistic, the employment rate of women at ages 32-36. As we noted in the introduction, employment rates of women in this age range grew rapidly from the ’25 to ’60 birth cohorts, and this was a key reason for their increase in education. But employment rates then stabilized. A particular challenge for our model is to explain why education of women continued to grow rapidly, even after employment rates stopped growing. As we see in Figure 4, employment rates of white women with at least some college education were very stable from the ’60 to ’80 cohorts, while those of women with high school of less education actually fell. The model captures these patterns well, even with fixed preferences across cohorts.

Figure 4: Fit to Employment Rates conditional on Education, Women, Age 32-36

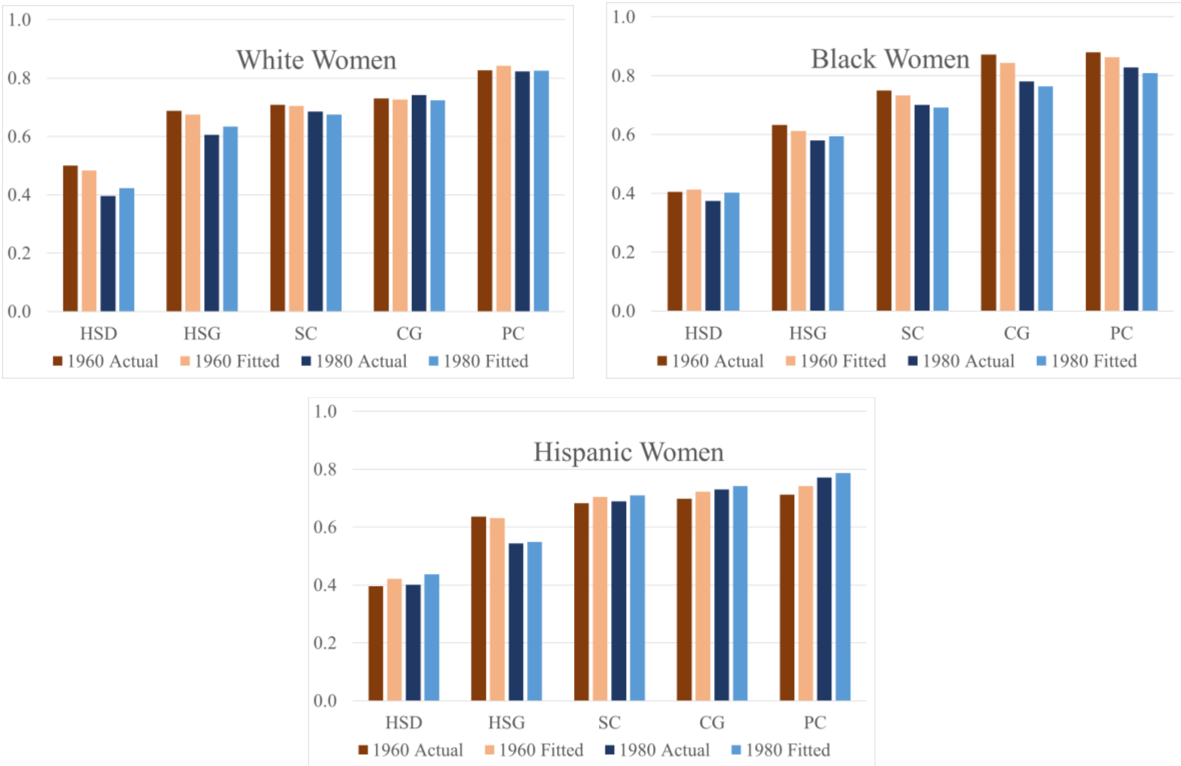


Figure 4 shows the patterns are different for Black and Hispanic women, and our model captures these differences (assuming common preferences across ethnic groups). For Black women employment fell slightly in all education groups. For Hispanic women increased among the college educated, and dropped for high school graduates. It is rather remarkable how well the model fits all these patterns using only the three exogenous factors (parental background, marriage market, labor market). This gives us some confidence in using the model to provide explanations for the key patterns in college graduation rates that we set out in the introduction.

V. Parameter Estimates

We find that having a college educated mother substantially increases daughter's tastes for college. Our estimates imply that utility from school is higher for both men and women whose mother graduated from college, but this effect is about 60% greater for women than for men. We also find that both men and women whose parents were married get more utility from marriage. As a result, women with college educated mothers are much more likely to graduate from college themselves, and both men and women whose parents were married are more likely to decide to marry as well. This creates important intergenerational links in both college attendance and marriage. Our model implies that differences in parental background are a key factor explaining differences in college attendance across ethnic groups, increases in education over time, and the growing gender education gap.

Our model also allows wage offer functions and job offer functions to differ by ethnic group, gender and cohort. Cohorts face different offer wage functions at a point in time because skills and training differ in important ways by cohort. For example, members of the 1980 birth cohort tend to have better computer skills (and better computers and machines at work) than members of the 1960 birth cohort, due to improvements in the technological environment over time. Returns to college will differ by cohort at a point in time because the set of skills taught in college differs by cohort, and skill-biased technical change will affect different cohorts differently because cohorts have different training. This differs from the usual treatment of technical change, which views it as shifting returns to college for all cohorts in a similar way at a point in time. We find important differences in the wage structure across cohorts. These differences in wage offer functions and job offer functions are important in explaining differences in education across ethnic, gender and cohort groups.

Our estimates uncover several interesting differences in wage offer functions by gender, ethnicity and cohort. Table 3 presents selected offer wage function estimates for Whites. As we see, starting wages of college graduate white women and men both improved substantially from the 1960 to 1980 birth cohort. But women still lag behind men by .12 log points. Experience returns of white women in the 1960 birth cohort were smaller than those for white men, but by the 1980 cohort they have almost caught up.

In the Appendix, we show that starting wages of Black and Hispanic women almost caught up to those for white women, but their experience returns did not. Differences in starting wages between White, Black and Hispanics men are modest, but differences in experience returns are substantial. Those of Hispanic men improved modestly but those of Black men stagnated. Thus experience returns of Whites (both men and women) are well above those of Blacks and Hispanics.

Table 3: Selected Offer Wage Function Estimates, Whites

		Intercept			Experience Coefficient		
		1960	**1980**		**1960**	**1980**	
Men	**High School**	10.00	10.05	+.05	.058	.074	+.016
	College	10.42	10.60	+.18	.076	.087	+.011
		+.42	+.55		+.018	+.013	
Women	**High School**	9.76	9.81	+.05	.041	.065	+.024
	College	10.30	10.48	+.18	.053	.080	+.027
		+.54	+.67		+.012	+.015	

Table 4 reports selected job offer rates. Job offer probabilities of white women were well below those of white men in the 1960 cohort. Since then, rates for women improved while those of men are almost unchanged, so by the 1980 cohort offer rates for white women are almost as high as those for white men. Job offer probabilities of Black men and women improved slightly but remain inferior to those of whites. Interestingly, job offer probabilities of Hispanic men now look similar to those of white men, but those of Hispanic women are in between those of Black and White women.

Table 4: Job offer rate by race and cohort, by experience and education

	Women 1960		Men 1960		women 1980		Men 1980	
	HSG	CG	HSG	CG	HSG	CG	HSG	CG
White								
EXP. = 3	42%	46%	52%	59%	48%	56%	51%	60%
EXP. = 5	44%	49%	56%	63%	51%	59%	55%	64%
EXP. = 10	51%	56%	66%	71%	57%	65%	65%	73%
Black								
EXP. = 3	30%	31%	34%	36%	34%	37%	39%	41%
EXP. = 5	31%	32%	35%	37%	36%	38%	41%	43%
EXP. = 10	33%	35%	38%	40%	39%	42%	46%	48%
Hispanics								
EXP. = 3	36%	41%	48%	53%	39%	46%	51%	58%
EXP. = 5	38%	43%	51%	57%	42%	49%	55%	62%
EXP. = 10	44%	49%	60%	65%	49%	57%	66%	72%

Table 5 reports selected job destruction rates. In the 1960 cohort job destruction rates of White women were higher than those of White men, but in the 1980 cohort they are almost identical (mostly because those of men got slightly worse). Job destruction rates for Blacks are well above those for Whites, and they have improved only slightly across cohorts. In contrast to Whites, the destruction rates of Blacks women are slightly better than those of Black men. The job destruction rates of Hispanics are slightly higher than those of Whites. Those of Hispanic men were unchanged, while those of Hispanic women improved slightly.

Table 5: Job destruction rate by race and cohort, by experience and education

	Women 1960		Men 1960		women 1960		Men 1960	
	HSG	CG	HSG	CG	HSG	CG	HSG	CG
White								
EXP. = 3	15%	12%	12%	10%	14%	11%	14%	12%
EXP. = 5	11%	9%	8%	6%	10%	8%	9%	8%
EXP. = 10	4%	4%	2%	2%	4%	3%	3%	3%
Black								
EXP. = 3	25%	23%	27%	25%	24%	22%	25%	23%
EXP. = 5	22%	21%	25%	23%	21%	19%	22%	20%
EXP. = 10	16%	15%	19%	18%	16%	14%	15%	14%
Hispanic								
EXP. = 3	17%	15%	15%	13%	17%	13%	15%	13%
EXP. = 5	13%	12%	11%	9%	13%	10%	11%	9%
EXP. = 10	7%	6%	5%	4%	6%	5%	4%	4%

A second important exogenous factor in our model is the marriage market. We allow marriage offer functions to vary in two key ways by cohorts and ethnic groups: First, we allow the probability of getting offers from potential spouses of different education levels to vary. But we find that this actually changed very little across cohorts, and differences across ethnic groups were stable. Thus, there was little change in the degree of assortative mating.

Second, we consider how the probability of getting marriage offers differs by age. As we see in Table 6, the probability a women can get marriage offers at older ages (especially 35 and 40) increased dramatically from the '60 to '80 cohort, and this was true for all ethnic groups. As we will see, this increased the incentive for women to graduate from college, as it became easier to delay marriage and fertility.

Table 6: Marriage Offer Probabilities by Age, Cohort and Race

	White		Black		Hispanic	
Age	1960	1980	1960	1980	1960	1980
25	24.0%	27.2%	16.3%	19.1%	22.0%	24.8%
30	12.4%	19.1%	9.5%	12.7%	12.8%	16.4%
35	4.7%	11.1%	4.5%	7.1%	5.8%	8.8%
40	1.3%	5.2%	1.7%	3.3%	2.0%	3.8%

The third and final exogenous factor in our model is parental background, which affects ability, tastes for school and tastes for marriage. The mapping from parental background to the low/medium/high ability types is invariant across cohort, gender and ethnicity. The estimated mapping implies, not surprisingly, that a person whose mother is a college graduate is more likely to be the high ability type. A person from a two-parent household is also more likely to be the high ability type, although this effect is not as strong as the effect of mother's education.

The mapping from mother's education to tastes for school is invariant to cohort and ethnicity, but importantly it differs by gender. Recall that we have:

(18) $\vartheta_{jt} = \vartheta_{0j} + TC \cdot I(E_t > HSG) + \vartheta_{1j} PE + \vartheta_{2j} \mu_j^W$ for j=m,f

We estimate the parameters ϑ_{1m} and ϑ_{1f} to be 257 and 402, respectively. Given our estimated CRRA utility function, this translates into a consumption equivalent of roughly $9,126 per year of school for men, and $20,786 for women. This could be interpreted as meaning people – especially women – whose mothers were college graduates enjoy school more, or get more utility from school because they are better at it (it is less work), or they simply value education more highly, or some combination of all three (our model cannot distinguish these stories). But the much large figure for women implies that mother's education has a bigger positive effect on education of daughters than sons.

Our estimates of parameter ϑ_{2j} imply that higher (labor market) ability types like school better, and this effect is stronger for men. In fact, our estimates imply that low and even medium ability men have a strong dislike for school. In contrast, medium ability women whose mothers did (did not) have a modest positive (negative) taste for school.

Finally, our estimates imply the fixed utility cost of marriage is 748 greater for people from single parent households, which translates into a consumption equivalent of roughly $65,753. It is worth noting that fixed costs of marriage are typically very large in models where agents search for spouses, as marital formation is not a common event. Thus, a large "love" shock is needed to overcome the large fixed cost and induce people to marry.

The impact of skill endowments and parental background on the value of school are illustrated in Table 7. The table reports the value of college for different types of agent in our model, focussing on Whites. We construct this table by running a counterfactual where we shut down the option to attend college. We then calculate the amount of initial assets that each type of agent must be given to compensate for the loss of the college option. This is the dollar amount the agent must be given to equalize their expected present value of lifetime utility under the baseline vs. the counterfactual. The values of college reported in Table 7 reflect all aspects of that value: labor market returns, marriage market returns and the utility from college attendance. We report values of college broken down by cohort, gender and type. The low skill endowment types in the model almost never go to college, so their option values are close to zero. Thus, we exclude them from the table.

Table 7: Value of College by Type, for Whites (PV in thousands $)

			1960					1980				
Skill endowment	Mother 's education	Mother's marital status	Type %	PV Women	CG rate	PV Men	CG rate	Type %	PV Women	CG rate	PV Men	CG rate
High	HS	Married	23.8	600	77.8	570	83.3	17.8	885	98.4	700	97.6
High	COL	Married	6.4	645	98.7	660	99.7	10.4	930	97.7	755	98.4
High	HS	Single	1.0	560	74.3	540	73.5	2.5	835	98.6	635	99.1
High	COL	Single	0.3	615	99.6	630	97.4	1.9	890	99.2	710	98.7
Medium	HS	Married	36.8	95	0.0	97	0.0	27.5	140	0.5	105	0.0
Medium	COL	Married	5.6	110	0.0	100	0.0	9.0	620	95.6	220	26.3
Medium	HS	Single	1.9	11	0.0	5	0.0	4.9	75	0.0	15	0.0
Medium	COL	Single	0.4	75	0.0	60	0.0	2.0	580	96.2	200	19.7

In the 1960 cohort, values of college were very similar for men and women. For instance, for a high-skill agent whose mother graduated from college and who came from a two-parent household, the value of college was $660k for men and $645k for women. The values are similar for other types as well (i.e., high and medium ability, mother did or did not graduate from college, dual or single parent household). The similarity of college values (for all types) explains why men and women graduated at similar rates in the 1960 cohort.

Amongst the high skill types whose mothers were college graduates, almost 100% graduated from college themselves. If the mother did not graduate from college, the college values fall by $90k for men and $45k for women.[29] In this group, about three-quarters of women and 80% of men graduate from college.[30] And, if the person was from a single parent household, the value of college drops by about $30k for both men and women, and graduation rates drop about 10% for men (but not for women). Note that the values of college for medium skill types are much smaller ($110k or less). Almost none of the medium ability men or women graduated from college in the 1960 cohort.

In the 1980 birth cohort, values of college are much higher, particularly for women. For instance, for a high-skill agent whose mother graduated from college and who came from a two-parent household, the value was $755k for men and $930k for women. Compared to 1960, that is a $95k increase for men but a $285k increase for women. Approximately three-

[29] It is surprising the drop is greater for men, as mother's college has a bigger impact on daughter's utility from school. But the labor market returns to college are greater for men. So an increase in the utility "cost" of college that reduces the rate of college attendance causes a bigger drop in present value of lifetime earnings of men.
[30] The mean value of college for high skill women from dual parent households with graduate mothers was $600k in the 1960 cohort. This value is calculated ex ante at age 25. Despite the positive ex ante value, about 75% of these type agents do not attend college due to adverse draws of the stochastic terms in the model.

quarters of high skill women whose mothers did not graduate from college graduated from college in the 1960 cohort, but in the 1980 cohort it is close to 100%, reflecting that their value of college increased from $600k to $885k.

The biggest change in the 1980 cohort is among medium skill women whose mother's graduated from college (5.9% of population in 1960, and 11% in 1980). Their value of college increased by roughly $500k from 1960 to 1980, due to the improvement in their labor market returns. This caused their college graduation rate to increase from near zero to above 95% For medium skill men whose mother's graduated from college, the graduation rate in the 1980 cohort was only 20 to 25%. The much higher graduation rate of this group of women vs. men accounts for the higher graduation rate of women overall. As we saw earlier, our estimates imply that medium skill women like school a lot more than men.

Note that the college graduation rate of women increased from 25.9% in the 1960 cohort to 44.9% in the 1980 cohort. In our model low, medium and high skill refer to the bottom, middle and top third of the labor market skill distribution (i.e., low medium and high intercepts in the offer wage function), defined within each gender/cohort/ethic group. So by construction, if the college graduation rate of women is 44.9% it must be that a substantial number of medium skill women are graduating from college.

Two cautions are in order in interpreting this result: First, we cannot infer that women attending college in the 1980 cohort were on average less skilled than women in the 1960 cohort, as absolute skill levels of medium and high skill types may have changed over time. Second, we cannot infer that women graduates on average are less skilled than male graduates, as types are defined *within* genders. Furthermore, labor market skill is only one dimension of skill, and the greater utility from college of women may reflect that they are more skilled at studying and learning college material.

Finally, In Table 8 we show how labor market skill endowment type proportions vary by mother's college and across cohorts. Notice how, as mothers become more educated, the proportion of agents who are high skill increases.

Table 8: Skill Type Proportions by Cohort, Whites

	Mother no college			Mother college			ALL		
	low	medium	high	low	medium	high	low	medium	high
1960	22.5%	38.7%	24.8%	1.3%	5.9%	6.8%	23.8%	44.6%	31.6%
1970	20.8%	33.8%	21.5%	2.4%	10.1%	11.5%	23.1%	43.9%	33.0%
1980	21.3%	32.4%	20.3%	2.8%	11.0%	12.3%	24.0%	43.4%	32.6%
1990	18.9%	26.6%	16.5%	4.4%	16.0%	17.6%	23.3%	42.6%	34.1%
2000	16.2%	23.9%	14.9%	4.9%	19.0%	21.1%	21.1%	42.9%	36.0%
2010	14.4%	21.3%	13.3%	5.6%	21.5%	23.9%	20.0%	42.8%	37.2%

VI. Explaining Gender, Cohort and Ethnic Gaps in College Attainment

In the following Sections we use the model to address the following three questions: (A) Why are do recent cohorts of women get more education than men?, (B) Why has the college graduation rate increased across cohorts? And why has it increased more for women?, (C) What explains gaps in college education between Whites, Blacks and Hispanics?

VI.A. Why are Recent Cohorts of Women getting more Education than Men?

First, we ask why women in the 1980 cohort graduated from college at a much higher rate than men – e.g., a gap of 44.9% vs. 35.9% for Whites. We address this question via counterfactual experiments where we equalize, in turn, the labor market opportunities and preferences for school of women and men. These results are reported in Table 9.

A striking result on Table 9 is that the *labor market* returns to college are actually greater for men than for women. If we give white women in the 1980 cohort the same wage offer function as men, their college graduation rate increases substantially from 44.9% to 51.1%.[31] Thus, if women had the same labor market opportunities as men, their college graduation rate would increase even further, and the gender education gap would widen dramatically! The gender gap in college graduation does not arise because women have a higher labor market return to college than men.

Table 9: Explaining Gender Differences in College Graduation

	White	**Black**	**Hispanic**
Fitted 1980 women's college rate	44.9%	24.5%	17.9%
Men's labor market parameters	51.7%	28.5%	19.7%
Men's utility from schooling parameters	34.4%	21.6%	17.4%
Both	**41.6%**	**26.4%**	**19.4%**

It is interesting to examine our offer wage function estimates in light of this result. According to our estimates (see Table 3), having a college degree vs. a high school degree increases the log wage function intercept for white women in the 1980 cohort from 9.810 to 10.476. This 0.666 log point difference implies the starting wage for college women is exp(.666)-1 = 95% higher than that for high school women. In contrast, for white men in the 1980 cohort we have instead a 10.595 - 10.051 = 0.544 log point gap, which means the starting wage for college men is exp(.544)-1 = 72% higher than for high school men.

[31] If we also give women the same job offer probabilities as men, it increases their college graduation rate only very slightly more, to 51.7%. This is because job offer probabilities for white men and women were very similar in the 1980 cohort.

Thus, the so-called "college premium" for women is about 23% higher than for men. This 23% gap might be taken to indicate that the labor market return to college is greater for women than for men. Just looking at the starting wage ignores differences in wage growth, but factoring in wage growth with experience doesn't change the calculation in any important way. Our estimates imply college educated men and women enjoy faster wage growth with experience than the high school educated, but the gap is similar for men and women, so the 23% gap is quite stable at all levels of experience. This may appear to contradict the finding from our counterfactual that the labor market return to college is greater for men. But the apparent contradiction only arises because economists have often tended to conflate the "college wage premium" with the actual labor market return to college education.

The reason the labor market return to college is greater for men is that, while women get a larger *percentage* gain in wages than men, men get a larger *absolute* gain in wages – because the high school men have higher earnings than the high school women. On top of that, the men have higher employment rates than women, which also causes their *absolute* increase in earnings to be greater. Of course, economic agents care about the absolute gains in earnings that result from college when calculating the return to college – they do not care about percentage gains in wages (except as far as these influence earnings).

So why do women go to college more than men? If we give white women the men's tastes for college their college graduation rate drops by 10.5 percentage points, from 44.9% to 34.4%. This is because our estimates imply that women get substantially more utility from school than men. Thus, it is not a higher return to college, but rather a greater taste for college attendance that causes white women in the 1980 cohort to graduate from college at a much higher rate than men. Given the structure of our model, this may subsume a number of factors: For example, woman may be better at studying (or like it more), so they get less disutility from putting in effort at college. Or they may place greater value on learning for its own sake. Or they may simply get more utility from social activities at college.

The bottom row of Table 9 reports a counterfactual where we give white women both the labor market constraints and tastes for school of white men. Their college graduation rate then drops from 44.9% to 41.6%. Thus, the drop due to tastes outweighs the increase due to better wage offers.[32] Nevertheless, this 41.6% graduation rate is still 5.7 points higher than the

[32] In the bottom panel of Table 5 we perform the reverse experiment where we give men the labor market constraints and tastes for school of women. As expected, giving men the wage and job offers of women reduces their college graduation rate (from 35.9% to 31.1%), and giving them women's tastes for school raises their college graduation rate (to 41.7%). Interesting, doing both increases their graduation rate further to 43.6%. This illustrates an interesting interaction effect: When taste for school is very strong, reducing labor market

35.9% rate of men. The 5.7 percentage point gap that remains reflects other factors that create different incentives for men vs. women to attend college, even after we equalize tastes and labor market opportunities. The gap reflects the fact that marriage market returns to college are greater for women than for men. In particular, as women are likely to have children and spend time out of the labor force, their gain from finding a college educated (high income) husband exceeds a male's gain from finding a college educated wife.

VI.A.1. Does the Explanation of the Gender Gap differ by Ethnic Group?

Consider next the results for Blacks. According to our estimates, college leads to larger wages gains for Black men than Black women. And job offer probabilities are slightly better for Black men. Thus, when we give Black women the wage offers and job offers of men their college graduation rate increases from 24.5 to 28.5. Again, the labor market returns to college are greater for men than women. If we give Black women the same tastes for school as men their college graduation rate drops to 21.6%. So again, Black women derive more utility from school attendance than Black men.

If we give Black women *both* the labor market opportunities and tastes for school of Black men, their college graduation rate *increases* from 24.5% to 26.4%. This contrasts to white women, whose college graduation rate dropped in this experiment. The difference arises because Black women's taste for school is not as strong as White women's. Recall that tastes for school depend on mother's education, and the effect is stronger for women. Black women's taste for school is less strong than white women's (on average) because their mothers are less educated.

When we give Black women *both* the labor market opportunities and tastes for school of Black men, their college graduation rate of 26.4% exceeds the 22.6% rate of Black men. But in contrast to whites, the marriage market does not provide a strong incentive for Black women to attend college, because their marriage probability after college is very low (So going to college does not do much to raise the chance of an offer from a college man). Instead, it is children that make the difference: Black women have an incentive to get more education than Black men so they can afford to raise children as singles.

opportunities increases college attendance, as agents often choose to go to school (rather than stay at home) when they get poor wage offers (or no offers).

VI.B. Why has the College Graduation Rate Increased? Especially for Women?

In the 1960 cohort, men and women graduated from college at roughly the same rate, and this was true for Whites, Blacks and Hispanics. But in the 1980 cohort, women graduated from college at a much higher rate than men, and this was true for all groups. In this section we ask the questions: "Why has the college graduation rate increased?" and: "Why has the gender gap in college graduation *increased* over time?" An important point of our paper is that the latter question is very different from the question we posed in the previous section, which was: "Why is the college graduation rate of women *currently* higher than that of men?" And as we'll see, the answers are very different.

In the 1960 cohort the college graduation rates of White women and men were 26% and 27%, respectively. But in the 1980 cohort these rates increased to 45% and 36%. Thus, the college graduation rate of men increased by a substantial 9 percentage points, but the rate for women increased by a staggering 19 percentage points, opening up a 9 point gender gap. The patterns are similar for Blacks and Hispanics (see Figure 1).[33] Here we ask what exogenous factors drove this huge increase in college graduation, and the opening of the gender gap.

We address these questions by analysing the marginal contributions of the three main exogenous processes (family background, labor market and marriage market) to changes in college graduation rates. Table 10 reports the marginal contribution of each factor to the change in the college graduation rate from the 1960 to the 1980 cohort. Recall that in this analysis we assume preferences are fixed across cohorts and ethnic groups – i.e., we only allow preferences to differ between men and women. Note that mother's marital status affects both ability and taste for marriage. We decided to allocate the first effect to the parental background category, and the second to the marriage market category.

Parental background: Consider first the results for Whites. Our results imply changes in parental background, which affects the labor market skill endowment and taste for school, caused the college graduation rate of White women to increase by 5 percentage points between the 1960 and 1980 cohorts. Changes in parental background also caused the college graduation rate of White men to increase, but by a much smaller 1.4 percentage points.

Why is the effect of parental background so much greater for women? As we discussed in Section V having a college educated mother increases tastes for school, and this effect is much stronger for daughters. Between the 1960 and 1980 birth cohorts, the college graduation

[33] For Blacks the increases for men and women, were is 8% and 13%, while for Hispanics the increases were 2% and 8%, respectively.

rate of mothers increased from 14% to 24% (see Table 1), and this increased tastes for schooling, especially for daughters. Having a college graduate mother also increases the probability a person is the high skill type, but it does so equally for men and women.

Table 10: Factors Driving Increasing Education Across Cohorts

	White		Black		Hispanic	
	Men	**Women**	**Men**	**Women**	**Men**	**Women**
College Rate Fitted - 1960	**27.0%**	**25.9%**	**14.7%**	**11.6%**	**10.3%**	**11.0%**
1) Family Background - A- mother's education	0.019	0.047	0.020	0.039	0.010	0.025
1) Family Background - B - mother's marital status	-0.001	-0.001	-0.006	-0.002	-0.007	-0.004
1) Family Background - B1 - mother's marital status (only ability)	-0.007	-0.004	-0.008	-0.004	-0.005	-0.003
1) Family Background - A+B1+C	**0.014**	**0.050**	**0.015**	**0.038**	**0.006**	**0.026**
2) Taxes (brackets, deductions, exemptions, no change in EITC)	0.002	0.007	0.000	0.004	0.001	0.003
2) Welfare rules (AFDC, EITC, welfare reform 1996: time limits+ "stigma")	0.001	0.008	0.003	0.023	0.001	0.017
2) Job offer function - D	0.022	0.048	0.028	0.056	0.009	0.012
2) Wage offer function - E	0.050	0.077	0.040	0.059	0.014	0.031
2) Labor Market - D + E + Taxes and Welfare	**0.054**	**0.104**	**0.057**	**0.085**	**0.015**	**0.041**
3) Marriage Market - B2 - mother's marital status (only marriage cost)	0.001	0.007	0.000	0.008	-0.006	-0.002
3) Marriage Market - F – change in education distribution	0.011	0.010	0.003	0.008	0.004	0.007
3) Marriage Market - G – change in offer by age	0.017	0.034	0.005	0.013	0.003	0.013
3) Marriage Market - F +G + B2 (marriage cost)	**0.021**	**0.044**	**0.010**	**0.029**	**0.004**	**0.021**
College Rate Fitted - 1980	**35.9%**	**44.9%**	**22.6%**	**24.5%**	**12.2%**	**17.9%**

The mother's marital status is a second dimension of parental background. It affects both the probability a person is high ability and the taste for marriage. But as we see in Table 6 these effects are very small compared to the effects of mother's education.[34]

The impacts of mother's education are similar for Blacks and Hispanics. For Blacks the fraction of college graduate mothers increased from 6% to 13%, and for Hispanics it increased from 7% to 11%, while for Hispanics the percentage of native born mothers increased from 42% to 52% (Recall that for Hispanics we assume that only having a native born college educated mother increases taste for school). Hence, changes in parental background caused the college graduation rate of Black women to increase by 3.8 pp, compared to 1.5 pp for Black men, while causing the college graduation rate of Hispanic women to increase by 2.6 pp, compared to 0.5 pp for Black men.

Labor Market: Next we consider the impact of changes in the labor market (offer wages and job offer probabilities, as well as taxes and welfare rules). As we see in Table 10, our results imply that changes in labor market opportunities caused the college graduation rate of White women to increase by 10.4 percentage points between the 1960 and 1980 cohorts.

[34] The rate of single parent households increased substantially between the 1960 and 1980 cohorts (from 6.4% to 11.8% for whites), and this had two opposite effects on college: It lowered ability, causing college to fall, but it also reduced tastes for marriage, which raised college education. When women know they are less likely to be married it creates an incentive to get more education.

Changes in labor market opportunities also caused the college graduation rate of White men to increase, but by a much smaller 5.4 percentage points.

Changes in taxes and welfare rules led to small increases in education for women, but negligible changes for men. For both men and women it is changes in the job offer probabilities and especially offer wages that are the main factors. So why did the labor market have a bigger effect on women?

As we saw in Section V, Table 3, both starting wages and experience returns for white male and female college graduates improved between the 1960 and 1980 birth cohorts. The college vs. high school wage premium for men increased from .42 to .54 log points, while for women it increased from .54 to .67, so for both men and women it increased by .13 log points. But the experience returns for college graduate women improved much more than for men (i.e., as we see in Table 3, for college women the experience coefficient increased by .027, compared to an increase of .011 for men). Experience returns of white college women in the 1960 birth cohort were much smaller than those for white college men, but by the 1980 cohort they have almost caught up. In addition, job offer probabilities of white women improved across the three cohorts, so by the 1980 cohort they are very similar to those for white men. These improvements in experience returns and job offer probabilities explain why changes in labor market opportunities led to large increases in education for women than for men.

If we look at Blacks the story is similar. Changes in labor market opportunities caused the college graduation rate of Black women to increase by 8.5 percentage points compared to 5.7 points for Black men. For Hispanics effects are smaller but still much larger for women: the figures are 4.1 pp for women and only 1.5 pp for men. The very low college graduation rate of Hispanic men, as well as its very slow growth from the 1960 to 1980 cohorts, is notable. According to our estimates, the college vs. high school wage premium for Hispanic men was only .46 log points in the 1980 cohort, and this is the lowest of any group. It compares to .54 log points for both black and white men, .67 for white women, and .55 for Hispanic women. Black women at .47 is the next smallest.

A key reason that college grew less for Black and Hispanic women than for white women is that experience returns for Black and Hispanic college women grew much less than for white women. The same was also true for Black and Hispanic men relative to white men.

Marriage Market: Finally we look at the contribution of the marriage market. As we see in Table 10, our results imply changes in marriage market opportunities caused the college graduation rate of White women to increase by 4.4 percentage points between the 1960 and 1980 cohorts, compared to a 2.1 pp increase for men. So all three factors (parental background, labor market, marriage market) caused the college graduation rate of White women to increase roughly twice as much as that of men. Table 10 decomposes the marriage market effect into that due to changes in assortative mating vs. changes in the probability of marriage offers by age. We see that the increase in the probability of getting marriage offers at older ages is the larger factor. This alone causes the graduation rate of White women to increase by 3.4 pp. College became more attractive for women because college attendance does not crowd out opportunities for marriage (and fertility) to the extent that it did in the past. The results for Blacks and Hispanics are similar, except the effects are more modest than for whites.

One popular explanation for the increase in women's education is increased returns to college in the marriage market, due to an increase in assortative mating. So it is interesting to explore further why we find changes in assortative mating are <u>not</u> a very important factor:

Table 11 reports assortative mating patterns for the '60 to '80 birth cohorts for whites. Clearly the fraction of couples where both spouses marriages are college graduates increased substantially (from 18.9% to 35.0% of all marriages). But this 85% increase in the prevalence of college couples is roughly consistent with what we would expect given the increase in college education across cohorts (i.e., the fraction of college graduate women increased 73% from 26.3% to 45.5%). The degree of assortative mating did not change to any appreciable degree. In fact, the conditional probability of a college women matching with a college man fell very slightly from 69% to 68%.

Table 11: Assortative Mating Patterns by Cohort, Whites only

1960		HUSBANDS		
		HSD + HSG	SC	CG + PC
WIVES	HSD +HSG	29.5%	9.3%	4.4%
	SC	9.5%	12.3%	7.7%
	CG + PC	3.3%	5.3%	18.9%

1970		HUSBANDS		
		HSD + HSG	SC	CG + PC
WIVES	HSD +HSG	19.3%	6.2%	2.9%
	SC	9.8%	13.0%	7.2%
	CG + PC	5.3%	8.1%	28.1%

1980		HUSBANDS		
		HSD + HSG	SC	CG + PC
WIVES	HSD +HSG	13.3%	4.7%	2.2%
	SC	9.0%	12.8%	6.4%
	CG + PC	6.0%	10.6%	35.0%

Finally, it is notable that if we sum the three marginal effects of parental background, labor market and marriage market, we get roughly the total increase in education, For example, for white women we have 26.0 + 5.0 + 10.4 + 4.4 = 46.2 compared to 45.0. This suggests that interaction effects among the three factors are modest.

VI.C. Explaining gaps in college graduation between Whites, Blacks and Hispanics

Here we focus on the 1980 cohort and assess what factors account for different college graduation rates across ethnic groups. The differences are substantial: the college graduation rates of White men and women were 35.9% and 44.9%. In contrast, as we see in Table 12, the rates for Black men and women were 22.6% and 24.5%, and those for Hispanics were 12.2% and 17.9%. We now examine the impact on college graduation rates of Blacks and Hispanics if we give them the same exogenous factors facing whites (parental background, labor market, marriage market). If all three are equated then college graduation rates are equalized, as these are the only ways the groups differ in our model. We assume all three ethnic groups have identical preferences, and of course they all face the same tax and welfare rules.

Table 12: Explaining College Gaps between Whites, Blacks and Hispanics, 1980 Cohort

	Black		**Hispanic**	
Fitted 1980 college rate	22.6%	24.5%	12.2%	17.9%
	Men	**Women**	**Men**	**Women**
1) Family Background - A- mother's education	0.033	0.055	0.027	0.051
1) Family Background - B - mother's marital status	0.017	-0.018	0.003	0.006
1) Family Background - B1 - mother's marital status (only ability)	0.014	0.018	0.002	0.004
1) Family Background - C - mother's immigration status			0.024	0.019
1) Family Background - A+B1+C	**0.042**	**0.074**	**0.047**	**0.069**
2) Job offer function - D	0.038	0.051	0.023	0.033
2) Wage offer function - E	0.029	0.038	0.111	0.103
2) Labor Market - D + E	**0.074**	**0.092**	**0.132**	**0.137**
3) Marriage Market - B2 - mother's marital status (only marriage cost)	-0.002	-0.021	0.003	0.008
3) Marriage Market - F – change in education distribution	0.003	0.007	0.009	0.009
3) Marriage Market - G – change in offer by age	0.002	0.009	0.012	0.013
3) Marriage Market - F +G + B2 (marriage cost)	**0.013**	**0.030**	**0.025**	**0.033**
White Fitted - 1980 college rate	35.9%	44.9%	35.9%	44.9%

In Table 12 we have split the impact of parents' education into two components: It's impact on skill endowments and taste for school, both of which enter in the top panel (family background). We also split the impact of parents' marital status into two components: its impact on skill endowments and its impact on fixed costs of marriage. We include the latter under marriage market changes. Thus, our decomposition has three main parts: (i) the effect of family background on skill endowments and tastes for school, (ii) effects of labor market opportunities (wage and job offers), and (iii) the effect of the marriage offer functions plus the effect of parents' marital status on fixed costs of marriage.

Parental Background: If we give Black women the same mother's education as Whites (i.e., we increase the mother's college graduation rate from 13% to 26%) we predict their college graduation rate would increase by 5.5 percentage points. The rate of single parent households was 61% for Blacks in the 1980 birth cohort, compared to only 18.7% for whites. This leads to lower ability and lower tastes for marriage. If we eliminate only the negative effect on ability, we predict the college graduation rate of Black women would increase by 1.8 points. The combined impact of equalizing mother's education and the rate of single parent households (ability effect only) raises Black women's college graduation rate by 7.4 points, eliminating 35% of the gap with Whites. If we look at Black men, we get an increase of 4.2 points, eliminating 32% of the gap with whites.

For Hispanics, the effect of equalizing mother's education is similar to Blacks. For Hispanics it is important to remember that we only count a mother as college educated if she was born in the US.[35] If we equalize mother's education, and assume all parents were born in the US, then the college graduation rate of Hispanic men and women increases by 2.8 and 5.1 points. If we also equalize the rate of single parent status (reducing it from 25.9% to 18.7%), the college graduation rate of Hispanic men and women increases by 4.7 and 6.9 points.

Labor Market: Next, we consider the impact of labor market opportunities, both the wage offer function and the job offer function. According to our estimates, Whites, Blacks and Hispanics in the 1980 cohort have fairly similar starting wages conditional on schooling, but the whites have much higher returns to experience. This is true for both men and women. Black men have inferior job offer functions to White and Hispanic men, both in terms of the offer probability for men with zero experience and the rate at which offer probabilities increase with experience. Black women also have inferior offer functions to White women in terms of both the offer probability for women with zero experience and the rate at which offer probabilities increase with experience. The Hispanic women are similar to the Blacks in terms of initial offer probabilities, but similar to Whites in how the probabilities increase with experience.

If we give Blacks the same labor market constraints as Whites it increases their rates of college graduation by 7.4 and 9.2 percentage points for men and women. This closes the college gap with whites by about 55% for men and 45% for women. Thus labor market opportunities explain roughly half the college gap between Blacks and Whites.

If we give Blacks the same offer wage functions as whites we get increases of 3 and 3.8 percentages points, while if we give then the same job offer rates we get increases of 3.8 and

[35] In 1980, 52% of Hispanic mothers were born in US, and their college rate is 11%, so we set the rate of college mothers to (.52)(.11) = 5.5%. In counterfactual A, we increase the mother's college rate from 5.5% to 26%, the rate for whites. In counterfactual C, where we assume all mothers are born in the US, we increased mothers college rate from 5.5% to 11%.

5.1 percentage points. So the difference in offer rates is actually more important. The overall results for Hispanics are similar, but for them the wage offer function matters much more than the job offer functions (as their baseline job offer functions are more similar to whites).

Marriage Market: Finally, we consider equalizing marriage market opportunities. Our estimates of the marriage offer functions imply that the probabilities of receiving offers from spouses of different types conditional on ones' own education does not differ very much across Whites, Blacks and Hispanics. [Assortative mating patterns do not differ much across these groups, conditional on marriage]. As a result, giving Blacks and Hispanics the offer functions of whites only modestly increase their education. We tend to see larger impacts when we equalize how offer probabilities vary with age. Whites are more likely to get marriage offers at older ages than Blacks or Hispanics, as we saw in Table 6. This facilitates college completion. So, when we apply the whites' offer probabilities to Blacks and Hispanics they tend to stay in school longer, increasing college graduation rates modestly.

In our model, parents' marital status effects both the skill endowment and the fixed cost of marriage. Conceptually, we argue that equalizing marriage market opportunities implies also equalizing the fixed cost of marriage across ethnic groups. Consider replacing the rate of single mothers for Blacks (61.5%) with the rate for whites (18.7%). This experiment reduces the fixed cost of marriage for blacks substantially, causing their marriage rate at age 32-36 to increase from 26% to 53%.

As we see in the bottom panel of Table 12, when we give Blacks the same marriage market opportunities as Whites, the college graduation rates of Black men and women increase by 1.4 and 3 percentage points, respectively. Interesting, for women, the effect of reducing fixed cost of marriage by itself is negative (minus 1.8 points). With marriage more likely, black women foresee that they will work less, which reduces the return to college. However, if a higher chance of marriage is combined with a higher chance of getting offers at older ages, the women are able to complete college and marry later. The later effect dominates, giving an overall positive effect.

For Hispanics equalizing marriage market constraints with whites increases college graduation rates by 2.6 and 3.3 percentage points for men and women, respectively. Overall, we find that equalizing marriage market constraints closes 10% and 15% of the college gap for black men and women, and 11% and 12% of the gap for Hispanic men and women.

Summary: The sum of the three exogenous factors (parental background, labor market, marriage market) closes 85% to 95% of the college graduation gap between whites, blacks and Hispanics, for both men and women. In our model, because we assume common preferences for all three ethnic groups, equalizing the three exogenous factors equalizes college graduation rates across the groups by construction. Thus, the model implies a very small complementarity when the three factors are changed simultaneously.[36]

VII. Will graduation rates continue to increase? Will the gender gap continue to grow?

In this section we use our model to address three questions about future trends in college graduation rates: Will graduation rates continue to increase? Will the gender gap continue to grow? Will Blacks and Hispanics catch up to Whites?

We use the family background (i.e., mothers' education, parents' marital status and immigration status) of the 1990, 2000 and 2010 cohorts (see Table 1), to predict their college graduation rates. In this exercise, we assume labor market and marriage market opportunities of these cohorts are the same as the 1980 cohort. Enough time has passed that we can see the collge graduation rates of the 1990 cohort (at age 30). Surprisingly, our predicted graduation rates for the 1990 cohort are very close to what we observe in the data, with a deviation of at most one percentage point (see Figure 5). The reason this prediction is so accurate is that wage paths for the 1990 cohort do not appear to be much different from the 1980 cohort (at least at the young ages where we can observe data). Therefore, it seems reasonable to also predict the collge graduation rates of the 2000 and 2010 cohorts based on the same assumptions as for the 1990 cohort – i.e., using only information on their family background.

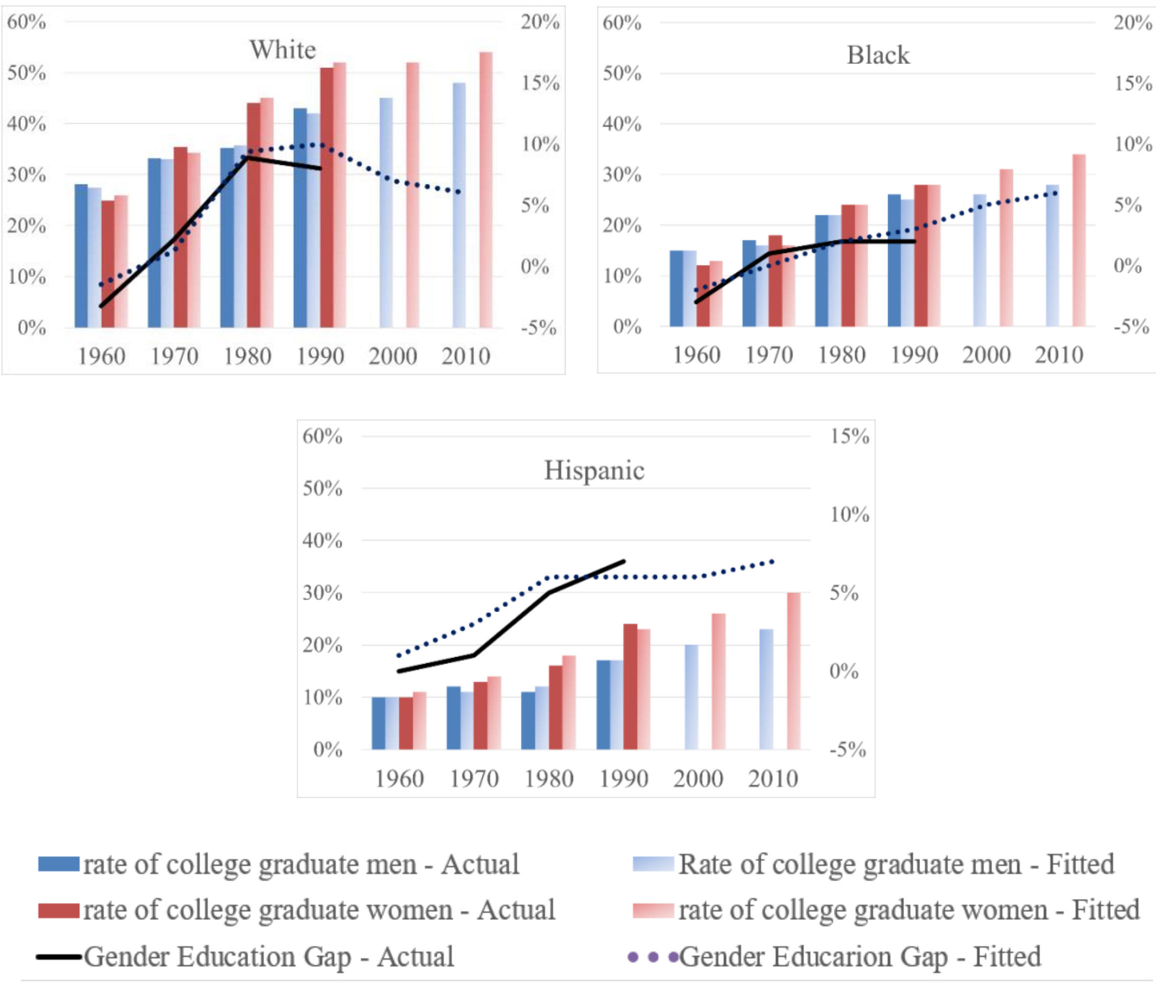
Consider first the results for Whites. The college graduation rate of White mother's increased from 38% in the 1990 cohort to 51% in the 2010 cohort. As we in Figure 5, our model predicts the growth in the college graduation rate of White women will slow down and stabilize at about 53% to 54% in the 2000 and 2010 cohorts. The reason we predict the graduation rate of white women will stabilize is that, by the 2000 cohort, a very high share of medium skill women with college graduate mothers are themselves graduating from college. It is difficult to induce a higher graduation rate amongst the marginal group, which consistes largely of medium skill women whose mothers did <u>not</u> graduate. These women get disutility from school (see Section V). We also predict the college graduation rate of White men will grow at a slighly faster rate than that of women. As a result, the gender gap will narrow slightly, from 9% in the 1980 cohort to 6% for the 2010 cohort.

---

[36] As an example of how the model decomposes education differences, consider the difference in college graduation rates between White and Black women in the 1980 cohort, 44.9 vs. 24.5. Differences in labor market returns to college explain 45% of the gap. Differences in skill endowments and tastes for schooling generated by the initial conditions (parent education and marital status) explain 36% of the gap. And differences in marriage market constraints (due to differences in marriage offer functions and tastes for marriage generated by parent marital status differences) explain 15% of the gap. The sum is 96%. In our model these three factors together explain the entire gap, so there is a very small positive interaction effect. Note that differences in preferences play no role in explaining the education gap between white and black women.

The story is rather different for Blacks and Hispanics. For Black women the college graduation rate of mother's increased from 13% in the 1980 cohort to 19% in the 1990 cohort and 28% in 2010. As a result, we predict the college graduation rate of Black women will increase from 24% in the 1980 cohort to 28% in the 1990 cohort and 34% in 2010. This brings the college graduation rate of Black women to roughly the rate of White women in the 1970 cohort. These substantial increases arise because the marginal Black women was still high skill in these cohorts, making it easier to induce them into attending college.

Figure 5: Actual and Predicted Graduation Rates, 1960 to 2010 cohorts

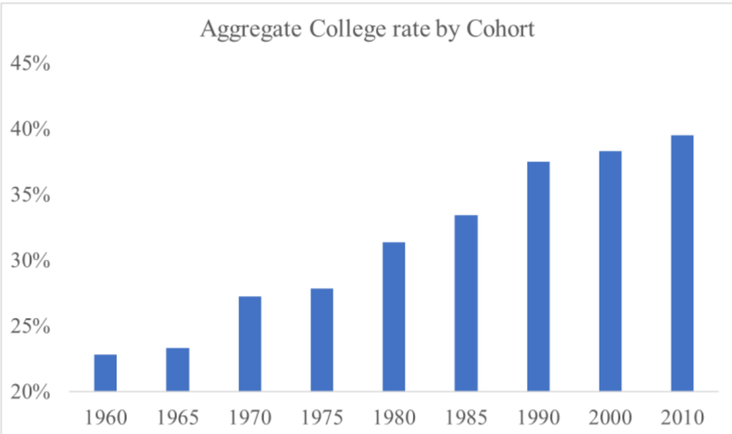


For Hispanic women the fraction of US born college graduate mothers increased from 5.7% in the 1980 cohort to 10.5% in the 1990 cohort to 17.3% in 2010. As a result, we predict the college graduation rate of Hispanic women will increase from 18% in the 1980 cohort, to 23% in 1990 to 30% in 2010. Again, these large increases occur because the marginal Hispanic woman was high skill in this period.

For Black men, we predict the college graduation rate will increase modestly, from 22% in the 1980 cohort to 28% in the 2010 cohort. Hence the gender gap in the graduation rate for Blacks will widen from 2 pp in the 1980 cohort to 6 pp in the 2010 cohort. For Hispanic men we predict the graduation rate will increase form 12% in the 1980 cohort to 23% in the 2010 cohort. This means the gender gap for Hispanics will increase only slightly from 6 to 7 pp.

In summary, we predict increases in college graduation rates for all six gender and ethnic combinations. The largest increases from 1980 to 2010 are for White men (12pp) and Hispanic women (12pp), followed by Hispanic men (11pp), Black women (10pp), White women (9pp) and finally Black men (6pp). All the above predictions are based on changes in family background, where increases in collge graduation rates of mothers across cohorts is the key factor driving the predicted increases in education. But the increase in single parent families, which tends to reduce ability, also plays a role. From the 1980 to 2010 cohorts, the rate of single parent households for Whites only increases slightly, from 18.7% to 21.9% (3 pp). But for Blacks the increase is 61.5% to 74.5% (13 pp), and for Hispanics is was 25.9% to 40.0% (14.1 pp). These large increases in single parent rates for Blacks and Hispanics is one tends to dampen their increases in education.

Figure 6: Aggregate College Graduation Rate



Finally, in Figure 6 we aggregate across genders and ethnic groups to construct the overall college graduation rate by cohort. This is historical from the 1960 to 1990 cohorts, and predicted for the 2000 and 2010 cohorts (whose true rates will not be observed until roughly 2030 and 2040). Our model predicts that the aggregate college graduaiton rate will only increase by 2 percentage points from the 1990 to 2010 cohorts (a 1 pp increase for for each).

This result may seem surprising, as our model predicts larger increases than 2pp for each individual group: If we compare the 1990 and 2010 cohorts, the largest predicted increase is for Hispanic women (7pp), followed by White men (6pp), Hispanic men (6pp), Black women (6pp), White women (4pp) and finally Black men (3pp).

The reason for this contrast is the dramatic drop in the population share of Whites, from 63% in the 1990 birth cohort to only 52% in the 2010 birth cohort. This is mainly driven by an increase in the share of Hispanics, from 25% in the 1990 cohort to 35% in the 2010 cohort. As the Hispanics have the lowest college graduation rate of any group, the increase in their population share drives down the aggregate rate.

VIII. College Tuition Subsidy Experiments

In this section we use the model to assess the size of tuition subsidies that would be needed to negate the impact of parental background on college graduation. According to our estimates, the annual cost of college implied by parameter TC in equation (18) is $58k per year. It is important to note that this is measured *relative* to high school, and that it includes tuition, room and board, travel costs, and the utility cost of college (i.e., the effort cost of studying). In addition to this, different types of agents get additional utility or disutility of school attendance (but we assume these differences are invariant across grade levels) – e.g., high ability women with college educated mothers get incremental utility from school at all levels.

As we saw in Table 12, our model implies that, in the 1980 birth cohort, differences in mother's college graduation accounted for 5.1 points of the college gap between Hispanic and White women. We predict that a tuition subsidy of $12.9k per year would increase the college graduation rate of Hispanic women by 5.1 pp, eliminating this source of difference with White women. Assuming that their children are in the 2000 birth cohort (20 years later), and that the subsidy is eliminated, we can further predict that the increase in mother's education in the 1980 cohort would cause the graduation rate of their daughters in next generation (2000 cohort) to increase by 3.5 pp, and their sons by 2.1 pp. Thus, there is a large passthrough of subsidy impact to the next generation.

Interestingly, the same size subsidy would only increase the college graduation rate of Black women by 2.3 pp. This is because, in the 1980 cohort, over 90% of high ability Black women were already graduating from college. In contrast, only about 60% of high ability Hispanic women were graduating – see Table 13 (Note: In the 1980 cohort the baseline attendance rate for Hispanic women was only 17.9%, compared to 24.5% for black women). Thus, it is more difficult to induce additional Black women to graduate using subsidies.

Table 13.A: Values of College for Blacks

			1960					1980				
Skill endowment	mother 's education	mother's marital status	type proportion	PV women	CG rate	PV men	CG rate	type proportion	PV women	CG rate	PV men	CG rate
High	HS	Married	27.0%	200	47.4%	260	52.2%	15.3%	680	89.6%	620	82.3%
High	COL	Married	1.6%	600	99.2%	640	97.3%	2.1%	790	97.5%	710	97.6%
High	HS	Single	12.4%	180	36.3%	250	46.5%	19.0%	670	96.4%	550	81.2%
High	COL	Single	0.9%	590	98.4%	620	98.5%	3.3%	775	98.3%	640	97.8%
Medium	HS	Married	14.7%	10	0.0%	15	0.0%	8.3%	70	0.0%	25	0.0%
Medium	COL	Married	0.3%	60	0.0%	45	0.0%	0.4%	90	0.0%	80	0.0%
Medium	HS	Single	16.0%	0	0.0%	0	0.0%	24.7%	0	0.0%	10	0.0%
Medium	COL	Single	0.4%	0	0.0%	0	0.0%	1.6%	20	0.0%	20	0.0%

Table 13. B: Values of College for Hispanics

			1960					1980				
Skill endowment	mother 's education	mother's marital status	type proportion	PV women	CG rate	PV men	CG rate	type proportion	PV women	CG rate	PV men	CG rate
High	HS	Married	24.0%	190	32.4%	240	32.1%	19.5%	510	57.3%	360	37.3%
High	COL	Married	3.0%	570	98.4%	600	98.3%	4.0%	760	95.9%	700	97.4%
High	HS	Single	2.2%	180	27.8%	230	27.6%	4.2%	490	53.2%	340	33.9%
High	COL	Single	0.3%	540	99.2%	585	99.3%	1.1%	740	98.2%	620	98.2%
Medium	HS	Married	37.1%	0	0.0%	0	0.0%	30.1%	35	0.0%	15	0.0%
Medium	COL	Married	2.6%	70	0.0%	60	0.0%	3.5%	85	0.0%	70	0.0%
Medium	HS	Single	4.2%	0	0.0%	0	0.0%	8.2%	0	0.0%	0	0.0%
Medium	COL	Single	0.4%	0	0.0%	0	0.0%	1.2%	15	0.0%	20	0.0%

Returning to Table 12, we see that differences in mother's college graduation accounted for 5.5 points of the college gap between Black and White women. We predict that a tuition subsidy of $26.3k per year would be required to increase the college graduation rate of Black women by this amount. We further predict that this would increase the graduation rate of their daughters in next generation (2000 cohort) by 2.6 pp (and their sons by 1.4 pp). Thus, there is again a large passthrough of subsidy impact to the next generation, although not as large as we saw for Hispanics.

IX. Conclusion

In this paper we have specified and estimated a model of individual and household decision making in which education, labor supply, marriage and fertility are all endogenous. We use the model to explain changes in college graduation rates in the 1960 through 1980 birth cohorts, by gender/ethnicity, based on three exogenous factors: family background, labor market opportunities and marriage market constraints. We discipline the model by requiring it to explain differences in all endogenous variables by gender, birth cohort (1960, 70, 80) and ethnicity (White, Black, Hispanic), using preferences that are fixed across cohorts and groups – differing only by gender. We use the model to assess the contribution of each exogenous factor to changes in graduation rates by cohort/gender/ethnic group. We also use the model to predict graduation rates in the 1990, 2000 and 2010 birth cohorts.

The first question is "Why do women graduate from college at a higher rate than men?" For example, in the 1980 cohort, White women graduated at a 45% rate compared to 36% for men, giving a 9 percentage point gender gap. For Hispanics the gap was 6 pp and for Blacks it was 2 pp. We find that labor market returns do not explain the gender gap: In fact, labor market returns to college are actually greater for men than women. However, the overall returns to college – factoring in labor market returns, marriage market returns and tastes for education – are greater for women than men. The main reason is that women simply get more utility from school than men. Given the structure of our model, this may subsume a number of factors: For example, woman may be better at studying (or like it more), so they get less disutility from putting in effort at college. Or they may place greater value on learning for its own sake. Or they may simply get more utility from social activities at college.

The second question we ask is "Why has the college graduation rate increased in general across the three cohorts?" For instance, in the 1960 cohort the college graduation rates of White women and men were 26% and 27%, respectively. But in the 1980 cohort these rates had increased to 45% and 36%. According to our model, the main factor was increasing labor market returns to education. Between the 1960 and 1980 birth cohorts, the college vs. high school offer wage premium for men increased from .42 to .54 log points, while for women it increased from .54 to .67, so for both men and women it increased by .13 log points. These changes in labor market opportunities caused the college graduation rate of White women (men) to increase by 10.4 (5.4) percentage points. The actual increases were 19 (9) points.

Two other factors also contributed importantly to the increase in the graduation rate: The first factor was the increase in mother's education, this increased both ability and tastes for school. The second factor, which was important for women, is a higher probability of getting marriage offers at older ages. This made it easier for women to delay marriage and fertility while pursuing a college degree.

The third question we ask is: "Why did the gender gap in college graduation *increase* over time?," or equivalently, "Why did the college graduation rate of women increase faster than that of men?" A key point of our paper is that this is very different from the question: "Why is the college graduation rate of women *currently* higher than that of men?" We find there are three key reasons that the graduate rate of women grew faster than that of men, leading to a widening gender education gap:

1) Our results imply changes in parental background, in particular the increasing share of mothers with a college degree, caused the college graduation rate of White women to increase by 5 percentage points between the 1960 and 1980 cohorts. Changes in parental background also caused the college graduation rate of White men to increase, but by a much smaller 1.4 percentage points. The effect is much greater for women because having a college educated mother increases tastes for school, and this effect is much stronger for daughters. (Having a college education mother also increases the labor market skill endowment, but this effect is similar for men and women). Similar patterns hold for Blacks and Hispanics.

2) Experience returns of White college women in the 1960 birth cohort were much smaller than those for White college men, but by the 1980 cohort they have almost caught up. In addition, job offer probabilities of white women improved across the three cohorts, so by the 1980 cohort they are very similar to those for white men. These improvements in experience returns and job offer probabilities explain why changes in labor market opportunities led to *larger* increases in education for women than for men. (Although, despite these improvements, women's *labor* market returns to college are still less than those of men.)

However, a key reason college grew less for Black and Hispanic women than for White women is that experience returns for Black and Hispanic college women grew much less than for White women. The same was also true for Black and Hispanic men relative to white men.

3) An increase in the probability of getting marriage offers at older ages caused the college graduation rate of White women to increase by 3.4 percentage points from the 1960 to 1980 birth cohorts. College became more attractive for women because college attendance did not crowd out opportunities for marriage (and fertility) to the extent that it did in the past. The same was true for Black and Hispanic women, but to a lesser degree, as their chances of getting offers at older ages are still lower than for whites.

The fourth question we ask is "Why do Blacks and Hispanics have much lower graduation rates than Whites?" Our model assumes common preferences across groups, so taste differences are not a factor. We find that differences in parental background account for 36% of the 20.4 percentage point college gap between Black and White women, and 32% of the 13.3 point gap between for Black and White men. For Black women the gap in mother's education alone accounts for 5.5 points of the gap, while impact of mother's marital status on ability accounts or an additional 1.8 points (because Blacks have a much higher rate of single parents). For Hispanics the role of parental background is a bit smaller, accounting for 26%

and 20% of the gap with Whites for men and women, respectively. This is because the single parent rate for Hispanics is similar to whites.

Labor market opportunities are the largest factor: They account for 45% of the college gap between Black and White women, and 56% of the gap between Black and White men. The most important difference is that Blacks have worse returns to experience than Whites. And Blacks have lower offer probabilities than Whites, both initially and in terms of how offer probabilities increase with experience. Results for Hispanics are similar overall, but their wage offer functions are relatively worse than Blacks while their job offer functions are closer to Whites. Hence it is offer wages that account for most of the gap between Hispanics and Whites.

Finally, the marriage market plays a relatively small but still significant role. If Black and Hispanic women had the same marriage market opportunities as White women, both in terms of the probability of getting offers and the chance of meeting husband with a college education, it would increase their marriage market returns to college education. This would cause their college graduation rates to increase by 3 points for Blacks and 3.3 points for Hispanics, closing 15% and 12%, respectively, of the gaps with White women.

We also use the model to predict future trends in college graduation rates. These are intergenerationally linked, as mother's college affects children's tastes for college (especially daughters), and also affects children's ability level. Thus, we can predict how increases in mother's education in one cohort (e.g., 1990) generate increases in education in their children's cohort (i.e., 2020). In general, we predict the recent large increases in women's graduation rates will cause their children's graduation rates to increase further. But the results differ substantially by ethnic group:

For Whites, we predict that growth in women's college graduation rate will slow down considerably in the 2000 and 2010 cohorts. This is because the marginal woman is now a medium ability person whose mother did not attend college. Such women get disutility from school attendance, and their labor market returns to college are not as great as for high ability women. So it is difficult to induce them to attend college and graduate. Thus, we predict the college graduation rate of white women will plateau at about 54%. We predict that the rate of White men will slowly catch up, so gender gaps amongst Whites will slowly narrow.

The story is different for Blacks and Hispanics. We predict the college graduation rate of Black women will increase from 28% in the 1990 cohort and 34% in the 2010 cohort. This will bring the college graduation rate of Black women to roughly the rate of White women in the 1970 cohort. Similarly, for Hispanic women we predict the graduation rate will increase from 23% in the 1990 cohort to 30% in 2010. These substantial increases arise because the marginal Black or Hispanic women is still a high ability type person in these cohorts, making it easier to induce them into attending college.

Thus, we predict the gender eduaction gap will continue to grow for Blacks and Hispanics, although the growth for Blacks is greater. In general, if we compare the 1990 and 2010 cohorts, the largest predicted increase in the college graduation rate is for Hispanic women (7pp), followed by White men (6pp), Hispanic men (6pp), Black women (6pp), White women (4pp) and finally Black men (3pp). However, if we aggregate across gender/ethnic groups, we predict the aggregate college graduation rate will only increase by 2 pp from the 1990 to 2010 cohorts. This is because Hispanics, who have the lowest graduation rate of any group, make up large shares of the 2000 and 2010 birth cohorts.

Finally, we use our model to predict the impact of tuition subsidies on college attendance. In contrast to prior work, the intergenerational linkage through mother's education allows us to predict how subsidies would affect college attendance not only for the generation directly affected by the subsidy but also for the next generation (after the subsidy is eliminated). We predict a tuition subsidy of $12.9k per year would increase the college graduation rate of Hispanic women in the 1980 birth cohort by 5.1 pp. Going forward, this would increase the college graduation rate of their daughters (sons) in the 2000 birth cohort (20 years later) by 3.5 pp, and their sons by 2.1 pp. Thus, there is a large passthrough of the subsidy impact to the next generation. This represents an important benefit of a tuition subsidy.

Interestingly, the same size subsidy would only increase the college graduation rate of Black women by 2.3 pp. This is because, in the 1980 cohort, over 90% of high ability Black women were already graduating from college. In contrast, only about 60% of high ability Hispanic women were graduating.

We predict that a tuition subsidy of $26.3k per year would increase the college graduation rate of Black women by 5.5 pp. We further predict that this would increase the graduation rate of their daughters (sons) in next generation (2000 cohort) by 2.6 pp, and their sons by 1.4 pp. Thus, there is again a large passthrough of subsidy impact to the next generation, although not as large as we saw for Hispanics.