SUPPLEMENTAL APPENDIX

NOT FOR PUBLICATION

A Extended Model with Skill-Biased Technical Change

Educational attainment and the evolution of the gender wage gap are central in this framework. Although the benchmark model uses both an exogenously closing gender wage gap and rising college premium, the model is easily amendable such that the general wage trends are the results of a general equilibrium model. The underlying mechanism for the closing gender wage gap and rising college premium is as follows.

Assume that men and women are born with two factors of production—human capital (brain) and raw labor (brawn), as in Guvenen and Kuruscu (2010) and Rendall (2010). A linear combination of these two labor inputs determines the wage rate an individual can earn in the labor market. Specifically, individuals are born with a fixed amount of brawn, r, that is the same within each gender but can differ across genders. In addition, individuals are born with some initial brain, h_0, and can acquire more by attending college and acquiring education, e. The exogenous driving force we are interested in is a specific form of skill-biased technical change (SBTC), which we model as a rise in the wage of brain relative to brawn. Table A.1 summarizes the key parameters by gender and education level.

Table A.1: Labor Market Endowments for Men and Women

	Uneducated	Educated
Men	r , h_0	r , eh_0
Wage	$w_t\left(\left(1-\theta_t\right)r+\theta_t h_0\right)$	$w_t\left(\left(1-\theta_t\right)r+\theta_t e h_0\right)$
Women	αr, h_0	αr, eh_0
Wage	$w_t\left(\left(1-\theta_t\right)\alpha r+\theta_t h_0\right)$	$w_t\left(\left(1-\theta_t\right)\alpha r+\theta_t e h_0\right)$

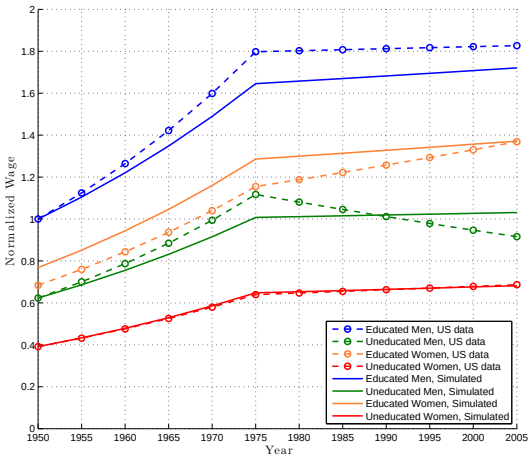
It can be shown that, with this structure, SBTC results in a closing of the gender wage gap and rising college premium. More specifically, we perform an experiment using a simplified model. Abstracting from the marriage market and the initial education decision, what would men and women's initial endowments and SBTC have to be in order to mimic the observed wage trends?

The simple model requires us to set six parameters $\{r, \alpha, h_{0,}, e, \theta_{1950}, \gamma_\theta\}$. We match the following targets:

1. Male brawn endowment, r, is set to match uneducated male wages in 1950;
2. Relative female brawn endowment, α, is matched to uneducated female wages in 1950;
3. Education factor, e, is matched to educated male wages in 1950;
4. Initial human capital, h_0, initial brain prices, θ_{1950}, and the five-year growth rate of brain prices, γ_θ, are set to match the general wage trends of educated, uneducated men and women from 1950 to 2005.

Since this exercise is illustrative and it abstracts from the marriage market and education decisions, and, therefore, also from any selection effect into the labor market, it will not be able to generate the precise wage trends. For this reason, wage trends from 1950 to 2005 are linearized (smoothed). As in the remainder of the paper, wages are assumed to have a general growth trend of 1.87% per year until 1975 and 0% thereafter. Figure A.1 plots the US linearized wage trends by gender and education with the simulated trends assuming a five-year growth rate of $\gamma_\theta = 0.013$ (our SBTC measure). The remaining parameters are the initial human capital level of $h_0 = 1.137$, an education factor $e = 2.324$, male brawn endowment of $r = 0.452$, gender differences in brawn of $\alpha = 0.316$, and initial brain prices of $\theta_t = 0.250$.

Figure A.1: Wages, US data and Simulation



The simulated wage trends clearly illustrate how this simple mechanism can capture most of the closing gender wage gap and rising college premium. Of course, since this simple computation has abstracted from the marriage market and education choice (ignoring any selection effects), the trends are linear in nature. With these caveats in mind, the experiment does remarkably well in matching the observed wage trends by gender and education.

B Population Dynamics and Equilibrium

Since individuals only marry within their cohort, we only need to keep track of the evolution of λ for a given cohort. Recall that

$$\lambda_m^t(\mathbf{z}_m; e_f) = \theta_m^{e_m e_f} \times P_m^t(\mathbf{z}_m),$$

where $P_m^t(\mathbf{z}_m)$ denotes the population fraction of men with characteristics $\mathbf{z}_m$ at age t. The initial share of educated women $P_f^0(\mathbf{z}_f)$ follows simply from the education decision. The share of educated women in a cohort is

$$P_f^0(s, 1, \overline{w}_f^u) = P\left(\beta \mathbb{E}_{t=1}\left(\widetilde{S}_1^f\left((s, 1, w_f^i), \lambda_m^1\right)\right) - \beta \mathbb{E}_{t=1}\left(S_1^f\left((s, 0, w_f^i), \lambda_m^1\right)\right) > \kappa\right).$$

All educated women at the beginning of college are single and earn $\overline{w}_f^u$ (the average wage conditional on no education). For uneducated women, the initial population only differs in that individuals have already drawn their actual ability level,

$$P_f^0(s, 0, w) = \left(1 - P_f^0(s, 1, \overline{w}_f^u)\right) g_f^u(w),$$

where $g_f^u(w)$ is the probability density of ability for an uneducated female.

The fraction of females that are single tomorrow is determined by two terms: females who do not meet any eligible men this year and those who meet eligible men but decide to remain single. That is:

$$P_f^{t+1}(s,e_f,w) = \left[\sum_{hh=s,d} \sum_{e_m} \int_{w_m} \left(1-\lambda_m^t(\mathbf{z}_m;e_f)\right) dG_m(w_m) \right.$$
$$\left. + \sum_{hh=s,d} \sum_{e_m} \int_{w_m} \lambda_m^t(\mathbf{z}_m;e_f) \int_{-\infty}^{\underline{B}^p} dF_0(b) dG_m(w_m) \right] P_f^t(s,e_f,w), \tag{8}$$

where $\lambda_m^t(\mathbf{z}_m;e_f)$ is as defined above (we suppress the arguments of the limits of integration). Similarly, the fraction of divorced females tomorrow depends on those currently divorced, those who do not marry between today and tomorrow, those who do not meet anyone, and the newly divorced. Since divorced individuals only meet somebody every five periods, we do not need to keep track of how long individuals have been divorced, but simply track the periods in the model. Also note that $P_f^0(d,e_f,w)=0$ by definition. Thus, we have the following:

$$P_f^{t+1}(d,e_f,w) = \left[\sum_{hh=s,d} \sum_{e_m} \left(\int_{\tilde{w}} \left(1-\lambda_m^t(\mathbf{z}_m;e_f)\right) dG_m(\tilde{w}) + \int_{\tilde{w}} \lambda_m^t(\mathbf{z}_m;e_f) \int_{-\infty}^{\underline{B}^p} dF_0(b) dG_m(\tilde{w}) \right) \right]$$
$$\times P_f^t(d,e_f,w) + \left(1 - P_f^t(d,e_f,w) - P_f^t(s,e_f,w)\right) \left[\sum_{e_m} \int_{\tilde{w}} \int_{\underline{B}^d}^{\infty} dF(b|b_{-1}) dG_m(\tilde{w}) \right], \tag{9}$$

where $\tilde{w}$ denotes men's wage in the integration. Last, we allow for cohorts to be of different sizes. Since the equilibrium should be stationary, we normalize the population size to one every period. Relative cohort sizes are given exogenously. The analogous expressions hold for men.

Equilibrium An equilibrium is a set of decision rules for women's college choice, e_f; time allocation decisions for single and divorced women that depend on their own state vector, $h_f^i(\mathbf{z}_f), \ell_f^i(\mathbf{z}_f)$, and $n_f^i(\mathbf{z}_f)$ for $i=s,d$, and when married that depend on both spouses' state vectors, $h_f^p(\mathbf{z}_m,\mathbf{z}_f), \ell_f^p(\mathbf{z}_m,\mathbf{z}_f)$, and $n_f^p(\mathbf{z}_m,\mathbf{z}_f)$; the same set of decision rules for men, defined analogously; marriage thresholds, $\underline{B}^p(\mathbf{z}_f,\mathbf{z}_m)$, and divorce thresholds, $\underline{B}^{\text{cd}}(\mathbf{z}_f,\mathbf{z}_m)$ and $\underline{B}^{\text{ud}}(\mathbf{z}_f,\mathbf{z}_m)$; value functions for women when they are married, single, and divorced, $J^f(b,\mathbf{z}_m;\mathbf{z}_f,\lambda_m^t)$, $S^f(\mathbf{z}_f,\lambda_m^t)$, and $D^f(\mathbf{z}_f,\lambda_m^t)$; value functions for men defined analogously; and the distribution of married, single, and divorced women, $P_f^t(\mathbf{z}_f)$ for $hh=p,s,d$; and the same for men, such that

- The value functions and decision rules solve individuals' utility maximization problems: 5-7; and
- Individuals' optimal choices are consistent with aggregate equilibrium shares: $P_m^t(\mathbf{z}_m)$ follows the laws of motion in (8) and (9); and symmetrically for men.

C Data Appendix

US targets (data) reported throughout the paper are obtained from four sources.

1. National Vital Statistics System of the National Center for Health Statistics (NCHS): marriage and divorce rates;
2. Survey of Income and Program Participation (SIPP): survival of marriages until 15th anniversary, remarriage rates, share of first marriages before age 34, fraction ever divorced, economies of scale;

3. Integrated Public Use Microdata Series - Current Population Survey (IPUMS-CPS): all other data post-1968; and
4. Integrated Public Use Microdata Series - Census (IPUMS-USA): same as IPUMS-CPS for 1950 and 1960.

C.1 NCHS's Marriage and Divorce Data 1968-1995

The National Bureau of Economic Research (NBER) hosts marriage and divorce data from the National Vital Statistics System of the National Center for Health Statistics (NCHS). The data contain information from marriage and divorce certificates. For small states it includes all records, for larger states it includes a sample. Due to a lack of funding by NCHS, data files are not available for years after 1995. For our purposes, the dataset records the date of marriage/divorce, education and age of bride and groom/divorced individual. The education of spouses is not reported after 1990. Population weights are available to obtain state-specific population sizes. Using the CPS, we compute current marital status-, (education-) and age-specific population sizes by state. Note that during the late 1960s through the 1970s the CPS aggregates smaller states into broader groups. The specific years when states were combined varies across regions of the United States. We use relative population shares from before and after the combined periods to assign population sizes to each grouped state. Combining the inflated NCHS weights by state and the CPS marital status and age specific population by state yields the denominator for each age-specific/education-specific marriage and divorce rate. Five-year moving averages for marriage rates per 1,000 unmarried (by age group 20-34 and education (college graduate versus non-college degree)) are reported.

C.2 SIPP 1984–2008

The SIPP surveys 1984-2008 report the beginning and separation dates for up to three marriages each survey. Furthermore, age, gender, and educational attainment are reported. By nature, most marital information is retrospective. However, only weights for the current survey year are available. Since all surveys have retrospective information, we compute specific statistics for a given year from several different surveys to check for inconsistencies in retrospective information.

Survival until 15th Anniversary. The survival rate until the 15th anniversary for individuals ages 20-34 is computed by tracking individuals whose marriages started when they were between 20 and 34 years of age. With the first survey recorded in 1984, it is possible to compute the survival rate starting in 1950 for individuals by tracking the marriage history of individuals ages 54–68 in 1984. Using the SIPP provided weights, five-year moving averages are reported. The resulting survival probabilities are comparable to existing National Vital Statistics (NVS) data for the whole population (ages 15 and above). Checking for robustness, the largest discrepancy across survey computation occurs in 1979, where the average survival rate across six surveys is 66.06% (with a minimum of 63.03 and a maximum of 69.97) and the standard deviation is 3.05%. In general, the standard deviation ranges from 3.05% to 0.17%, with the average standard deviation being 1.27 percentage points.

Remarriage Rates. The NVS reports remarriages for divorced and widowed women. We do the same here—although, note that the number of widowed individuals ages 20-34 is generally negligible. Individuals who have at least one previously failed marriage and have not yet remarried are assumed to be eligible to remarry. For the actual marriages, the same strategy as for the survival rate is used: we use the retrospective

Table A.2: Children at Home and Economies of Scale by Marital Status and Gender

Marital Status	Own Children	Step or Other Children	Economies of Scale Base
Married	1.31	0.09	2.36
Married (Div Woman)	0.94	0.78	2.71
Divorced Man	0.26	0.03	1.13
Divorced Woman	1.28	0.04	1.64
Single Man	0.04	0.01	1.02
Single Woman	0.36	0.02	1.18

martial history information to track the number of remarriages in a given year for all individuals that were ages 20-34 at the time of remarriage. Remarriage rates are available from NVS reports from 1900 to 1988 for women ages 18-54. Computing remarriage rates using our method, the data coincide going back only until 1970. Using retrospective information and current survey year SIPP weights for the remarriage rates is less accurate than for survival until the 15th anniversary. For the survival rate, we only need to track how long marriages that started in a certain year lasted. For the remarriage rate, we need to know the number of divorced individuals for each year we want to compute the remarriage rate. Since weights are not retrospective, using the weights of, for example, individuals in their 60s in 1986 to compute the number of divorced individuals in 1950 may be unreliable. Therefore, we only use 10 years of retrospective information in computing remarriage rates. With 10 retrospective years the results are robust. For example, for women ages 20-34 the average standard deviation across different survey years for a given data point is 1.76 percentage points.[36]

First Marriages. To compute the share of individuals who marry before age 34, we compute the ratio of individuals who married before age 34 to individuals who married at least once by the age of 54. Again, as before we use the retrospective marital information to determine when individuals married for the first time.

Fraction Ever Divorced. We use the current survey year and retrospective martial history to determine the fraction of individuals that has been divorced at least once, regardless of their current marital status.

Economies of Scales. The economies of scales are computed using individuals ages 25 to 34. SIPP surveys report the number of biological as well as stepchildren living at home. We compute the average number of children from 1984 to 2008 living at home by current marital status and past divorce history. To compute the economies of scales, it is assumed that the second adult adds 0.7 equivalent members to a household, each biological child adds 0.5, and a stepchild or other child adds 0.7 members to a household. Since married couples have a small share of children 0.09 that are not biological children, we drop these children.

In general, couples that remarry tend to have more children living with them, since they have similar numbers of biological children than never-divorced couples, but also tend to have stepchildren living with

[36]The standard deviations across sample are, in relative terms, about three times larger than for the survival rates.

Table A.3: Children by Marital Status and Gender (25-64)

		Children type:		
	Total Own	Own at Home	Step at Home	Other at Home
Marital Status:				
Married	2.00	1.30	0.02	0.06
Married (Div-Div)	1.91	0.44	0.57	0.08
Married (Div M)	1.83	1.02	0.10	0.06
Married (Div F)	1.31	0.84	0.41	0.10
Divorced Man	1.70	0.26	0.01	0.02
Divorced Woman	1.99	0.99	0.01	0.05
Single Man	0.26	0.05	0.00	0.01
Single Woman	0.65	0.37	0.00	0.02

them. Table A.3 details the number and type of children by current and prior marital status for all individuals ages 25-64. The table reports averages from the 1984 to 2008 surveys.

C.3 IPUMS-CPS and Census

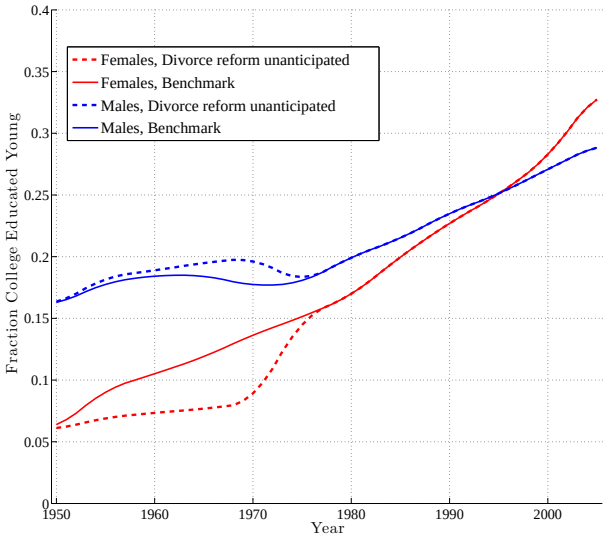
All remaining US data are obtained from IPUMS-CPS 1968 to 2005 and IPUMS-USA 1950, 1960 surveys. For all computations we use individuals ages 20-34, except for the hours targets where we use individuals ages 25 - 34, assuming most individuals have completed their education by age 25. Individuals in the armed forces, currently at school, or widows are dropped from the sample. As are all individuals with missing education information. All summary statistics use weights provided by the CPS or Census Bureau.

Stock of Divorced, Married, and Single. In computing stocks of individuals by marital status, we make the following assumptions regarding status. In the CPS and Census data the category "separated" pertains only to married persons who were legally separated or parted because of marital discord. Therefore, we treat both separated and divorced individuals as divorced. Further, the distinction between "married, spouse present" and "married, spouse absent" was made during the interview itself. The second category included individuals that stayed most of the week at a different residence. Therefore, we treat both categories as married in the paper. Unmarried respondents in households with unrelated adults have been asked about cohabitation only since 2007. Therefore, cohabiting individuals are classified as either single, divorced, or separated.

Education. All individuals with at least four years of college or a bachelor's degree are coded as having a college degree. Individuals ages 25 - 34 are used to compute the fraction of college-educated individuals each year, assuming most individuals have finished their education by age 25 to 34. The same education classification is used when computing couple types by educational attainment.

Hours Worked. Hours are computed for all individuals (including those who did not work). IPUMS-CPS reports usual hours worked per week in the last year starting in 1976. Prior to 1976 only hours worked last week are reported. We follow McGrattan and Rogerson (1998) to compute a constant hours measure over time.

Figure A.2: Anticipation of Divorce Reform on College Attainment Prior to 1970s



Wages. Weekly wages are computed for full-time workers, that is, individuals that worked at least 1,400 hours, using hours information from above and weeks worked last year. Top-coded observations on earnings are multiplied by 1.45. All wages are deflated by the US Bureau of Labor consumer price index. Assuming a 40-hour workweek, weekly wages of less than half of the minimum wage, at the time, are dropped from the sample. Variance in wages in logarithm is equal for all gender and education groups and set at 0.25, roughly the average.

D Alternative Assumptions

D.1 Unanticipated Divorce Reform

As seen in Figure A.2, when individuals do not anticipate divorce reform and the related rise in divorce risk (in equilibrium), the educational choices of men and women before the 1970s would have been quite different. But the effect would be in opposite directions. Whereas anticipating higher divorce risk caused women to increase their education level in the 1950s and 1960s, it had the opposite effect on men: they lowered their education level. This is because women have a stronger incentive to insure against divorce by accumulation education (i.e., divorce risks played a central role in explaining the reversal of the gender education gap).

References

Guvenen, F. and B. Kuruscu (2010, September). A Quantitative Analysis of the Evolution of the U.S. Wage Distribution: 1970-2000. In *NBER Macroeconomics Annual 2009, Volume 24*, pp. 227–276. University of Chicago Press.

McGrattan, E. R. and R. Rogerson (1998). Changes in Hours Worked Since 1950. *Quarterly Review 22*(1), 2–19.

Rendall, M. (2010, Sep). Brain versus Brawn: the Realization of Women's Comparative Advantage. IEW - Working Papers 491, Institute for Empirical Research in Economics - IEW.