SCHOOLING, MARRIAGE, AND MALE AND FEMALE CONSUMPTION

Gustaf Bruze*

October 26, 2008

Abstract:

A marriage matching model is estimated to quantify the share of returns to education that is realized through marriage. In the model, more educated agents earn higher wages in the labor market, and are more productive in housework. Men and women who marry benefit from the presence of household public goods, complementarities in household production, and the division of labor between spouses. The predictions of the model are matched with NLSY data on sorting in marriage, and data on the allocation of time from a time use study. Counterfactual analysis for men and women at age 40, suggests that better marital outcomes generate 65 percent of the return to education for women around middle age and 20 percent of the corresponding return for men.

* I thank the members of my thesis committee, Casey Mulligan, Ali Hortaçsu, and Jeremy Fox for their advice and support. All remaining errors are my own.

1) Introduction

Men and women who spend time in school receive many different future returns to their investment. In this paper, I provide quantitative estimates of the returns to education that are only realized by men and women who marry. More specifically, I use data on the generation of American men and women born around 1960, and estimate how much of their return to education at middle age that was realized through improved marital outcomes.

To answer this question, I estimate a static matching model which is based on the marriage model of Becker (1973). In the estimated model, all agents choose how to allocate their time in an optimal way, and the equilibrium marriage matching is stable. An important concept in the model is the marital surplus that is realized when a man and a woman marry, and which consists of three parts. First, some goods that are consumed in marriage are public, second, married agents benefit from the division of labor, and third, the time inputs of men and women in housework are complementary.

The model economy is populated by men and women who differ in terms of their educational attainment, wage rates, non-labor income, and housework productivity. Since no analytical expression is available for the expected matching of a marriage model with such a heterogeneity of male and female types, I estimate the parameters of the marriage model with the method of simulated moments. The estimated model is then used to construct counterfactual outcomes for men and women when additional years of schooling raise their own wages and housework productivity, but do not improve their marital outcomes. The results from this exercise indicate that variations in the quality of marital outcomes generate 65 percent of the consumption difference between middle aged women at different levels of schooling, and 20 percent of the corresponding difference for middle aged men.

Previous attempts to quantify how schooling affects individual consumption through marriage, include Goldin (1992) and Lefgren and McIntyre (2006)[1]. These authors consider outcomes for women only, under the assumption that men and women split total household income in half, and

[1] The concept of marriage market returns to education has a long history in the human capital literature. An early reference is Becker (1960). Other recent papers that address marriage returns to education are Ge (2007) and Lafortune (2008).

use their shares for the consumption of a private good. Both studies find that women earn in the order of 50% of their returns to education through marriage.

The estimates in this paper are derived from a model in which the value of housework is also considered. In addition, the division of resources in marriage is estimated rather than assumed by using the method developed in the collective household literature (Chiappori 1988). The gains from a richer specification however, come at the cost of additional and difficult measurement problems. As a robustness check, I also estimate how schooling affects consumption through marriage by using other and more traditional welfare measures such as full income. The main results of this paper turn out to be similar across a range of such alternative measures.

From a methodological point of view, this paper is part of a recent literature that tries to place the collective model of household behavior within the context of a general equilibrium marriage matching model. A previous example is the study by Del Boca and Flinn (2006), who combine a model of household time allocations with the matching model of Gale and Shapley (1962) to evaluate if married couples behave in a cooperative manner. A second example can be found in Choo et al. (2008), who integrate the marriage model of Choo and Siow (2006) with the collective household model, and show that the sharing of resources in marriage can be estimated not only with data on the labor supplies of husbands and wives, but also with data on who marries whom across a set of isolated marriage markets.

The rest of this paper is organized as follows. In section 2 below, I present the model that is estimated in this paper, and section 3 presents some key properties of the model equilibrium. Section 4 contains a description of the data, section 5 discusses the issue of identification, and section 6 presents the estimation method that is used in this paper. Section 7 presents the main stylized patterns of sorting in the US marriage market, and section 8 contains the results from the estimated model. Finally, section 9 concludes with a discussion of ways in which the estimates in this paper can be extended and improved.

2) Model

Basic setup

The economy is static and populated by men and women who are located across a set of K isolated markets. In any given market k, there are a total of M_k men indexed by $i = 1,2, \dots, M_k$, and a total of F_k women indexed by $j = 1,2, \dots, F_k$. Agents have the choice of remaining single, or marrying a member of the opposite sex in their own market. All agents are characterized by their years of schooling s, their non-labor income y, their wage rate w, and a vector of additional characteristics that are specified below.

Men and women derive utility from the consumption of a private consumption good c, leisure l, and a public household good q. Both men and women rank bundles of these commodities according to the log utility function

$$u(c, l, q) = \phi \cdot log(c) + \delta \cdot log(l) + (1 - \phi - \delta) \cdot log(q)$$

Furthermore, agents have a unit time endowment which can be used for leisure, housework h, and market work $1 - h - l$. I use the subscript o for singles and normalize the price of the market good for single agents to one. The budget constraint for a single agent is thus

$$c_o = y_o + w_o \cdot (1 - h_o - l_o)$$

In real life, a married couple can obtain more consumption services from a given amount of dollars, than if the two household members had lived apart as singles[2]. I capture these effects in the model, by letting married agents face a price p for the private consumption good that is lower than the price which single agents face. With subscripts m and f for husbands and wives respectively, the budget constraint for a married couple can then be written as

$$p \cdot \left(c_m + c_f\right) = y_m + y_f + w_m \cdot (1 - h_m - l_m) + w_f \cdot \left(1 - h_f - l_f\right)$$

[2] Examples of the sources of such consumption savings, include the fact that a couple can share consumer durables (television sets, radios, microwave ovens, etc), whereas two people who live as singles each need to acquire these consumer goods to obtain the same levels of consumption services that they enjoy when they live together.

As is conventional in the household economics literature, I use a square root formula for household consumption savings, so that the price p is equal to

$$p = \frac{1}{\sqrt{2}} \approx 0.7$$

Wages

The wages of men and women are modeled with standard Mincer equations. The log hourly wages of an agent depends on his or her years of schooling, experience, experience squared, IQ, social background[3] θ, and a market dummy. The log hourly wage for man i in market k, is thus given by

$$log\left(w_m^{i,k}\right) = \beta_m^{1,k} + \beta_m^2 \cdot s^{i,k} + \beta_m^3 \cdot exp^{i,k} + \beta_m^4 \cdot exp2^{i,k} + \beta_m^5 \cdot IQ^{i,k} + \beta_m^6 \cdot \theta^{i,k} + \varepsilon_m^{i,k}$$

and the log hourly wage for woman j in market k, is given by

$$log\left(w_f^{j,k}\right) = \beta_f^{1,k} + \beta_f^2 \cdot s^{j,k} + \beta_f^3 \cdot exp^{j,k} + \beta_f^4 \cdot exp2^{j,k} + \beta_f^5 \cdot IQ^{j,k} + \beta_f^6 \cdot \theta^{j,k} + \varepsilon_f^{j,k}$$

In these equations, the wage error terms ε_m and ε_f are independent across agents and markets, independent of all the personal characteristics of an agent, and have normal distributions

$$\varepsilon_m \sim N(0, \sigma_m^2) \quad , \quad \varepsilon_f \sim N\left(0, \sigma_f^2\right)$$

Household technology

The public household good q is produced with the housework inputs of the household members. To capture the effect of schooling on non-market productivity[4], I let the effective amount of human capital that each agent has available for housework grow exponentially with the years of schooling of that agent, and (in a restrictive way) the IQ and social background of that agent. More formally, I write the effective amount of human capital of an agent as

[3] Social background is a factor based on the years of schooling and occupations of the agent's mother and father.
[4] A large literature documents that more educated men and women are more productive in non-market work. Two recent examples are Elias (2005), and Ehrlich et al. (2008).

$$e^{\gamma \cdot s+\pi \cdot v_n} \qquad n = m, f$$

where

$$v_n \equiv \beta_n^5 \cdot IQ + \beta_n^6 \cdot \theta + \varepsilon_n \qquad n = m, f$$

and where γ and π are parameters to be estimated. For singles, I assume a linear household technology

$$q = e^{\gamma \cdot s_n+\pi \cdot v_n} \cdot h_n \qquad n = m, f$$

and for married agents, the household production function is of the CES form

$$q = \left((e^{\gamma \cdot s_m+\pi \cdot v_m} \cdot h_m)^v + \left(e^{\gamma \cdot s_f+\pi \cdot v_f} \cdot h_f\right)^v\right)^{1/v} \quad , \qquad v \leq 1$$

With such a specification, a married couple in which only one of the household members devotes time to housework, produces the same quantity of the household good as that household member would have produced, had he or she lived apart as a single agent. Furthermore, the extent to which the housework of the husband and the wife in a couple are direct complements, depends on the parameter v which also determines the elasticity of substitution in production. In section 5 below, I discuss how this parameter is identified from the data.

Single households

Agents who are single obtain utility through the consumption of the market good, leisure, and the household good. Formally, single agents choose their market work, leisure and housework to maximize their utility subject to their budget constraint, household technology constraint, and time constraints. If V_o is the utility obtained by a single agent, then this utility can be written as

$$V_o(w_o, s_o, y_o) = \max\ \phi \cdot log(c_o) + \delta \cdot log(l_o) + (1-\phi-\delta) \cdot log(q)$$

s.t) $$c_o = y_o + w_o \cdot (1 - h_o - l_o)$$

$$q = e^{\gamma \cdot s_n+\pi \cdot v_n} \cdot h_n \qquad n = m, f$$

$$0 \leq 1 - h_o - l_o \leq 1,\ \ l_o \leq 1,\ \ h_o \leq 1$$

Married couples

Following the collective household literature[5], I assume that allocations for married agents are Pareto efficient and use the labor supply and leisure of men and women as indicators of how couples share their resources. If U_m and U_f are the utilities that a husband and a wife obtain from the consumption of the market good, leisure, and the household good only, then by definition, these two utilities satisfy

P) $U_m = \max\ \phi \cdot log(c_m) + \delta \cdot log(l_m) + (1-\phi-\delta) \cdot log(q)$

s.t) $\phi \cdot log(c_f) + \delta \cdot log(l_f) + (1-\phi-\delta) \cdot log(q) \geq U_f$

$$p \cdot (c_m + c_f) = y_m + y_f + w_m \cdot (1-h_m-l_m) + w_f \cdot (1-h_f-l_f)$$

$$q = \left((e^{\gamma \cdot s_m + \pi \cdot v_m} \cdot h_m)^v + (e^{\gamma \cdot s_f + \pi \cdot v_f} \cdot h_f)^v\right)^{1/v} \quad , \quad v \leq 1$$

$$0 \leq 1-h_n-l_n \leq 1 \qquad n = m,f$$

$$l_n \leq 1 \ , \quad h_n \leq 1 \qquad n = m,f$$

In addition to the utility obtained from the consumption of commodities, married agents also obtain utility from the match with their spouse. I let V_m and V_f denote the total utility for men and women in marriage, and define this total utility as the sum of the utility obtained from the consumption of commodities, and the marriage match utility.

In detail, I assume that the total utility $\mathrm{V}_m^{i,j}$ that man i in a given market obtains if he marries woman j in that market and receives utility $U_m^{i,j}$ from the consumption of commodities, is

$$\mathrm{V}_m^{i,j} = U_m^{i,j} + \tau_s \cdot s_m \cdot s_f + \tau_{IQ} \cdot IQ_m \cdot IQ_f + \tau_\theta \cdot \theta_m \cdot \theta_f + \eta_m^i$$

By analogy, the total utility $\mathrm{V}_f^{i,j}$ that woman j in a given market obtains if she marries man i in that market and receives utility $U_f^{i,j}$ from the consumption of commodities, is

[5] See for example Chiappori (1988), and Chiappori et al (2002).

$$\mathrm{V}_f^{i,j} = U_f^{i,j} + \tau_s \cdot s_m \cdot s_f + \tau_{IQ} \cdot IQ_m \cdot IQ_f + \tau_\theta \cdot \theta_m \cdot \theta_f + \eta_f^j$$

In these equations, τ_s, τ_{IQ}, τ_θ are free parameters that contribute to the complementarities of the years of schooling, IQ, and social background of the husband and the wife in marriage, and which will be estimated from the patterns of sorting on these traits in equilibrium.

In addition, η_m and η_f are random utility terms that captures the heterogeneity in preferences for marriage among men and women. These two random variables are independent across men and women, and normally distributed with parameters

$$\eta_n \sim N(\mu_n, \sigma^2) \qquad , \qquad n = m, f$$

Marriage matching

To make the division of resources in marriage endogenous, I use the matching model of Crawford and Knoer (1981). For each potential couple, there are a finite number of permitted allocations that differ in terms of how the total resources available in marriage are divided between the husband and the wife. Since allocations are assumed to be Pareto optimal, these allocations all lie along the utility possibility frontier of the couple.

More formally, let

$$\Pi_u^{i,j} = \left\{ \left(U_{m,a}^{i,j}, U_{f,a}^{i,j} \right) \right\}_{a=1,A}$$

be the set of A distinct pairs of utilities obtained from the consumption of commodities, that correspond to the permitted allocations for the couple (i, j) in a given market. Each husband utility $U_{m,a}^{i,j}$ is the value of the Pareto problem (P) specified above, given that the wife obtains utility $U_{f,a}^{i,j}$ in marriage from the consumption of commodities.

I order these allocations so that the sequence of utilities for the husband is strictly decreasing

$$U_{m,1}^{i,j} > U_{m,2}^{i,j} > \cdots > U_{m,A}^{i,j}$$

and the sequence of utilities for the wife is strictly increasing. Figure 1 below illustrates three such pairs of utilities

$$\left(U_m, U_f\right) \in \left\{\left(U_{m,a}, U_{f,a}\right)\right\}_{a=1,2,3}$$

along the utility possibility frontier of a married couple.

To each of these permitted utilities from the consumption of commodities, one can add the utility of the marriage match itself to the husband and the wife, to obtain the set

$$\Pi_v^{i,j} = \left\{\left(V_{m,a}^{i,j}, V_{f,a}^{i,j}\right)\ \right\}_{a=1,A}$$

of total permitted utilities in marriage. Ultimately, men and women care only about these total utilities in marriage when making decisions about if they should remain single or get married, and if so, with whom. The set of permitted allocations for a couple however, differ only in terms of the allocation of commodities between them, and not in terms of the utility of the marriage match itself.

A matching in a given market is defined as a specification of the men and women who are single, a one-to-one function g from the set of married men to the set of married women, such that if man i and woman j are married, then

$$j = g(i) \quad , \qquad i = g^{-1}(j) \equiv f(j) \quad ,$$

and a division of utility

$$\left(U_m^{i,j}, U_f^{i,j}\right) \in \Pi_u^{i,j}$$

for every married couple (i, j) in the market.

A matching is said to be individually rational, if there is no married agent that would prefer to be single. Moreover, a matching in a given market is said to be blocked by man i and woman j in that market, if there is a permitted division of resources $\left(U_m^{i,j}, U_f^{i,j}\right) \in \Pi^{i,j}$ between them, such that they both weakly prefer marriage under the permitted division to the matching, and at least one of them strictly prefers marriage under the permitted division. Finally, if an individually rational matching is not blocked by any pair of man and woman, then it is said to be stable.

In a framework such as the marriage model presented above, Crawford and Knoer (1981) prove that the set of stable matchings is non-empty. The proof proceeds by constructing an algorithm through which agents on one side of the market (men or women), propose to members of the opposite sex in stages by promising an allocation that delivers a given total utility to their potential future spouse. Every time an offer of marriage is rejected, the proposing party is forced to either raise his or her offer, propose to someone else, or remain single.

As in the college admissions model of Gale and Shapley (1962), the matching model that is employed in this paper has multiple stable matchings. I select the unique matching which is obtained when men propose to women (the details of the algorithm are described in Appendix A).

3) Properties of the equilibrium

Singles

For agents who end up being single, the solution to their time allocation problem leads to a simple demand system in leisure, housework, and market work. For the econometric specification, I include demographic controls X and a classical measurement error ϵ in all reduced form time allocation equations, and write this demand system as

$$l_o = \delta + \delta \cdot \frac{y_o}{w_o} + X \cdot \Psi_{0,l} + \varepsilon_{0,l}$$

$$h_o = (1 - \phi - \delta) + (1 - \phi - \delta) \cdot \frac{y_o}{w_o} + X \cdot \Psi_{0,h} + \varepsilon_{0,h}$$

$$1 - l_o - h_o = \phi - (1 - \phi) \cdot \frac{y_o}{w_o} + X \cdot \Psi_{0,1-l-h} + \varepsilon_{0,1-l-h}$$

Married couples

For married couples, analytical expressions for the solution to their time allocation problems are not available, as these allocations depend on the division of resources in marriage, which in turn is the endogenous outcome of the marriage matching process. To characterize the time allocations of married agents, I instead rely on the general properties of an efficient allocation.

As is emphasized by the collective household literature, Pareto efficiency implies that the time allocation of a married couple can be thought of as if it was generated by a two-step process. First, the husband and the wife decide how much time they are each going to devote to housework, and thus, how much of the household good they will consume. Then, in a second stage, the couple divides the remaining resources between themselves and each household member chooses how to allocate his or her time between leisure and market work[6].

Formally, the problem that the husband and the wife solve in the second stage of this process is[7]

$$max \quad \phi \cdot log(c_n) + \delta \cdot log(l_n) \qquad\qquad n = m, f$$

$$\text{s.t)} \qquad p \cdot c_n + w_n \cdot l_n = \lambda_n$$

where λ_m and λ_f are the resources devoted to the husband and the wife respectively. Due to the budget constraint of the couple, these two shares have to satisfy

$$\lambda_m + \lambda_f = y_m + y_f + w_m \cdot (1 - h_m) + w_f \cdot (1 - h_f)$$

With this notation, the individual time that is devoted to leisure by the husband and the wife, can be written as a function of the resources that the husband and the wife receive in equilibrium

$$l_m = \left(\frac{\delta}{(\phi + \delta) \cdot w_m}\right) \cdot \lambda_m + X \cdot \Psi_{m,l} + \varepsilon_{m,l}$$

$$l_f = \left(\frac{\delta}{(\phi + \delta) \cdot w_f}\right) \cdot \lambda_f + X \cdot \Psi_{f,l} + \varepsilon_{f,l}$$

The equilibrium shares λ_m and λ_f are complicated objects that depend on the observable and unobservable characteristics of all agents in a given market, and the distributions of the random utility terms over spouses for these agents. I do not specify a given functional form for these sharing rules as part of the structural model. Rather, I treat the shares for men and women as unobserved random variables, and estimate the parameters of the model through simulation.

---

[6] This is an application of the second welfare theorem.

[7] Due to the separability of the household good from consumption and leisure that follows with log preferences, the solution to this second stage decision problem is independent of the amount of the household good that the husband and the wife consume.

In the case of housework, the time inputs of husbands and wives can be characterized analytically when both of these household members supply time to market work. For such couples, efficiency in household production requires that the ratio of the marginal products of the housework of the husband and of the wife, is proportional to the ratio of their wages:

$$\frac{\partial q}{\partial h_m}\Big/\frac{\partial q}{\partial h_f} = \frac{w_m}{w_f}$$

This result is reminiscent of the solution to the cost minimization problem of a firm, which chooses its inputs so that the marginal rate of technical substitution equals the ratio of factor prices. In the case of a household at an interior solution, the costs of the time inputs to housework are the wage rates of the husband and the wife. With the CES production function, this relationship can be used to obtain an expression for the optimal ratio of male to female housework at an interior equilibrium, which is

$$\frac{h_m}{h_f} = \left(\frac{w_f}{w_m}\right)^{\frac{1}{1-v}} \cdot \left(\frac{e^{\gamma \cdot s_m + \pi \cdot v_m}}{e^{\gamma \cdot s_f + \pi \cdot v_f}}\right)^{\frac{v}{1-v}}$$

Since the household production function is constant returns to scale, the ratio of the marginal products of housework is invariant to the scale of production, so that the quantity of the household good q drops out of the condition above. This makes it possible to identify the parameters of the household production function form data on inputs and factor costs alone (see Pollak and Wachter 1975). Economic theory restricts the response of the ratio of male to female housework as a function of changes in the ratio of male to female hourly wages (since $v \leq 1$) , but the effect of the relative educational attainment of the husband and the wife on the ratio of male to female housework cannot be signed.

Turning finally to levels of housework, closed form solutions are once again available for couples in which both the husband and the wife supply time to market work. The amount of housework that they perform, is equal to

$$h_m = (1 - \phi - \delta) \cdot (w_m + w_f + y_m + y_f) \cdot \left(\frac{e^{\gamma \cdot s_f + \pi \cdot v_f}}{w_f}\right)^{\frac{1}{1-v}} \cdot z^{\frac{v}{1-v}} + X \cdot \Psi_{m,h} + \varepsilon_{m,h}$$

$$h_f = (1 - \phi - \delta) \cdot (w_m + w_f + y_m + y_f) \cdot \left(\frac{e^{\gamma \cdot s_m + \pi \cdot v_m}}{w_m}\right)^{\frac{1}{1-v}} \cdot z^{\frac{v}{1-v}} + X \cdot \Psi_{f,h} + \varepsilon_{f,h}$$

where z is the constant cost per unit of the household public good, which follows from the constant returns to scale assumption.

4) Data

To construct a set of marriage markets for men and women, I use cross-sectional data for 2000 from the National Longitudinal Survey of Youth 1979 (NLSY79). I limit myself to white men and women in the representative part of the survey, for whom the median age difference between husbands and wives is 2 years. I therefore divide all respondents into six different markets based on their birth year, so that men who were born in 1957 can marry women born in 1959, men born in 1958 can marry women born in 1960, and so on. This gives a total of six markets based on the ages of men and women which are displayed in Table 1.

For each of these six cohort marriage markets, I draw a total of 160 single or married men and women[8]. With the notation from section two above, this implies that I set

$$M_k + F_k = 160$$

To make sure that each of the six cohort samples are representative of the relevant population, I chose the number of single men, single women, and married couples in each market to match the sex ratio, and the ratio of singles to married agents in the 5% sample of the 2000 US Census. I also selected the number of agents in each cohort sample, so that the educational attainment of singles and married couples corresponded to the distribution of these agents by educational attainment in the 2000 Census (more details about the sampling procedure can be found in Appendix C).

[8] Members of cohabiting households were not sampled. Since there are few marriages between whites and blacks, I also limited the sample to white respondents only. The total number of agents can be increased.

The NLSY79 does contain flow variables for the non-labor income that each household member received in the year of 2000. Instead of using this variable, I computed a new non-labor income variable as a 10% earnings flow on the entire stock of wealth of each household, minus an imputed value of last year's net savings. In the case of married couples, I then divided this non-labor income stream so that husbands received 60% of the non-labor income, and wives received 40%[9]. After these imputations, the overall ratio of non-labor income to wage earnings was 25% for the agents in the six constructed NLSY cohort marriage markets.

Estimation of the marriage model also requires data on the time allocations of men and women. I work with the data set "Time Use in Economic and Social Accounts 1975-76", which was collected by the Survey Research Center at the University of Michigan[10]. This data set has information on the time use of both spouses in a couple and good demographic controls, but the sample size is rather small. Other larger data sets are available but they only sample one person per household.

The sample in the time use data was constructed to represent the adult US population above age 18 when the study was conducted. Each respondent and his or her spouse were asked to fill in a time diary in the four waves of the survey that were conducted during an entire year. I only use data from respondents who appeared in at least three waves of the study, and where these interviews were conducted on a weekday, a Saturday, and a Sunday[11].

The principal investigators of the time use study break down all time use for respondents into the categories of market work, housework, leisure, personal care, and unaccounted for time. The details of this classification can be found in Appendix B. I assume that each agent has 14 hours available per day for the three activities of market work, housework, and leisure, and create the leisure variable as the residual of the total available 14 hours per day, minus the recorded market and housework time in the data.

Not all respondents provided complete time diaries in which the sum of all recorded activities during a day was equal to 24 hours. I dropped observations where more than 90 minutes per day

[9] I also experimented with other divisions, but this had no effect on the overall results.
[10] I am in the process of changing this data set to time use data from the American Time Use Survey (2003-2007).
[11] The time use study was intentionally set up to sample from these weekdays as US time use patterns have been found to vary systematically across weekdays, Saturdays, and Sundays.

were unaccounted for, observations with missing demographic information, and observations for student and retired respondents. After all these additional requirements, I was left with 694 observations of time use during a representative week for men and women in the ages of 20 to 65 years.

5) Identification

The marriage matching model implies a set of reduced form equations for the time allocations of single agents. These equations together with data on time devoted to market work, leisure and housework, identify the preferences that men and women have over the market consumption good, leisure, and the public household good (it should be added that this is not the only variation in the data that identifies the preferences of men and women over goods).

The parameters of the household production function are in turn identified by the efficiency condition for male and female housework in couples where both spouses are working. In particular, the parameter v which determines the elasticity of substitution for male and female housework, can be recovered with data on how households in a cross-section adjust their time inputs as a function of the ratio of wage rates that they face. The same kind of adjustments also identifies the effect that years of schooling have on male and female housework productivity.

Once the parameters of the utility function $u(c, l. q)$ and the household production function have been identified, the utilities $\left(U_m, U_f\right)$ which men and women obtain from the consumption of goods can be inferred from the labor supply of the husband and wife in a given couple. This follows since the reduced form labor supply equations for married men and women are monotonic functions of the husband and wife resources obtained in marriage, which can be inverted to obtain the pair of utilities $\left(U_m, U_f\right)$ up to a positive monotonic transformation.

For all married couples finally, the stability of the observed marriage matching makes it possible to derive a set of inequalities that have to hold as necessary conditions in equilibrium if no pairs of married men and women are to be willing to block the observed equilibrium matching. These sets of inequalities together with the choices that men and women make when marrying a particular spouse, parametrically identify the preferences that men and women have over marriage matches (see Sørensen 2007 for a similar argument).

6) Estimation

Brief overview of method

Due to the rich heterogeneity of male and female types in the marriage market, I am unable to derive analytical expressions for many expected outcomes of the marriage model that could form the basis of a generalized method of moments estimator. Instead, I rely on simulation techniques and estimate the marriage matching model with the method of simulated moments (McFadden 1989, and Pakes and Pollard 1989)[12].

For a brief overview of the estimation method, let Y denote a vector of endogenous variables, let X denote a vector of exogenous variables, let θ be a vector of b parameters to be estimated that appear in the conditional density $f(y|x,\theta)$ of Y given X, and finally, let (y_n, x_n) be a random sample of $n = 1,2,\ldots,N$ observations from this density.

When a set of $d \geq b$ moment functions $g(y,x,\theta)$ are available, and when the expectation of these functions given x is equal to zero if and only if the function g is evaluated at the true parameter vector $\theta = \theta_0$, then the parameter vector θ can be estimated consistently with the generalized method of moments estimator

$$\hat{\theta}_{GMM} = \left(\frac{1}{N}\cdot\sum_{i=1,N} g(y_i,x_i,\theta)\right)' \cdot \Omega \cdot \left(\frac{1}{N}\cdot\sum_{i=1,N} g(y_i,x_i,\theta)\right)$$

for some positive definite weighting matrix Ω.

In the case of the marriage model that is estimated in this paper, the function $g(y,x,\theta)$ cannot be expressed analytically because it involves expectations that are analytically intractable. However, the structural marriage model can be simulated to form random draws from the

[12] A standard likelihood approach would require the simultaneous evaluation of high-dimensional integral over the unobserved random utility terms of all agents in a market. This is not feasible from a computational point of view.

conditional density $f(y|x,\theta)$ of Y given X. In such a situation, the method of simulated moments proceeds by replacing intractable expectations in moment conditions, with approximations of these expectations that have been formed as sample averages across a large number of simulations with the structural marriage model.

For example, I lack an analytical expression for the expected marital outcome of a man or a woman with a given set of observable characteristics. To form an approximation of this expectation, I simulate the marriage model a large number of times, form the fraction of simulations in which the man or the woman turns out to be married, and use this fraction in place of the expected marital status. One of the moment conditions employed then sets the difference between the observed and predicted marital status of men and women equal to zero.

More generally, if $\hat{g}(y,x,\theta)$ is the moment function that has been formed by replacing analytically intractable parts with averages across a large number r_N of Monte Carlo simulations, then the simulated method of moments estimator of θ is equal to

$$\hat{\theta}_{SMM} = \left(\frac{1}{N} \cdot \sum_{i=1,N} \hat{g}(y_i, x_i, \theta)\right)' \cdot \Omega \cdot \left(\frac{1}{N} \cdot \sum_{i=1,N} \hat{g}(y_i, x_i, \theta)\right)$$

Under suitable regularity conditions, the parameter estimates obtained from such a procedure are consistent as the ratio of the square root of the number of observations, over the number of simulations goes to zero:

$$\frac{\sqrt{N}}{r_N} \to 0$$

In the current version of this paper, I present estimates of θ by weighting all moment conditions equally (that is, by setting the weighting matrix Ω equal to the unit matrix.)

Computation of equilibrium

The structure of the estimated model does not rule out the possibility of corner allocations where one or both of the spouses in a married couple chooses not to work in the labor market. When I compute the stable matching of the marriage model, I do allow for corner allocations where married women choose not to devote any time to market work. Since middle age couples where only the wife is working are extremely rare though, I disregard efficient marriages which involve the husband not working in the labor market[13]. For the same reason, I simply assume an interior allocation for single agents in which they work in the labor market.

For each potential couple, I allow for ten permitted allocations with a different division of resources between the husband and the wife[14]. I first compute the maximum utility that a wife can obtain in marriage while the husband receives at least his reservation utility. I then partition the interval of permitted offers from men to women so that these offers end up being spread with an equal dollar equivalent distance between them, over the interval from a woman's reservation utility to the maximum utility that she can receive in marriage.

To compute a stable matching in a market is time consuming because it involves solving a large number of numerical optimization problems for all conceivable couples (not only those that end up being married). In this version of the paper, I form simulated moments by generating $r_N = 300$ Monte Carlo simulations with the marriage model. More information about the details of these computations can be found in Appendix D.

Pre-estimation of wage equations

In a first stage, I estimate the parameters of the male and female wage equations by pooling all observations that are available in the panel structure of the NLSY data. I add a set of region dummies and year dummies to the specification of the Mincer equations that were presented in the model section above, and estimate these equations with wage data for white men and women in the representative part of the NLSY data who are 30 years or older.

[13] At the estimated parameters, less than 2% of all potential marriages have efficient allocations in which the husband would not be supplying time to market work.

[14] I also experimented with more permitted allocations per couple, but such a change had no impact on the result.

For women, I first impute wages offers for married women who are not working by using the two-step method of Heckman (1974). I then run two separate OLS regressions on the set of all single and married men and women. The results from these two regressions are presented in tables 2 and 3 below. As can be seen in these tables, men earn a return of about 6.4% on their investment in the labor market when controlling for social background and IQ scores, while the corresponding figure for women is 6.7%.

Since the imputed log hourly wages in the NLSY data for men and women are likely to include big measurement error components, I follow Flinn and Del Boca (2006) and reduce the variance of log hourly wages by half in the estimation of the marriage model (the size of the reduction is motivated by the evidence from Bound et al (1994), who compare self-reported and “true” hourly wages for factory workers and find that half of the variance of reported log hourly wages is due to measurement error).

Moment conditions based on time allocations

To estimate the parameters of the matching model, I use a number of moments based on the time allocations of men and women. For market work and housework, I use ten moments based on the mean time supplied to these two activities for single agents, married men with working wives, married men with non-working wives, married working women, and married non-working women. In all of these moments, I include demographic controls for the number of children between ages zero and four in each household, the number of children of ages five to seventeen, and a dummy for home ownership.

As the household production function is identified with data on the ratio of male to female housework, I also include two moments for the housework of husbands and wives in couples where both of these household members supply time to market work. For that purpose, I rewrite the expression for the ratio of male to female housework in these couples in regression form, so that it reads

$$log\left(\frac{h_m}{h_f}\right) = \left(\frac{1}{v}\right) \cdot log\left(\frac{w_m \cdot h_m}{w_f \cdot h_f}\right) + \gamma \cdot \left(s_f - s_m\right) + \pi \cdot \left(v_f - v_m\right) + \ X \cdot \Psi_{mf,h} + \varepsilon_{mf,h}$$

Since housework appears on both sides of this equality, and wages and housework are likely to be measured with error, I instrument for the first independent variable by using the lagged annual earnings of the respondent[15]. I then form two moments based on the orthogonality of the error term $\varepsilon_{mf,h}$ in the equation above on the one hand, and the lagged earnings variable and the difference between the years of schooling of the wife and the husband, on the other hand.

I also include two additional moments based on the labor force participation of women. The first of these has a moment function equal to the difference between the observed and predicted labor force status of each agent, and the second moment function equals the product of an agent's years of schooling, and the difference between his or her observed and expected labor force participation.

Moment conditions for marital behavior

To capture the marital behavior of men and women in the estimated model, I include a moment based on the difference between the observed marital status of an agent, and his or her expected marital status. I also include a moment equal to the product of an agent's years of schooling, and the difference between his or her observed and expected marital status. A third moment condition is based on the amount of sorting on years of schooling in marriage. I introduce a function that is equal to the product of the deviation of the wife's years of schooling from the mean and the deviation of the husband's years of schooling from the mean if the agent is married, and equal to zero otherwise. I then form a moment condition equal to the difference between the observed and predicted values of this function.

7) Patterns of sorting in data

This section presents the general patterns of sorting on years of schooling and hourly wages in the US marriage market. I use data from the 1980 census, the representative sample of the National Longitudinal Survey of Youth 1979 (NLSY79), and the March supplement of the

[15] This variable was not used to compute the hourly wages of men and women.

Current Population Survey (CPS). All of the correlations are summarized in Table 4 below, and are computed for white households[16].

I first compute the correlation for years of schooling in the 1980 census and for households where the wife is 40 years old. In a sample of over 40000 households, the correlation is 0.62. The reported years of schooling in the census data are likely to contain measurement error. To correct for such errors, I use the NLSY79 and regress years of husband schooling on wife schooling and vice versa, first by OLS and then by using parental schooling as an instrument for years of schooling. If measurement errors are classical and parental schooling is a valid instrument, then the ratio of OLS to IV estimates provides a consistent estimate of the noise to signal ratio in reported husband and wife years of schooling. With no adjustment for measurement errors in the NLSY79 data, the correlation for years of schooling in marriage is 0.59. When this correlation is adjusted upwards in accordance with the estimated noise to signal ratio in schooling, the new correlation for husband and wife schooling is 0.75.

To examine correlations for husband and wife hourly wages, I start by working with households from the 1980 census in which the wife is 40 years old, and in which the hourly wage rates of the wife and the husband are between 50 cents and 100 USD. The correlation for hourly wages in this group of households is 0.14. I then try to correct for the presence of measurement errors in wages by using the panel structure of the NLSY79. Men and women are likely to sort on permanent rather than temporary wages, and sorting on permanent wages is more important from a welfare point of view since permanent income determines consumption levels.

For all white men and women above age 24 in the NLSY79, I run gender specific regressions of log hourly wages on a constant, the age and age squared of the respondent, and year dummies X:

$$log\left(w_n^{i,t}\right) = \alpha_n^0 + \alpha_n^1 \cdot age^{i,t} + \alpha_n^1 \cdot age2^{i,t} + \delta_n \cdot X + e_n^{i,t} \quad , \quad n = m, f$$

After this step, I construct the log of permanent wages for these men and women in 2000 as the predicted log wage in that year, plus the mean of the error terms over all available respondent observations. I then compute correlations for wages in marriage in the cases where a wage

[16] Household in which the wife is labeled as white.

observation is available for the wife in 2000, and a wage observation is available for the husband in that year, plus in four other years of the NLSY survey.[17]

When I use permanent wages, the correlation for husband and wife wages in the NLSY79 is 0.42. To determine how much the reduction of measurement errors increase the reported correlation, I also compute the correlation for wages with the exact same individuals, but only using the pair of wage observations from 2000 for the husband and the wife. The correlation for husband and wife wages is then 0.33.

Since these two correlations for wages from the NLSY79 are considerably higher than the correlations that I found in the 1980 census data, I use the March supplement of the CPS in each five year interval from 1980 to 2000, to determine if there has been a clear upward trend in sorting on wages in the US over the recent decades. In each CPS survey, I use households with white women in the ages of 36 to 45, and compute correlations for hourly wages in marriage. As can be seen in Table 4, there appears to be no clear upward trend in these numbers. More likely, the higher correlations for wages that I obtain in the NLSY79 reflect a higher quality of the data, smaller measurement errors, and thus, higher correlations for husband and wife wages.

To summarize the findings in this section, the correlation of husband and wife schooling in marriage is fairly stable across different data sets. With adjustments for measurement errors, it is in the order of 0.75. The correlation for hourly wages differs more between data sets, possibly reflecting differences in the quality of the data. When I adjust for measurement errors, this correlation appears to be in the order of 0.3 - 0.4.

8) Results

Overall estimation results

In Table 5, I display the values of the estimated preference and household production parameters in the model. Men and women who obtain an additional year of schooling are estimated to

[17] I do not require any minimum number of observations for women since it would select on women with a strong attachment to the labor force. On average, there are 11 observations per husband used to compute the correlation on permanent wages, and 4 observations per wife.

increase their productivity in housework by about 7 %. The household production parameter v is estimated to be 0.57, which implies that the elasticity of substitution for male and female housework is in the order of 2.33. The estimated value of v also implies that a married couple produces 60% more of the household good than what the husband and the wife would do by themselves, if they supplied the same amount of time to housework as singles. This degree of complementarity for male and female housework is large, but does not appear to be unrealistic.

Table 6 shows a summary of the observed and predicted levels of sorting in marriage. As can be seen in the table, the predicted covariance of years of schooling in marriage, is very close to the observed covariance in the data. The correlation for years of schooling in the model is in turn 0.68, and the correlation for hourly wages is 0.30. These numbers are close, and perhaps a little bit lower than the numbers that are found in the data and that were summarized above. Finally, the degree of sorting on non-labor income in the model is 0.15.

In Table 7, I display the observed time allocations for three different types of households, which exhibit the well known pattern of specialization as a function of comparative advantage. In couples where both the husband and the wife are working, husbands specialize towards market work and perform less housework than their single counterparts. Their wives on the other hand, perform less market work and more housework than their single counterparts. In couples where only the husband works, the patterns of specialization are even stronger, so that the husbands perform the highest amount of market work across all household types, and the wives perform the highest amount of housework.

In the bottom part of Table 7, I also display the predicted patterns of time allocations from the model. These data exhibit the same type of specialization that can be observed in the data, and the model performs fairly well except in the case of husbands whose wives are not working, and for whom the predicted amount of housework is too low.

In Figures 2 and 3, I show the observed and predicted fractions of men and women who are married as a function of their years of schooling. In the data, more schooling raises marriage rates for both men and women, and these overall patterns are replicated by the model. Overall, the fraction of men married is 71% in the model, and close to the overall fraction of men married in the data which is 72%. Finally, in Figure 4, I compare the fraction of married women who are working in the data and in the model. In both cases, more educated women are more likely to working in the labor market.

Counterfactuals

In the estimated model, the years of schooling of men and women affect their marital experiences through their frequency of marriage, the characteristics of their spouses (due to positive assortative matching), and the share of resources that they obtain in marriage. To construct counterfactual marital outcomes that capture all three of these effects, I introduce an additional imaginary man and woman in each market, vary their years of schooling across a large number of simulations with the model, and record their consumption outcomes.

In these simulations, I assign wages to the imaginary agents by using the estimated Mincer equations, adjusting the experience and years of schooling of the imaginary agents, and setting the value of the error terms in these equations equal to zero. Furthermore, I assign housework productivity to the imaginary agents in accordance with their years of schooling, give them the mean level of non-labor income, IQ, and social background by gender across all cohorts, and draw random utility terms over spouses for them from the estimated distributions in the model. With these settings, the outcomes for the imaginary agents across simulations only depend on their years of schooling, since all other of their characteristics are being held constant.

To illustrate how I use these imaginary agents to construct counterfactuals, I consider the case of an imaginary man who increases his education from s_A to s_B years of schooling. I first simulate the model S times at both these levels of schooling for the imaginary man, record his utility from the consumption of the market good, leisure, and the household good in each simulation, and form the mean utilities u_A and u_B for him across these two sets of simulations

$$u_A = \frac{1}{S} \cdot \sum_{t=1,S} u_A^t \quad , \quad u_B = \frac{1}{S} \cdot \sum_{t=1,S} u_B^t$$

I then perform a third set of counterfactual simulations in which I assign the imaginary man s_B years of schooling and increase his wage and housework productivity in accordance with this new level of schooling, but hold the quality of his marital experience fixed. I explain this procedure in more detail by separately considering the simulations in which the imaginary man ends up being single with s_A years of schooling, and the simulations in which he ends up being married.

In the cases where the imaginary man ends up being single, his utility with s_A years of schooling is given by

$$V_o(w(s_A), s_A, y)$$

In this expression, V_0 is the value function for single agents that was defined above, and $w(s)$ is a function meant to represent the relationship between wages and years of schooling that is explicit in the Mincer equation. For these cases, I construct the counterfactual outcome of the imaginary agent by increasing his wage and housework productivity, so that his counterfactual utility becomes

$$V_o(w(s_B), s_B, y)$$

In the case where the imaginary man ends up being married with s_A years of schooling, I construct a new counterfactual marriage for him in which he has s_B years of schooling, receives a higher wage and is more productive in housework, but remains married to the same wife as when he had s_A years of schooling. I furthermore divide the resources in the new counterfactual marriage so that the ratio of the dollar value of the husband's and the wife's total consumption of the market good, leisure, and the household good in the counterfactual marriage, is the same as the corresponding ratio in the original marriage when the imaginary man had s_A years of schooling.

This procedure is illustrated in Figure 6 below, in which I represent the original marriage of the imaginary man when he has s_A years of schooling with the point A_A along the utility possibility frontier of UPF_A. I also include a typical marriage for this man when he has s_B years of schooling and obtains the allocation represented by the point A_B along the utility possibility frontier of UPF_B. Finally, the figure also contains the counterfactual marriage described above, which is illustrated by the point A_{CF} along the intermediate counterfactual utility possibility frontier of UPF_{CF}.

With this complete set of counterfactual outcomes for the imaginary man that cover both the cases in which he was single and married with s_A years of schooling, I can form his mean utility from the consumption of the market good, leisure, and the household good across all counterfactual simulations as

$$u_{CF} = \frac{1}{S} \cdot \sum_{t=1,S} u_{CF}^{t}$$

I then define and compute the share of the consumption difference for the imaginary man that is caused by an improvement in his marital experience as he increases his years of schooling from s_A to s_B, as

$$share = 1 - \frac{\Lambda(u_{CF}) - \Lambda(\mathrm{u_A})}{\Lambda(\mathrm{u_B}) - \Lambda(\mathrm{u_A})}$$

where $\Lambda(u)$ is the inverse of the expenditure function that converts utilities into dollars. In words, this share is the total difference between the consumption of the man when he has s_A and s_B years of schooling, minus the share of this consumption difference that would be present if the man increased his education but experienced no improvement in his marriage outcomes.

The procedure above gives me the share of consumption differences that is due to improvements in marital outcomes for one man in one cohort market as he increases his education from s_A to s_B years of schooling. I repeat this set of simulations for all the imaginary men and women in the model, and change their educational attainment in increments of one year of schooling, from 8 to 9, 9 to 10, 10 to 11, and so on, up to 17 to 18 years of schooling. I then take the median[18] of all these estimated shares for men and women respectively, and use them as my overall share estimate.

In Table 8 below, I report the end result of all these simulations. In the case of men, the table indicates that improvements in marital outcomes generate 18% of the differences in consumption between men at different levels of schooling, and that for women, the corresponding figure is 64%.

[18] I work with medians rather than means, to minimize the effects of a few outliers with extreme estimated shares.

Alternative counterfactuals

As I argued in the introduction to this paper, the estimates that I just presented are derived from a model which contains many features that appear to be relevant when analyzing marriage, such as the value of housework, the division of labor between spouses, and the way in which husbands and wives share resources in marriage. On the other hand, the inclusion of these effects causes many difficult measurement problems. For this reason, I also use the estimated model to simulate the same type of counterfactual outcomes as above, but with alternative consumption measures for the imaginary agents which can be constructed from variables that are easier to quantify.

The first of these alternative measures has previously been used by Goldin (1992), and Lefgren and McIntyre (2006), and is constructed from earnings and non-labor income. With the notation of this paper, the alternative consumption measure for single agents $\widehat{c_0}$ is

$$\hat{c}_o = y_o + w_o \cdot (1 - h_o - l_o)$$

and for married men and women, the measure is half of total household income

$$\hat{c}_m = \hat{c}_f = \frac{1}{2} \cdot \left(y_m + y_f + w_m \cdot (1 - h_m - l_m) + w_f \cdot \left(1 - h_f - l_f\right)\right)$$

In Table 8, I present the results when this alternative measure is used in place of the dollar value of the total consumption of the market good, leisure, and the household good. As is shown in the table, the share of improved marital outcomes in total consumption differences is now estimated to be 17% for men and 63% for women.

In addition, I also make a slight change and include consumption savings for married agents that are given by a square root formula. The consumption of husbands and wives is then equal to

$$\hat{c}_m = \hat{c}_f = \frac{1}{2} \cdot \left(\frac{1}{p}\right) \cdot \left(y_m + y_f + w_m \cdot (1 - h_m - l_m) + w_f \cdot \left(1 - h_f - l_f\right)\right)$$

The shares of improved marital outcomes in the consumption differences for men and women do not change much, and are now 14% and 68% respectively.

Finally, I also use a consumption measure for men and women that is based on full income. The wage rates used to construct this measure come from the estimated Mincer equations of the model, and as was pointed out above, the estimated wages for women were produced from a model that controls for selection in and out of the labor force. For singles, the consumption measure is equal to hourly wages

$$\hat{c}_o = w_o$$

and for married agents, the consumption of the husband and the wife is equal to half of their full income

$$\hat{c}_m = \hat{c}_f = \frac{1}{2} \cdot (w_m + w_f)$$

The last row of Table 8 presents the result with a full income measure, and indicates that the share of improved marital outcomes in the total consumption differences for men is 21%, while the corresponding share for women is 57%.

9) Conclusion

In this paper, I have provided estimates of the share of returns to education that is realized through marriage. At middle age, US women of the 1960 generation appear to have earned in the order of 65% of their return to schooling through marriage, whereas the number for men was in the order of 20%.

I conclude with a brief discussion of two of the ways in which the estimates presented in this paper can be extended and improved. First of all, returns to education are captured over the entire life cycle, whereas this paper only considered the payoffs to men and women around middle age. As more educated agents delay their marriage relative to their less educated counterparts, this potentially implies that the numbers presented in this paper are biased upwards.

Another limitation is that the model estimated in this paper does not consider the possibility of divorce and remarriage. More educated agents have slightly more stable marriages (see for example Lefgren and McIntyre 2006) which generates a bias in the opposite direction. Which of these two effects dominate is difficult to determine without a more detailed life cycle analysis.

APPENDIX A. Algorithm for Computing Stable Equilibria

In a given market k, there are a total of M_k men indexed by $i = 1,2, \dots, M_k$, and a total of F_k women indexed by $j = 1,2, \dots, F_k$. A man i and a woman j can obtain distinct permitted utilities from the consumption of commodities that are in the set

$$\Pi_u^{i,j,k} = \left\{ \left(U_{m,a}^{i,j,k}, U_{f,a}^{i,j,k} \right) \right\}_{a=1,A}$$

Each husband utility $U_{m,a}^{i,j,k}$ is the value of the Pareto problem (P) specified above, given that the wife obtains utility $U_{f,a}^{i,j,k}$ in marriage from the consumption of commodities. The utilities of the husband are ordered so that they are strictly decreasing.

$$U_{m,1}^{i,j,k} > U_{m,2}^{i,j,k} > \cdots > U_{m,A}^{i,j,k}$$

and the utilities of the wife are strictly increasing. To each of these permitted utilities from the consumption of commodities, one can add the utility of the marriage match itself to the husband and the wife, to obtain the set

$$\Pi_v^{i,j,k} = \left\{ \left(V_{m,a}^{i,j,k}, V_{f,a}^{i,j,k} \right) \right\}_{a=1,A}$$

of total permitted utilities in marriage. Ultimately, men and women care only about these total utilities in marriage when making decisions about if they should remain single or get married, and if so, with whom.

An algorithm for constructing a stable matching proceeds in discrete stages $t = 0,1,2, \dots$ In any stage, a man i in market k is permitted to propose a marriage to woman j in that market with a division of utility that is indexed by

$$d^{i,j,k}(t) \in \{1,2, \dots, A\}.$$

Since no marriage will ever take place between a husband and the wife for whom there is no marital surplus, I simply disregard such couples in what follows. For the couples that can benefit from marriage, I assume that the first division of resources $\left(V_{m,1}^{i,j,k}, V_{f,1}^{i,j,k}\right)$ assigns a utility to the wife so that her total utility is equal to her utility as single

$$V_{f,1}^{i,j,k} = V_o^{j,k}$$

I also assume that the permitted divisions of utility for a couple contain at least one division for which the husband prefers to be single, rather than to marry under that division of utility.

The Crawford and Knoer algorithm then proceeds as follows[19]:

R1. The permitted offers from men to women give women their utility as single. With the notation used above, $d^{i,j,k}(0) = 1\ ,\ \forall\ i, j, k$. Unless otherwise noted below, $d^{i,j,k}(t)$ is constant.

R2. Each man initially chooses the better of two alternatives: to makes an offer to his favorite woman given the schedule of permitted divisions of utility $\left[\, d^{i,j,k}(t)\right]$, or to remain single.

R3. Each woman who receives one or more offers chooses the better of two alternatives: either to remain single, or reject all but her favorite offer which she tentatively accepts.

R4. Offers not rejected in previous stages remain in force. If woman j in market k rejected an offer from man i in that market in stage $t-1$, then

$$d^{i,j,k}(t) = d^{i,j,k}(t-1) + 1$$

If not,

$$d^{i,j,k}(t) = d^{i,j,k}(t-1)$$

Rejected men continue by once again choosing the better of two alternatives: to remain single, or to make offers to their favorite women, taking into account their current permitted offers.

R5. The process stops when no rejections are issued in some stage. Women then accept the offers that remain in force from the men they have not rejected, and single men and women remain single.

---

[19] This exposition is a slight modification of the presentation found in Crawford and Knoer (1981).

Appendix B. Breakdown of Time Use into Major Categories

Work			
1	normal work/work at home	2	job search, unemployment benefits, welfare, food stamps
5	second job	6	lunch at workplace
7	before and/or after work, other work related	8	coffee breaks at work place
9	travel to work, travel for job search, etc		
Housework			
10	meal preparation	11	meal cleanup
12	indoor cleaning	13	outdoor cleaning
14	laundry	16	repairs, maintenance
17	care of house plants	19	other indoor housework
20	baby care	21	child care
22	helping children	23	reading to/help with homework
24	indoor playing with children	25	outdoor playing
26	medical care of children	27	other child care
29	travel in connection with child care	30	shopping for all other than household goods
31	shopping for household goods	34	obtaining financial services/other government services
35	obtaining household services	37	obtaining other services
38	errands	39	travel related to obtaining services
Personal Needs and Care			
32	personal care (hair dresser, beauty salons etc)	33	medical care
40	washing and dressing	41	medical care to self and others in household (not children)
42	non-medical care to others	43	meals at home
44	meals away from home	45	night sleep
46	naps and resting	49	travel related to care to others
Time Unaccounted For			
48	unaccounted time		
Leisure			
50	full time studies	51	other classes
54	homework	56	other education
59	travel related to studies	60	professional organizations
61	special interests	62	political activities
63	organizational work	64	religious groups
65	religious activities	66	fraternal organizations
67	family organizations	68	other organizations
69	travel related to organizational activities	70	Sports
71	spectacles/events	72	Movies
73	theater	74	Museums
75	visiting with others	76	Party
77	at bars/dancing	78	other events
79	travel related to socializing	80	active sports
81	outdoors	82	walking, biking
83	hobbies	84	domestic crafts
85	art and literature	86	music/drama/dance
87	games	88	classes, other
89	travel, active leisure	90	Radio
91	TV	92	records, tapes
93	reading books	94	reading magazines
95	reading newspapers	96	conversations, phone
97	letters	98	other passive leisure
99	travel, passive leisure		

APPENDIX C. Cohort Data from the NLSY

To construct separate markets with men and women who can remain single, or marry a member of the opposite sex, I use 2000 data on white respondents from the representative part of the NLSY79. The agents who appear in the marriage markets are all respondents in the original NLSY data set, but I also use data for their spouses to construct some moments used in the estimation process.

For the respondents, all the required data is available, except for data on their housework. The AFQT scores used in this paper are the cohort and gender adjusted percentiles of the raw total test scores. For social background, I estimate a single principal factor model with the data that is available on the years of schooling of the respondent's father and mother, and the earnings of these two parents (imputed from their occupations) which I denote by θ in the text above.

For the spouses of the respondents, there is no information on AFQT scores and social background. I therefore impute values for these two variables by forming cells based on the education, gender, occupation, and spouse education of all the respondents in the representative part of the NLSY survey. In the imputation, I use both the mean and the variance of these traits for each cell.

To make sure that each cohort sample is representative of the overall US population, I compute sex ratios, fractions of men and women married, and the educational distribution of the entire cohort of white men and women in the 5% sample of the 2000 Census. Each separate cohort based marriage market is then constructed by drawing a number of single and married men and women to replicate these summary statistics.

Single agents and married men were included in the constructed sample if they worked last year. Married women were included regardless of their labor force status. All the included men and women in the constructed data set were either single agents who lived alone, or married agents (no cohabiting men and women were included).

APPENDIX D. Computation of Equilibrium

The equilibrium of the model was computed through the following steps:

1) The utility of being single was computed for all men and women.

2) The maximum utility that a woman could obtain in every marriage while the husband received his reservation utility as single was computed. Both allocations where the wife did and did not work were considered.

3) The maximum utility that a man could obtain in every marriage while the wife received her reservation utility as single was computed. Both allocations where the wife did and did not work were considered.

4) For marriages with a surplus, the interval between the highest obtainable utility for the wife and her reservation utility as single, was partitioned with 10 equispaced points in dollar equivalent terms.

5) For each of these ten utility values, the maximum utility that could be obtained by the husband was computed as the maximum over allocations where the wife did and did not work.

6) The matching algorithm from Appendix A was applied to the men and women in a given market and the values along the utility possibility frontier that were computed in the five steps above.

Due to Inada conditions, any allocation that satisfies the first order conditions of the overall household Pareto problem (P) satisfies the non-negativity constraints for male and female housework and leisure. If the non-negativity constraint for male market work was binding in a couple, then a marriage between the husband and the wife was ruled out, that is, I discarded efficient allocations in which only the wife would have supplied time to market work. Such allocations appeared in less than 2% of all potential marriages in the model at the estimated parameters.

To experiment with the model, I also computed stable matchings when the utility possibility frontier of a couple was split into 20 and 30 allocations respectively. This had no impact on the results. To minimize the computational time, I therefore estimated the model with only 10 allocations along the utility possibility frontier of each couple.

REFERENCES

[] Becker, G. S., 1960, “Investing in Education and Research: Underinvestment in College Education?”, American Economic Review, 50(2), pp 346-354

[] Becker, G. S., 1973, “A Theory of Marriage: Part I”, Journal of Political Economy, 81(4), pp 813-846

[] Bound, J., Brown, C., Duncan, G., Rodgers, W., 1994, “Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data”, Journal of Labor Economics”, 12(3), pp 345-68

[] Chiappori, P-A., 1988, “Rational Household Labor Supply”, Econometrica, 56(1), pp 63-90

[] Chiappori, P-A., Fortin, B., and Lacroix, G., 2002, “Marriage Market, Divorce Legislation, and Household Labor Supply”, Journal of Political Economy, 110(1), pp 37-72

[] Choo, E., and Siow, A., 2006, “Who Marries Whom and Why?”, Journal of Political Economy, 114(1), pp 175-201

[] Choo, E., Seitz, S., and Siow, A., 2008, “The Collective Marriage Matching Model: Identification, Estimation, and Testing”, Working Paper

[] Crawford, V. P., and Knoer, E. M., 1981, “Job Matching with Heterogeneous Firms and Workers”, Econometrica, 49(2), pp 437-450

[] Del Boca, D., and Flinn, C. J., 2006, “Household Time Allocation and Modes of Behavior: A Theory of Sorts”, Working Paper

[] Ehrlich, I., Hamlen, W., and Yong, Y., 2008, “Asset Management, Human Capital, and the Market for Risky Assets”, Journal of Human Capital, 2(3), pp 217-261

[] Elias, J., 2005, “The Effects of Ability and Family Background on Non-monetary Returns to Education”, Ph.D. Dissertation, Department of Economics, University of Chicago

[] Gale, D., and Shapley, L., 1962, “College Admissions and the Stability of Marriage”, American Mathematical Monthly, 69, pp 9-15

[] Ge, S., 2007, “Women’s College Decisions: How Much Does Marriage Matter”, Working Paper

[] Goldin, C., 1992, “The Meaning of College in the Lives of American Women: The Past Hundred Years”, NBER working paper no. 4099

[] Heckman, J., 1974, “Shadow Prices, Market Wages, and Labor Supply”, Econometrica, 42(4), pp 679-694

[] Kenny, L., 1983, “The Accumulation of Human Capital during Marriage by Males”, Economic Inquiry 21, pp 223-31

[] Lafortune, J., 2008, “Making Yourself Attractive: Pre-Marital Investments and the Returns to Education in the Marriage Market”, Working Paper

[] Lefgren, L. and McIntyre, F., 2006, “Examining the Relationship between Women’s Education and Marriage Outcomes”, Journal of Labor Economics, 24(4), pp 787-830

[] McFadden, D., 1989, “A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration”, Econometrica, 57(5), pp 995-1026

[] Pakes, A., and Pollard, D., 1989, “Simulation and the Asymptotics of Optimization Estimators”, Econometrica, 57(5), pp 1027-1057

[] Pollak, R. A., and Wachter, M. L., 1975, “The Relevance of the Household Production Function and Its Implications for the Allocation of Time”, Journal of Political Economy, 83(2), pp 255-277

[] Sørensen, M., 2007, “How Smart Is Smart Money? A Two-Sided Matching Model of Venture Capital”, Journal of Finance, 62(6), pp 2725-2762

Table 1: Cohort Based Marriage Markets in the NLSY Data

Market	Male cohort	Female cohort	Male age	Female age	# Men	# Women
1	1957	1959	43	41	79	81
2	1958	1960	42	40	79	81
3	1959	1961	41	39	79	81
4	1960	1962	40	38	81	79
5	1961	1963	39	37	81	79
6	1962	1964	38	36	81	79

Notes: Composition of men and women in the six constructed marriage markets in the NLSY data.

Table 2: Male Wage Equation

Dependent Variable	Log male wages
Years of schooling	0.064
	(0.003)
Social background	0.069
	(0.008)
AFQT score	0.249
	(0.026)
Experience	0.063
	(0.007)
Experience squared	-0.001
	(2.5E-0.4)
North East Region	0.095
	(0.017)
North Central Region	-0.029
	(0.015)
South Region	-0.020
	(0.016)
Constant	1.151
	(0.069)
Year dummies 1988-2002	
# Observations	12598
R-squared	0.22

Notes: Regression of log male wages on male characteristics with pooled data over period of 1988 to 2002. Numbers in parentheses are standard errors.

Table 3. Female Wage Equation

Dependent variable	Log female wages
Years of schooling	0.067
	(0.001)
Social background	0.022
	(0.004)
AFQT score	0.286
	(0.012)
Experience	0.073
	(0.002)
Experience squared	-0.001
	(8.3E-05)
North East Region	0.076
	(0.008)
North Central Region	-0.111
	(0.007)
South Region	-0.013
	(0.007)
Constant	0.758
	(0.023)
Year dummies 1988-2002	
# Observations	15069
R-squared	0.66

Notes: Second stage regression of log female wages on female characteristics with pooled data over period of 1988 to 2002. Numbers in parentheses are standard errors.

Table 4: Observed Sorting on Schooling and Wages

	Correlation	Observations
Years of schooling		
1980 Census	0.62	42,251
Matching data set	0.64	620
NLSY79	0.59	1,901
NLSY79, IV	0.75	1,901
Hourly Wages		
1980 Census	0.14	22,647
Matching Data set	0.055	383
NLSY79, annual data	0.33	1,532
NLSY79, permanent wages	0.42	1,532
CPS 80	0.1	3,007
CPS 85	0.15	3,070
CPS 90	0.093	3,820
CPS 95	0.016	4,607
CPS 2000	0.12	4,228

Notes: The samples used to compute these correlations are described in detail in section 7.

Table 5: Estimated Preference and Production Parameters

Preferences over goods		
0.45	(0.056)	(consumption weight in preferences)
0.38	(0.071)	(leisure weight in preferences)
Household technology		
0.071	(0.919)	(schooling and housework productivity)
0.57	(0.508)	(CES parameter housework)

Notes: Estimated coefficients, and standard errors in parentheses.

Table 6: Observed and Predicted Levels of Sorting

	Observed Sorting	Predicted Sorting
Covariance Years of Schooling	6.10	6.07
Correlation Years of Schooling	0.75	0.68
Correlation Hourly Wages	0.30	0.30-0.40
Correlation Non-Labor Income	0.15	n.a.

Notes: Data on observed levels of sorting are taken from section 7 of the text. Predicted levels of sorting come from the estimated model.

Table 7: Observed and Predicted Time Allocations

	Market Work	Housework	Leisure
Observed Time Allocations			
Singles	0.45	0.16	0.39
Husbands (working wives)	0.48	0.15	0.42
Husbands (non-working wives)	0.54	0.15	0.31
Working wives	0.3	0.32	0.38
Non-working wives	0	0.45	0.55
Predicted Time Allocations			
Singles	0.42	0.18	0.4
Husbands (working wives)	0.43	0.15	0.42
Husbands (non-working wives)	0.58	0.06	0.36
Working wives	0.34	0.27	0.39
Non-working wives	0	0.49	0.51

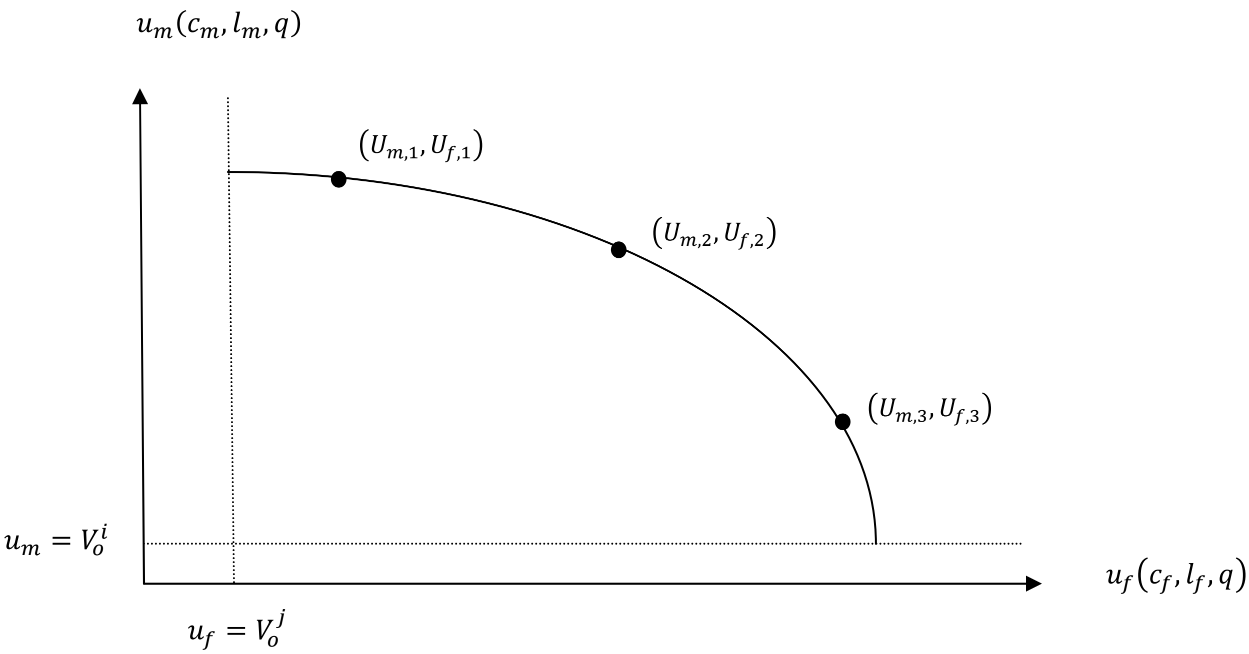
Notes: Figures refer to fraction of total time spent in the three activities of market work, housework, and leisure.

Table 8: Share of Returns to Schooling Earned through Marriage

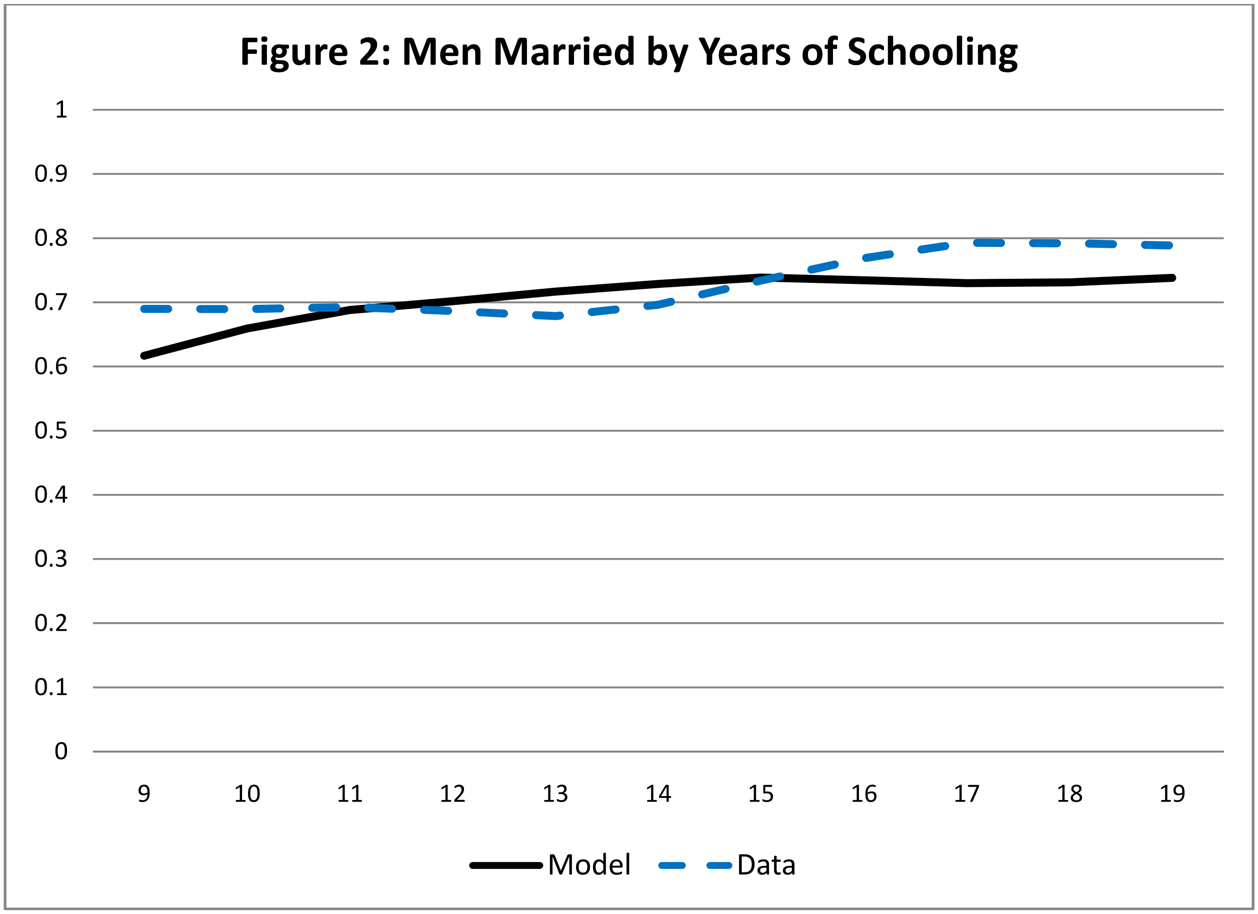
Consumption measure	Men	Women
Dollar Value of Consumption According to Model	18%	64%
Half of Earnings and Non-Labor Income (No Consumption Savings)	17%	63%
Half of Earnings and Non-Labor Income (With Consumption Savings)	14%	68%
Half of Full Income	21%	57%

Notes: Figures for men and women refer to share of total return to education earned through marriage at age 40.

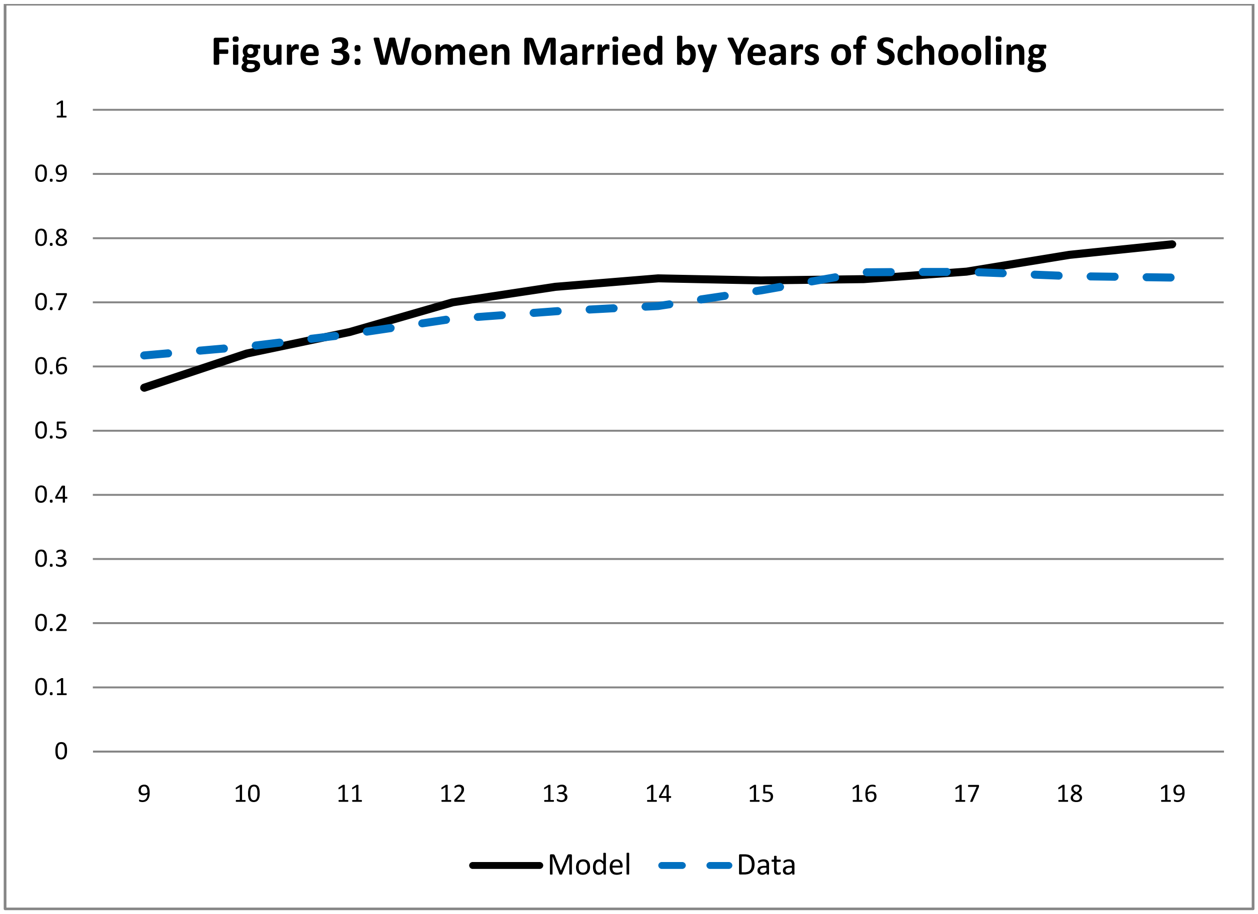
Figure 1. Stylized Utility Possibility Frontier for a Married Couple



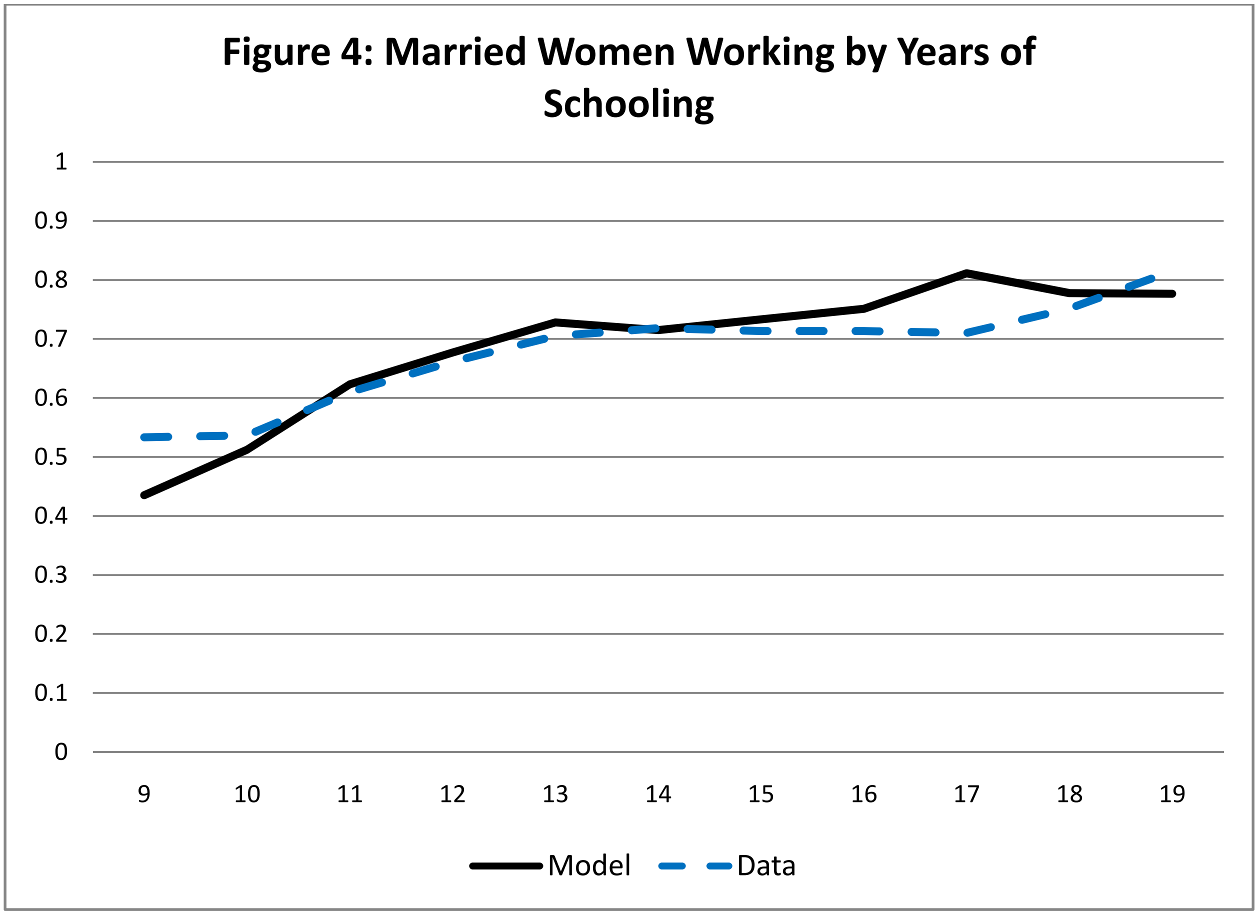
Notes: Three permitted divisions of utility along a stylized utility possibility frontier of a married couple. The utility of the husband is plotted along the y-axis, and the utility of the wife is plotted along the x-axis. Only the part of the frontier between the reservation utilities of the husband and the wife (their utilities if they choose to remain single) is plotted since no marriages will takes place outside of this area.



Notes: Fraction men married by years of schooling in model and in data (3 year moving average for both series).

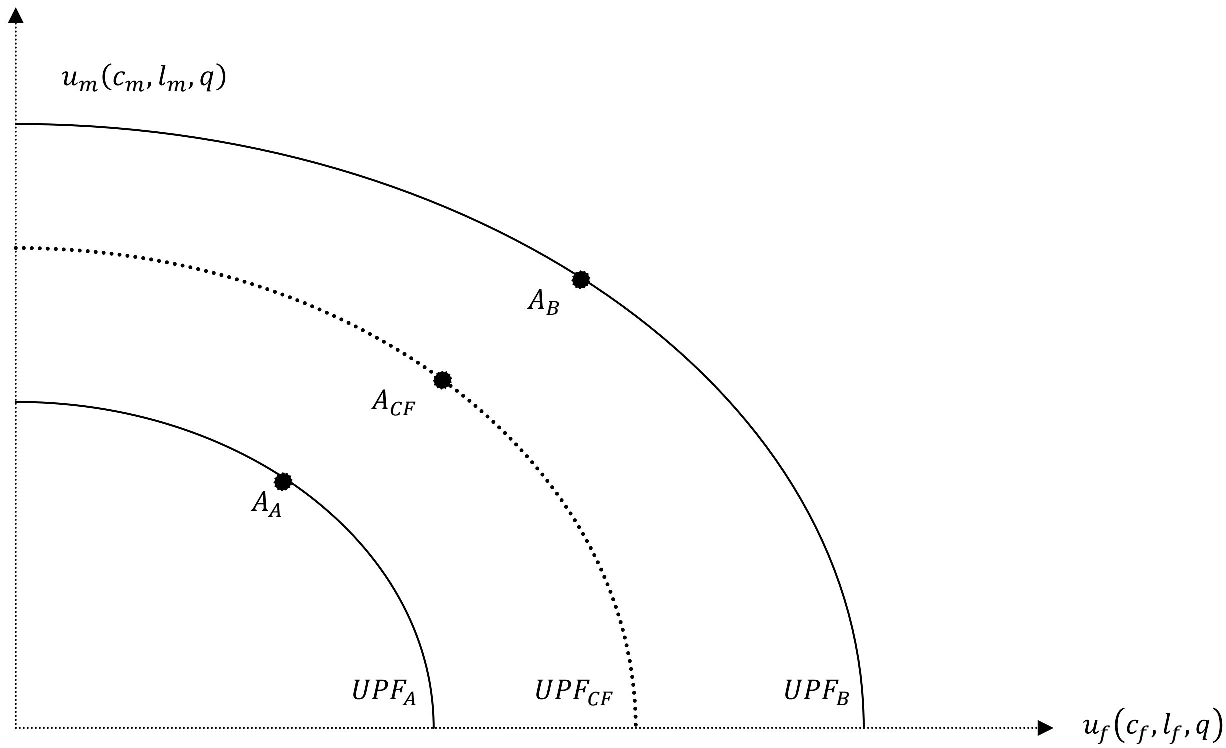


Notes: Fraction women married by years of schooling in model and in data (3 year moving average for both series).



Notes: Fraction married women working by years of schooling in model and in data (3 year moving average for both series).

Figure 5. Examples of Actual and Counterfactual Marital Outcomes



Notes: Figure shows two simulated actual marital outcomes for an imaginary man A_A and A_B at two different levels of schooling. Intermediate counterfactual marital outcome is $A_{CF}$when man obtains the higher years of schooling, but is still married to the same wife as in the allocation A_A. In the counterfactual simulations, the position of the counterfactual allocation A_{CF} along the utility possibility frontier UPF_{CF} is chosen so that the ratio of the dollar value of the husband and wife consumption in allocation A_{CF}, is the same as for the allocation A_A.